

China Guidelines for Pharmacoeconomic Evaluations

2020 Edition

Sponsored by

Chinese Pharmaceutical Association

China Society for Pharmacoeconomics and Outcomes Research

ISPOR Beijing Chapter

Dec 2020

Editor-in-Chief

Gordon G. Liu, PhD
PKU BOYA Distinguished Professor, National School of Development
Peking University
Beijing, China

Associate Editors

Shanlian Hu, PhD
Professor, School of Public Health
Fudan University
Shanghai, China

Jiuhong Wu, PhD
Professor, Department of Pharmacy
306 Hospital
Beijing, China

Jing Wu, PhD
Professor, School of Pharmaceutical Science and Technology
Tianjin University
Tianjin, China

Zhaohui Dong, PhD
Director, China Academy of Labor and Social Security
Ministry of Human Resources and Social Security
Beijing, China

Hongchao Li, PhD
Assistant Professor, School of International Pharmaceutical Business
China Pharmaceutical University
Nanjing, China

Working Group Members (by alphabetical order)

Yingyao Chen, PhD
Professor, School of Public Health
Fudan University
Shanghai, China

Haijing Guan, PhD
Senior research fellow, China Center for Health Economic Research
Peking University

Beijing, China

Sheng Han, PhD

Director of the Academic Department, International Research Center for Medicinal
Administration
Peking University
Beijing, China

Xiaoning He, PhD

Assistant Professor, School of Pharmaceutical Science and Technology
Tianjin University
Tianjin, China

Min Hu, PhD

Associate Professor, School of Public Health
Fudan University
Shanghai, China

Hong Li, PhD, MPH

Adjunct Associate Professor, James L. Winkle College of Pharmacy,
University of Cincinnati
Ohio, US

Minghui Li, PhD

Assistant Professor, College of Pharmacy
University of Tennessee
Memphis TN USA

Shunping Li, PhD

Professor, Centre for Health Management and Policy Research, School of Public Health,
Cheeloo College of Medicine
Shandong University
Jinan, China

Binyan Sui, PhD

Associate Researcher, Division of Health Policy Evaluation and Technology Assessment
China National Health Development and Research Centre
Beijing, China

Xin Sun, PhD

Professor, Center for Clinical Epidemiology and Evidence-based Medicine
West China Hospital, Sichuan University
Chengdu, China

Hongyan Wu, PhD
Professor, School of Medicine and Health Management
Guizhou Medical University
Guiyang, China

Jipan Xie, MD PhD
Vice President, Analysis Group, Inc.
Los Angeles, US

Li Yang, PhD, Msc
Professor, School of Public Health
Peking University
Beijing, China

Xiaomeng Yue, PhD, BPharm
James L. Winkle College of Pharmacy
University of Cincinnati
Cincinnati, US

Yu Zeng, PhD
President/Professor, Health Industry Institute of South China Sea
Hainan Medical University
Haikou, Hainan, China

Tiantian Zhang, PhD
Associate Professor, College of Pharmacy
Jinan University
Guangzhou, China

Fei-Li Zhao, PhD
Conjoint Senior Lecturer, School of Biomedical Sciences and Pharmacy
The University of Newcastle
Callaghan, NSW Australia

Table of Contents

Introduction.....	1
Instructions	4
Executive Summary.....	6
Chapter 1: Study Questions	15
1.1 Background.....	16
1.2 Objectives and Questions	16
1.3 Perspective.....	16
1.4 Target Population.....	17
1.5 Interventions and Comparators.....	18
Chapter 2: Study Design.....	20
2.1 Types of Study.....	21
2.2 Assumptions	23
2.3 Sample Size	23
2.4 Time Horizon.....	24
Chapter 3: Cost.....	26
3.1 Cost Identification	27
3.2 Cost Measurement	28
3.3 Cost Valuation	29
Chapter 4: Discounting.....	31
Chapter 5: Health Outcomes.....	34
5.1 Efficacy/Effectiveness	35
5.2 Utility.....	37
5.3 Benefit.....	41
Chapter 6: Evaluation Techniques.....	43
6.1 Types of Evaluations	44
6.2 Incremental Analysis	45
Chapter 7: Model Analysis	48
7.1 Model Questions.....	50

7.2 Choice of Modeling Techniques.....	50
7.3 Model Development	52
7.4 Data Sources	54
7.5 Model Assumptions	57
7.6 Model Validation.....	58
7.7 Model Transparency	60
7.8 Model Localization or Adaptation.....	61
Chapter 8: Variability and Uncertainty	63
8.1 Variability Analysis.....	64
8.2 Subjects of Uncertainty Analysis.....	64
8.3 Methods of Uncertainty Analysis	65
8.4 Presentation and Interpretation.....	67
Chapter 9: Equity.....	69
Chapter 10: Generalizability.....	72
Chapter 11: Budget Impact Analysis.....	76
11.1 Perspective.....	78
11.2 Target Population.....	78
11.3 Scenarios.....	79
11.4 Time Horizon.....	79
11.5 Market Share.....	80
11.6 Costs	80
11.7 Computing Framework.....	81
11.8 Uncertainty and Scenario Analyses	81
11.9 Validation	82
11.10 Data Sources and Hierarchy	82
References	85
Appendix 1: Standard Reporting Format.....	89
Appendix 2: Pharmacoeconomic Evaluation Quality Checklist	96

Introduction

Pharmaceutical technologies play an important role in maintaining and promoting the population health. Research shows that with the increasing incomes and technology development, medical expenditure has grown the fastest among all consumptions. It not only exceeds the income growth rate, but also other consumption growth rate. Worldwide, the total health expenditure accounts for approximately 6.3% of gross domestic product in recent years. The proportion is over 10% in developed countries (WHO, 2018), while approximately 6.4% in China(National Health Commission of People's Republic of China, 2018). Of note, the drug expenditure in China contributes to a much higher percentage of the total health expenditure than other countries and regions. It is expected that, in the future, health technology assessments of pharmaceutical products will play a more central role in resource allocation in China. Therefore, how to scientifically evaluate and efficiently allocate healthcare resources will become a critical topic in the sustainable development of economy and health care in China.

Pharmacoeconomics is an interdisciplinary subject, which studies how to achieve maximum health improvement with limited resources. Using the theories from applied economics, pharmacoeconomics compares and analyzes the economic costs and health returns of pharmaceutical technologies in a systematic and scientific manner and aims to inform about the optimal strategy for decision-making and improve the overall efficiency of healthcare resource allocation. In general, the economic costs in a pharmacoeconomic evaluation include all healthcare resources consumed during the diagnosis and treatment of a disease. The healthcare returns can be measured as effectiveness, utility, and benefit. In addition to the costs and healthcare returns described above, The Professional Society for Health Economics and Outcomes Research (ISPOR) has proposed another value framework with additional factors (Neumann et al., 2018).

In the 21st century, with the rapid growth of China's economy, various major reforms for social development are also moving forward. In 2009, China launched a new round of reform for the national healthcare system. The reform aimed at further improving the accessibility and service quality for hospital visits while better managing its economic costs, thus improving the overall efficiency of healthcare resource allocation. Because of the importance of pharmacotherapy in current clinical practice in China, pharmaceutical policy has become the focus in China's healthcare system reform at this stage. Multiple government agencies have emphasized the importance of the pharmaceutical policy and indicated that pharmacoeconomic evaluations will play a critical role in the decisions for the essential healthcare system, health insurance system, and essential drug policies. In 2009, *Opinions of the Communist Party of China (CPC) Central Committee and the State Council on Deepening the Health Care System*

Reform ([2009] No. 6) (hereinafter referred to as *The Opinions*) was published. It states that China will "establish a scientific and appropriate system to establish drug prices" and "gradually require pharmacoeconomic evaluations for new drugs and patented drugs before the price is determined". In 2016, the CPC Central Committee and the State Council published the blueprint for "Healthy China 2030", which called for an applied outcomes assessment and technology evaluation system. At the same time, in the *13th Five-Year Plan for Deepening the Health Care System Reform*, it again emphasized that pharmacoeconomic evaluations should be an integral part in drug price negotiations and in the determination of the essential drug list. In 2017, the Ministry of Human Resources and Social Security of the People's Republic of China initiated the drug reimbursement negotiation, in which reports of pharmacoeconomic evaluations and budget impact analyses were officially included as the supporting evidence in determining prices. In 2019, the newly established National Healthcare Security Administration published the *2019 Work Plan for National Healthcare Insurance Drug Reimbursement List Adjustment* which clearly stated that "drugs in the same class should be compared based on the principles of pharmacoeconomic evaluations, and the drugs with demonstrated clinical necessity, safety and efficacy, and reasonable price should be prioritized (for the consideration of the reimbursement list)." At the same time, during the systematic update and revision of the national reimbursement list, the National Healthcare Security Administration also invited pharmacoeconomic experts to participate in the process. In the near future, the *14th Five-Year Plan* will be implemented, where "Healthy China 2030" and deepening the national healthcare system reform are the core elements in the social development. Therefore, pharmacoeconomic evaluations will play an increasingly important role.

Due to its delayed introduction in China, pharmacoeconomics has not been systematically applied in the healthcare decision-making process in China. However, pharmacoeconomics is at the stage of rapid adoption and development in China. In recent years, the number of pharmacoeconomic papers published by Chinese researchers increased rapidly. There are hundreds of papers published in English journals, with the quality comparable to the international average. In addition, the number and capability of Chinese researchers continue to grow (Thomas et al., 2019). However, the quality of pharmacoeconomic studies in China varies substantially. There have not been standards that guide these studies, and thus there remains substantial room for the improvement of the overall quality. Given the importance of pharmacoeconomic evaluations in the real-world practice and decision-making, the scientific rigor and standardization of the methods will be the key that influences the results and value of pharmacoeconomic evaluations. Experiences from developed countries have shown that without standards for systematic research and evaluations, the quality of results varies, and variations across studies, such as study design and reporting standards, can lead to different

conclusions. Such variations will affect the comparability across different studies and their scientific rigor and reference value for healthcare decision-making. To date, 44 countries and regions have set up their own guidelines for pharmacoeconomic evaluations to guide and standardize pharmacoeconomic studies within their respective countries/regions (ISPOR, 2019). Therefore, it is critical to develop the China-specific pharmacoeconomic evaluation guidelines to guide pharmacoeconomic research in China, which will help standardize the pharmacoeconomic studies and improve their value as the scientific guidance for healthcare decision-making.

The guidelines for pharmacoeconomic evaluations are developed based on pharmacoeconomic theories and include the standards by which evaluations of pharmaceutical therapies should abide. The current Guidelines aim to provide a general framework and standards for pharmacoeconomic evaluations, serving as a methodological guide for implementing pharmacoeconomic research, and a set of standards for research quality assessment. For a systematic study of the fundamental knowledge about pharmacoeconomics, one can refer to relevant books on pharmacoeconomics (Chen et al., 2006; Hu et al., 2009; Sun et al., 2015; Wu et al., 2017). As pharmacoeconomic research continues to evolve, a number of methodological issues need to be improved, and the guidelines regarding these issues may vary across different countries (Knies et al., 2010). We hope the current Guidelines will benefit the development of pharmacoeconomics in China, improve the efficiency of healthcare resource allocation, and facilitate the development of healthcare services in China.

Instructions

The current Guidelines are revised and updated based on the *China Guidelines for Pharmacoeconomic Evaluations (2011 Edition)*, and provide methodological guides for pharmacoeconomic evaluations and general standards for the economic evaluation of pharmacotherapy-related programs. The Guidelines are intended for two groups of users. The first group consists of pharmacoeconomic evaluation researchers in China, who can standardize and improve the quality of their research by following the Guidelines. The second group includes decision-makers from relevant healthcare agencies in China, who can refer to the Guidelines to evaluate the quality of the pharmacoeconomic research submitted by other organizations. These agencies may include health insurance management departments, drug price management departments, essential drug policy management departments, and new drug review and evaluation departments. These agencies can require or recommend relevant companies to submit pharmacoeconomic evaluation reports of their drugs at the time of launch, price negotiation, or reimbursement negotiation. These government agencies can also create a national pharmacoeconomic expert review committee, which will be responsible for evaluating pharmacoeconomic evaluation reports submitted by pharmaceutical companies and generating the final reports for the economic values of the pharmacotherapy-related programs that will be used in decision-making. This Guidelines can be used as standards which the review committee will used to evaluate the quality of the pharmacoeconomic reports submitted by companies.

As an emerging interdisciplinary subject, pharmacoeconomics is still rapidly evolving. At the same time, China's healthcare policy is constantly being updated. Therefore, the *China Guidelines for Pharmacoeconomic Evaluations* need to be continuously updated and improved. To this end, we will maintain an open and dynamic model to update the *China Guidelines for Pharmacoeconomic Evaluations* as needed. The model allows us to continuously receive comments on the Guidelines from different communities, including academia, government, research institutes, and industry, organize regular meetings with experts to discuss these comments, and further revise and publish new editions of the Guidelines.

If you have any comments or suggestions on the 2020 edition of the *China Pharmacoeconomic Evaluation Guidelines* (hereinafter referred to as *The Guidelines [2020 Edition]*), please contact China_PEG2018@163.com.

China Guidelines for Pharmacoeconomic Evaluations Working Group

Dec, 2020

Executive Summary

The Guidelines (2020 Edition) includes six parts, introduction, instructions, executive summary, body, references, and appendices. The body is written in accordance with the main methods and techniques involved in a pharmacoeconomic evaluation, and includes a total of eleven chapters.

Chapter 1 focuses on the study questions. The first step in a pharmacoeconomic evaluation is to clarify the study questions, including the background, objectives, research questions, perspective, target population, interventions, comparators, and time horizon. The background should provide an overview of the epidemiology and economic burden of the disease, the main interventions and their efficacy and safety, the recommended treatment regimens according to the local clinical guidelines and the ones in other countries/regions, the current status of pharmacoeconomic evaluations of the interventions in global literature and the value of the study. Researchers should clearly state the main study objectives and questions to be investigated in the pharmacoeconomic evaluation, and clearly define the study perspectives according to the study objectives and the intended recipients of the report. These should be consistent throughout the study. Commonly used perspectives include societal perspective, healthcare system perspective, payer perspective, healthcare institution perspective and patient perspective, etc., among which, societal perspective and healthcare system perspective are recommended in economic evaluations in China. The study should clarify the target population for the pharmacoeconomic evaluation and its inclusion and exclusion criteria. When describing the target population, it is recommended to include epidemiological characteristics of the patients, such as age, sex, disease type and severity, presence of comorbidities or risk factors, and socioeconomic status. Evaluations should be performed at the level of the overall target population or in patient subgroups, if needed. Descriptions of interventions and comparators should include information such as formulation, dose, dose frequency, treatment route, concomitant medications, and treatment background. It is recommended that the selection of comparators should prioritize standard or conventional treatment regimens for that indication. If there is no effective treatment, or intervention is not recommended for certain diseases, “no intervention” can be used as the comparator. In such cases, the rationale for using placebo as the comparator should be justified.

Chapter 2 focuses on study design. Based on whether simulation is used, the studies can be classified into two categories: modeling studies and studies based on individual patient-level data. The studies based on individual patient-level data can further be classified into prospective studies and retrospective studies. Moreover, based on whether a study includes an intervention, prospective studies can be classified into observation studies and experimental studies, which include piggyback study alongside the randomized controlled trial (RCT) and pragmatic clinical trial (PCT). Sufficient details should be provided regarding the rationale and

justifications for the key assumptions related to the study design or model estimation. With regard to the sample size, it should be adjusted based on the needs for a pharmaco-economic evaluation in a piggyback study alongside a RCT. In a programmatic clinical trial or other types of prospective studies, the minimum sample size should be assessed based on the real-world distribution of relevant parameters. In a retrospective study, a minimum sample size is usually not considered. Modeling studies do not need sample size estimation. The time horizon and its rationale should be stated in the study design. The time horizon needs to reasonably reflect a disease's natural progression, and the duration should be long enough to observe all the impact of an intervention on costs and health outcomes to a patient.

Chapter 3 focuses on cost. The cost analysis mainly includes cost identification, cost measurement, and cost valuation. Costs in pharmaco-economic evaluation include direct cost, indirect cost, and intangible cost. Direct costs also include direct medical costs and direct non-medical costs. The scope of cost identification should be consistent with the study perspective and study duration. It should include all the current and future costs related to the intervention during the study time horizon. If an intervention prolongs life, the cost analysis should include the disease-related costs and the intervention costs incurred during the extended lifespan. If an intervention leads to adverse drug reactions (ADRs) is measured, all costs related to addressing the ADRs should be included, especially costs related to monitoring serious ADRs. In a piggyback study parallel to an RCT, costs that directly result from the RCT but will not occur in the real-world clinical practice should be identified and excluded. When estimating costs, we should first list the types of resource utilizations related to the intervention, define the unit for each type, and then estimate the quantity of each type of resource utilizations based on the defined unit. Whenever possible, the costs should be estimated based on the data from a Chinese population. If such data is unavailable, data from other countries should be adjusted to make it more suitable for China. During cost valuation, the quantity of each type of resource utilization is multiplied by its unit price and then the costs for all types are summed to obtain the total cost. Unit price should be obtained corresponding to the defined unit for resource utilization. It is recommended that unit prices should be obtained from the latest pricing information published by the government or an authoritative source, such as the final price for provincial tendering or price determined during the national reimbursement negotiation. If the drug has not been launched in China, it is recommended to use the manufacturer's suggested price for analysis. Use of another pricing system should be clearly indicated, and justifications for such sources should be provided. It is recommended to use the human capital approach (HCA) to perform calculations for the indirect cost related to the disease.

Chapter 4 focuses discounting. When the time horizon is more than one year, studies should discount cost and health outcomes that occur in the future (i.e., converting future costs

and health outcomes to the values at baseline). The same discount rate is recommended for both cost and health outcomes. It is recommended to use 5% per year as the discount rate for the base case. In addition, a sensitivity analysis should be conducted by varying the discount rate within the range of 0%–8%. Justifications should be provided if other discount rates are used.

Chapter 5 focuses on health outcomes. Health outcomes can be estimated using three categories of measurements, efficacy/effectiveness, utility, and benefit. Efficacy should be based on the best available evidence (i.e., the best evidence among the clinical efficacy studies and effectiveness studies). For a new intervention, when clinical efficacy data from an RCT is available and applicable, the data from the RCT is preferred. For interventions that have been on the market for some years, when updated efficacy data is not available or applicable, effectiveness data from real-world studies should be used. Clinical efficacy data obtained from a systematic review or meta-analysis of RCT has a higher level of clinical evidence in evidence-based medicine and is thus preferred. RCT data based on Chinese populations or international multi-centered RCT data with Chinese populations included are also preferred. Efficacy data from RCTs of direct head-to-head comparisons between the intervention group and the control group are preferred. Studies including the final end-points are preferred to be used in pharmacoeconomic evaluations. It is recommended to use quality-adjusted life year (QALY) as the measurement that incorporates utilities. Survival time and health utility values should be reported before QALY is presented. The measurement of health utility includes direct measurement and indirect measurement, and the latter is preferred. Commonly used health utility instruments in indirect measurement include EuroQol-5 Dimensions (EQ-5D), Short-Form Six-Dimensions (SF-6D), etc. For children, EQ-5D-Y is recommended as it is a health utility instrument specifically for children. Direct measurement could be performed when there is no suitable tool for indirect measurement to obtain health utility values for certain diseases or symptoms. Commonly used direct measurement methods include standard gamble (SG), time trade-off (TTO), and discrete choice experiment (DCE), etc. When an indirect measurement is used, generic utility instruments, such as EQ-5D-3L, EQ-5D-5L, and SF-6D V2, etc., are preferred, if there is evidence showing that these instruments have good reliability and validity in the target disease. On the other hand, if evidence shows that generic utility instruments are insufficient to assess characteristics of the patient population or disease symptoms, disease-specific utility instruments could be used. It is recommended that health utilities should be estimated using the scoring algorithms based on the preference of the general population and the scoring algorithms based on a Chinese population is preferred. When utilities cannot be obtained through direct measurement, they can be extracted from the published studies through a systematic literature review. Benefit is the quantification of health outcomes using monetary terms. The benefits of a treatment regimen include direct benefit,

indirect benefit, and intangible benefit. Direct benefit quantifies the gains as the actual monetary exchanges resulting from an intervention. Indirect benefit and intangible benefit quantify gains for which no actual monetary exchanges occur, and their calculation usually relies on methods such as human capital approach (HCA) or willingness to pay (WTP).

Chapter 6 focuses on evaluation techniques. Evaluation methods include cost-minimization analysis (CMA), cost-effectiveness analysis (CEA), cost-utility analysis (CUA) and cost-benefit analysis (CBA), etc. Researchers should choose an appropriate evaluation technique according to the characteristics of the intervention, the availability of data, and the evaluation objectives and requirements of the study. If possible, a cost-utility analysis (CUA) should be conducted. Other techniques, such as CEA, CMA or CBA can also be used but the justifications should be stated. Researchers can perform an evaluation using two or more techniques. They can also use one technique as the main method but also include other techniques, and then compare and analyze the differences in the results from different evaluation techniques. In CUA and CEA, the decision-making is based on results from the incremental analysis. Incremental analysis is the comparison of costs and outcomes between the intervention and the comparator. If the intervention has a lower cost and a better outcome compared to the comparator, it is the dominant regimen. In contrast, if the intervention has a higher cost and a worse outcome compared to the comparator, it is the strictly dominated regimen. If the intervention has both a higher cost and a better outcome compared to the comparator, the incremental cost effectiveness ratio (ICER) between the two regimens needs to be calculated. If the ICER is smaller or equal to the threshold value, then the intervention is cost-effective than the comparator. If the ICER is larger than the threshold value, the intervention is not cost-effective compared to the comparator. In an analysis of incremental outcomes, 1-3 times national GDP per capita is recommended as the willingness-to-pay threshold per QALY.

Chapter 7 focuses on model analysis. Pharmacoeconomic evaluation models usually use methods, such as graphs, equations, etc. to abstractly simulate the natural progression of a disease and the effect of interventions on its progression. It focuses on the interventions and the important clinical events as well as the health outcomes, changes, and resource use incurred during this process. The description of questions for decision-making is the starting point of model construction and analysis. It is recommended that a study should clearly state the disease, the model objectives, the target population, the intervention, the study perspective, scope of the simulation, health and other outcomes as well as the time horizon. Economic evaluations can be conducted using different types of models. The most common types include decision tree model, the Markov model, the discrete events simulation model (DES), the partitioned survival model (PSM), and dynamic models, etc. Before choosing the modeling technique, researchers

should first understand the clinical characteristics of the disease and then consider the availability of data. Model conceptualization should be based on the disease course and the effect of the intervention on disease progression. The model structure is normally simplified compared to the actual disease course. A model structure diagram should be presented to demonstrate the model. Researchers should systematically identify, collect, and evaluate the data used in the model, and describe the sources of all inputs in the model and the rationale for using these sources. When there are multiple sources for model inputs, different factors should be considered when choosing the appropriate source, such as the quality of parameters, the characteristics of the population in the data source, the country or region for data collection, the practice setting for data collection, the duration of data collection, etc. To the extent possible, these factors should be consistent between the studies sourcing to the parameters and the model. If necessary, clinical experts should be consulted, and a sensitivity analysis or variability analysis should be performed. Clinical data sources should be comparable among different treatment arms in a model. Researchers should describe and explain the assumptions regarding causality, generalizability, scope, structure, and data, etc. in the model. An uncertainty analysis should be performed to assess the key assumptions. Researchers should conduct model validation, including the model's face validity, internal validity, external validity, cross validity and predictive validity. A model should have sufficient transparency in order to allow readers, reviewers, and healthcare policy-makers to evaluate the credibility of the model and determine whether the model results are suitable to inform decision-making given the specific setting faced by decision-makers.

Chapter 8 focuses on variability and uncertainty. Variability refers to the variation in parameters related to the differences in treatment background that may affect the evaluation results. Variability cannot be eliminated completely. A sensitivity analysis or scenario analysis can be performed to assess variability caused by differences in region or background. Variability due to the heterogeneity of patients should be handled at the stage of study design by dividing patients into smaller but more homogenous subgroups. In addition, researchers should conduct a comprehensive analysis of different types of uncertainties in the pharmacoeconomic evaluation, including uncertainties in methodology, parameters, and modeling. Uncertainties in methodology and modeling are often assessed with scenario analysis. The uncertainty in parameter can be assessed using a deterministic sensitivity analysis (DSA), such as one-way sensitivity analysis, multi-way sensitivity analysis and extreme value analysis, or probabilistic sensitivity analysis (PSA) with Monte Carlo simulation. In a deterministic sensitivity analysis, sufficient rationale should be provided to determine the range of variation in the parameters. In probabilistic sensitivity analysis, a large number of parameters should be included to the extent possible. The probability distribution, the distribution parameter, and the

number of Monte Carlo iterations should be described and justified. When there are multiple uncertain factors, a tornado diagram could be used to present the results of a deterministic one-way sensitivity analysis. It is suggested to use the cost-effectiveness acceptability curve (CEAC) or cost-effectiveness scatter plot to present the results of a probabilistic sensitivity analysis in a pharmacoeconomic evaluation. The results from both the sensitivity analysis and the base-case analysis are equally important; thus researchers should avoid drawing conclusions based mainly on the base-case results.

Chapter 9 focuses on equity. In pharmacoeconomic evaluation, equity means that the values of all lives, life years, and QALYs affected by an intervention are (assumed) to be equivalent, regardless of the age, sex, or social status of individuals in the target population. If possible, the base-case results should be evaluated for equity. There are two methods to address equity issues. The first method is to perform a sensitivity analysis to illustrate the effect of equity assumptions on the results. The second method is to perform a subgroup analysis using pre-specified factors in order to compare equity-related characteristics between subgroups that benefit more versus less from the intervention, such as age, sex, race/ethnicity, region, socioeconomic status, health status, and other population characteristics. When the effectiveness varies among subgroups, and it is possible to implement the intervention in different subgroups, the cost-effectiveness results should be reported for each subgroup.

Chapter 10 focuses on generalizability. When data (e.g., economic, clinical, and humanistic data) is generated based on other healthcare settings (including other countries, regions or healthcare systems), researchers need to assess its suitability for the healthcare setting in the current study. If data adjustment to the current healthcare setting is required, the methods used for the adjustment should be described and its suitability should be demonstrated. Epidemiological data often varies geographically. When only non-national epidemiological data can be obtained, researchers should evaluate beforehand whether applying such data in the existing study is likely to result in bias. If bias exists, researchers should quantify the bias to the extent possible. Regarding the applicability of clinical data, researchers should clarify the differences between efficacy observed in the clinical trials and effectiveness in the real world, especially when using efficacy data from a phase III clinical trials in pharmacoeconomic evaluations. When using data from global multi-center clinical trials in the pharmacoeconomic evaluations, researchers should consider whether they should use pooled data from multiple countries or the data from the country or region that is most suitable for the decision makers. When applying intervention cost data obtained from a certain country or region, researchers need to pay attention to the variation in cost data between different countries or regions and make corresponding corrections and adjustments. Regional variations in healthcare organizations or agencies and their levels (tertiary, secondary, primary) will have an impact on

the generalizability of the data.

Chapter 11 focuses on budget impact analysis (BIA). BIA is the evaluation of the impact on expenditure of a healthcare system after a new intervention enters a healthcare system (e.g., the reimbursement list). The perspective of a BIA is normally a budget holder's perspective. Depending on the need from the decision-makers, the perspective can be defined as different levels of government payers from national to local level, commercial insurance organizations, or a medical institution of a certain type in a certain area. Different study perspectives will affect the range of the budget impact estimate. Budget impact analysis should clarify the target population. The target population size should be estimated based on the inclusion and exclusion criteria and other applicable patient characteristics. Two market scenarios of BIA should be clearly defined, namely, the "without entry scenario" where the new intervention is not included in the coverage of a budget holder, and the "with entry scenario" where the new intervention is within the coverage of a budget holder. Both scenarios should take into account expected market changes, including the launch of other interventions into the market, withdrawal of similar drugs from the market, and possible alternative treatments. The time horizon is usually between 3 to 5 years. Discounting is not recommended. Researchers should report the market shares in both scenarios, i.e., "without entry" and "with entry" scenarios, and the three types of market share changes (substitution, combination, and expansion) that are expected from the entry of the new intervention. The market shares in the "without entry scenario" are generally obtained from studies using real-world data. The market share in the "with entry scenario" predicts the market share of the new intervention in the target population and the market shares of all interventions in the target population; it is normally based on specific assumptions. Researchers should ensure the transparency of the prediction method and should describe in detail the assumptions, the reference data and the selected prediction model. Costs in BIA include two parts, the first is cost of the intervention itself, and the second is the impact of the intervention on other costs, including disease-related costs and indirect costs. However, indirect costs are not recommended to include in the analysis. The calculation framework in a BIA is generally presented in Excel. To the extent possible, the model should be presented as a "cost calculator" that clearly lays out each cost component so that it can be easily understood by decision makers. Researchers need to record and present decisions related to the selection of model structure and underlying assumptions, assess uncertainties through scenario analyses (by changing structural assumptions) and one-way/multi-way sensitivity analyses (by changing selected parameter input values), and, if necessary, conduct probabilistic sensitivity analyses. In addition, validation should be performed for the core analysis and model inputs data, including face validity, technical validation and external validity. Finally, in data source selection, data should be suitable to address the questions from the decision makers'

perspectives. It is recommended to use the high-quality data in the same region and the same population.

The appendices to the Guidelines include two sections: Appendix 1, the template for standard reporting, and Appendix 2.

Chapter 1: Study Questions

The first step in a pharmacoeconomic evaluation is to identify the primary study questions, including the background, objectives, study questions, perspective, target population, interventions, comparators, and time horizon. In addition to the primary study questions, secondary study questions, such as the different effects of interventions on distinct subgroups or by different treatments (e.g., monotherapy and combination therapy) may also be identified (CADTH, 2017).

1.1 Background

The background should provide the following information: an overview of the epidemiology and economic burden of the disease, main interventions (including pharmacotherapy and non-pharmacotherapy), their efficacy and safety, the recommended treatment regimens based on the local clinical guidelines and those in other countries/regions, the current status of pharmacoeconomic evaluation of the interventions in global literature (general conclusions and limitations), and the value of the study (necessity and importance).

1.2 Objectives and Questions

Researchers should clearly state the main study objectives and questions to be addressed in the pharmacoeconomic evaluation. All questions should be framed in an answerable and testable fashion.

Explanations

The research questions in a pharmacoeconomic evaluation usually include population/participants, intervention, comparator, outcomes and study setting, etc. For example, is it cost-effective to reduce the high blood pressure control from the original goal (systolic blood pressure ≤ 140 mmgh) to the intensive goal (systolic blood pressure ≤ 120 mmgh) among hypertension patients in the United States? (Bress et al., 2017)

1.3 Perspective

1.3.1 Researchers should clearly define the study perspective according to the study objectives and the intended audience of the report. Study perspectives mainly include the societal perspective, the healthcare system perspective, the payer perspective, the health care provider's perspective, and the patient's perspective.

1.3.2 Evaluations from a societal perspective and a healthcare system perspective are recommended. However, researchers could choose a suitable study perspective based on the study objectives. All pharmacoeconomic evaluations that will be used for public policy-making should be conducted from a societal perspective.

1.3.3 More than one perspective can be used in a pharmacoeconomic evaluation, but the perspective should remain consistent throughout the study.

Explanations

(1) The perspective plays an important role in pharmacoeconomic evaluations. Once the study perspective is determined, the evaluation process, including the study design, the analytical method, and calculation of costs and of effectiveness will be decided accordingly. The range and estimation of costs vary considerably across different perspectives.

(2) Based on the social welfare theory, a societal perspective is the optimal perspective in pharmacoeconomic evaluations. This means that regardless of who are the investors and beneficiaries, the analysis should include all costs and benefits, including those occurring outside of the healthcare system. For evaluations based on non-societal perspectives, changes in the cost-effectiveness with a societal perspective can be discussed as needed.

(3) For evaluations from different perspectives, how to handle transfer costs is question that need to be addressed. From a societal perspective, common examples of transfer costs include maternity insurance, worker's compensation, commercial medical insurance, and disease allowance in social welfare in other countries. From a societal perspective, these types of transfer costs should not be included in the cost calculation in order to prevent double counting. However, from a payer perspective (e.g., healthcare system or an insurance agency), such transfer costs must be included because they cannot be compensated (Genduso et al, 1996).

1.4 Target Population

1.4.1 The study needs to clarify the target population for the pharmacoeconomic evaluation and its inclusion and exclusion criteria. Generally, the target population is consistent with the drug's indication. When describing the target population, epidemiological characteristics of patients should be included, such as age, sex, disease type and severity, presence of comorbidities or risk factors, and socioeconomic status.

1.4.2 Pharmacoeconomic evaluations are usually performed in the overall target population level or, if needed, in patient subgroups defined by clinical characteristics. Subgroups may be defined by population characteristics, disease subtypes, severity, and presence of comorbidities.

1.4.3 Restrictions in clinical trials may lead to differences between clinical trial populations and real-world patients receiving the interventions. If there are such differences, the impact of different population on results should be further investigated.

Explanations

Subgroup variations provide important information to policy-makers helping them identify individuals that should be prioritized for an intervention. However, variables used to define subgroups and the number of subgroups are limited by the sample size. Therefore, researchers need to balance precision of the analysis and statistical power. Evidence shows that a subgroup analysis should be conducted when there are variations in clinical effectiveness and cost-effectiveness among different patient groups. A subgroup analysis should be conducted based on a priori list of variables; a subgroup analysis should be minimized in a simple analysis. The following questions can help identify whether differences among subgroups are true differences (Oxman et al., 1992):

- (1) Is the difference clinically meaningful (e.g., leading to different treatment recommendations among subgroups)?
- (2) Are the differences among subgroups statistically significant?
- (3) Is the subgroup hypothesis specified before the current analysis?
- (4) Should the type of study minimize the use of subgroup analysis?
- (5) Are the differences among subgroups observed in the same study (instead from two separate studies)?
- (6) Are the differences among subgroups consistent with other studies?
- (7) Is there indirect biological evidence supporting the differences among subgroups?

1.5 Interventions and Comparators

1.5.1 Descriptions of interventions and comparators should include information such as formulation, dose, dose frequency, treatment route, concomitant medications, and treatment background. It is recommended that the selection of a comparator should prioritize standard treatment for the same indication. If there is no existing standard treatment, conventional treatment in the clinical practice may be considered. If there is no effective treatment or an intervention is not recommended (e.g., watchful waiting approach in prostate cancer), placebo (i.e., no intervention) can be considered as a comparator in a pharmacoeconomic evaluation. In

such case, the rationale for no treatment in this disease should be justified. The comparator treatment should be described by its generic name, with its brand name listed.

1.5.2 If a new intervention is in the same class as the existing treatments, in principle, the comparator should be the standard treatment or the most commonly used treatment in the same class. If the intervention targets a new indication without standard treatments, the comparator should be the treatment with the most similar indication. If the objective of the study is to include a new intervention in the reimbursement drug list or a hospital formulary list, the comparator should be the available treatments on the lists. If the comparators cannot be decided for a rare disease treatment based on the above rules, existing treatments in the real-world setting can be considered as the comparators.

1.5.3 If the new intervention to be evaluated is Chinese medicine or Chinese patented medicine (non-adjuvant therapy), it is recommended to select Western medicine with the same or similar indications, as the comparator in a pharmacoeconomic evaluation. Chinese medicine of the same indications can also be considered.

Explanations

A pharmacoeconomic evaluation is based on comparisons among different treatment regimens; therefore, the results depend on the choice of comparators, which will substantially affect the intervention's therapeutic value and economic value. Ideally, a new drug should be compared with the most cost-effective treatment available. In reality, there are many options for comparators, such as the routine treatment and the standard treatment, etc. The "routine treatment" should be the most commonly used clinical treatment or the one with the largest market share while the "standard treatment" is the treatment that has been proven to be most effective among routine treatments. It is common to have more than one comparator in a single study.

Chapter 2: Study Design

2.1 Types of Study

2.1.1 Pharmacoeconomic evaluations can be classified as modeling-based studies and individual-level data-based studies according to whether simulation is used.

2.1.2 Individual-level data-based studies can be classified as prospective studies and retrospective studies. Prospective studies can be further divided into prospective experimental studies and prospective observational studies based on whether the study has an intervention. Moreover, experimental studies include piggyback studies alongside the randomized controlled trials (RCT) and pragmatic clinical trials (PCT).

Explanations

(1) A piggyback study alongside an RCT combines a pharmacoeconomic evaluation and a clinical trial, which is usually carried out during Phase III clinical trials, but can occasionally be included in Phase II or Phase IV clinical trials. It is a widely adopted study design and provides high credibility and internal validity due to the rigorous double-blinded randomized controlled design in the clinical trials. For the pre-marketing economic evaluation of a new treatment, a piggyback study is the best option, because it provides timely evidence to support post-launch decision making for drug pricing and reimbursement access. However, a piggyback study has its limitations. For instance, its external validity is low.

(2) A pharmacoeconomic evaluation based on a PCT means the study is based on the use of treatments in a real-world setting (MacPherson, 2004). Compared to piggyback studies alongside RCT, this type of study is more flexible in study design. More specifically, it has less stringent inclusion and exclusion criteria, uses active controls rather than placebo as the comparator, randomizes patients into treatment groups, but usually does not strictly control the implementation of intervention. It can include more outcomes supporting reimbursement decision-making. On the other hand, the limitations of PCT studies include longer follow-up periods, higher costs, and lower internal validity. However, if affordable, a PCT study remains a highly recommended study type for pharmacoeconomic evaluations.

(3) An observational study is a non-randomized clinical study. Patient information used for pharmacoeconomic evaluations is prospectively collected based on the data collection protocol without any intervention in physician's clinical decision. Although this type of study possesses good external validity, it has lower internal validity due to the fewer restrictions on external factors, poor patient compliance and various confounding factors. Moreover, since patients are not randomized, the differences in baseline characteristics between the intervention group and the control group may impact costs and health outcomes, which increases the difficulty of the analysis.

(4) A retrospective study is also a non-randomized clinical study. In pharmacoeconomic evaluations, it usually refers to a retrospective cohort study. Patients are selected from various databases or existing records such as hospital medical records, where their information is collected and used to compare cost and effectiveness between the intervention group and the control group. The advantages of this type of studies are that data are usually obtained from existing databases, leading to lower cost, shorter study duration, and higher external validity. However, unlike in prospective studies, researchers cannot control the quality of data collected nor define patient characteristics for the sample in a retrospective study. Therefore, there will be differences between the samples for the intervention group and the control group, which may bias the result. Additionally, the existing data used in retrospective studies are often not collected for the purpose of pharmacoeconomic evaluations, and thus cannot meet the requirements of the design for such studies in most cases.

(5) A mixed study is mainly a combination of the aforementioned study design methods, which usually obtain clinical effectiveness data on patients from clinical trials or observational studies and cost data from retrospective studies or cross-sectional studies. A mixed study has the advantage of using multiple data sources to address study questions that cannot be solved with a single data source, which is an economical and efficient option for pharmacoeconomic evaluations. When a prospective study is not feasible, mixed studies can be used as an alternative option. However, as the sources for cost and effectiveness data are different, these differences may bias results.

(6) A modeling study is the most common study design in cost-effectiveness studies. A decision analytic model “uses mathematical relationships to define a series of possible consequences that would flow from a set of alternative options being evaluated. Based on the inputs into the model, the likelihood of each consequence is expressed in terms of probabilities, and each consequence has a cost and an outcome. It is thus possible to calculate the expected cost and expected outcome of each option under evaluation” (Briggs et al., 2006). Decision analytic models include decision tree model, Markov model, discrete events simulation model, partitioned survival model and dynamic transmission model, etc. Modeling data could be obtained from a clinical trial, epidemiological research, meta-analysis, a real-world study, expert opinions, small-scale field investigation, and a literature review. This design is helpful when multiple outcomes (e.g., efficacy, safety) are considered, clinical trial results need to be extrapolated, the time horizon is long, or the study budget is limited. As it saves time and effort, a modeling study may provide relatively comprehensive information in a short period of time, and thus has broad applications.

2.2 Assumptions

Key assumptions related to the study design or model estimation should be explained in detail and provided with justifications.

Explanations

There are many uncertainties and unevaluable parameters in pharmacoeconomic evaluations, especially in a modeling study. These uncertainties and parameters will significantly affect the reliability and the robustness of the model estimation and evaluation results. Therefore, reasonable assumptions need to be made and the rationale should be stated. Study assumptions for pharmacoeconomic evaluations could be made for various aspects, including study perspective, analytical technique, target population, comparator selection, time horizon, cost calculation, discounting, and clinical indicators (Briggs, 2000; Sculpher et al., 2000).

2.3 Sample Size

2.3.1 In piggyback studies alongside RCTs, researchers should adjust the sample size of the clinical trial to meet the needs of a pharmacoeconomic evaluation as much as possible. In PCTs and prospective observational studies, researchers need to conduct primary data collection and the marginal costs for data collection is usually high. Therefore, they need to consider the minimum sample size requirement, especially the distribution of relevant parameters in the real world. In retrospective studies, data are generally from a database with a large sample such as hospital electronic medical records or a medical claims database. The marginal costs for obtaining additional data is low. Usually, samples obtained from these studies are much larger than the minimum required sample size. Therefore, the minimum sample size requirement is not necessary for such studies. In a modeling study, a hypothetical population is used in the model, and thus the sample size does not need to be calculated.

2.3.2 Generally, depending on the study question, the sample sizes for pharmacoeconomic evaluations should be larger than the minimum sample size required by RCTs. When related parameters are available, the sample size should be estimated using the specific formula for a sample size calculation in a trial including pharmacoeconomic end-points (Backhouse, 2002; Jin, 1993). If it is difficult to obtain the parameters used to estimate sample size for a pharmacoeconomic endpoint, the sample size of each patient group in such a study should not be smaller than the one estimated based on the primary endpoint for a clinical trial or cohort

study.

Explanations

Sample size is based on the study question. Generally, sample sizes for pharmacoeconomic evaluations should be larger than the minimum sample size required for RCTs, because pharmacoeconomic trials and RCTs are different in comparators, study subject, study background, end-point, effect size, observation timeframe, acceptable levels of Type I and Type II errors, and statistical methods. More specifically, i. pharmacoeconomic trial normally uses active control instead of placebo as a comparator; ii. it allows a more general treatment population, with more patient heterogeneity; iii. it uses the end-point of economic analysis, and cost outcomes often have a skewed distribution with more variation; iv. its evaluation result is a comprehensive index that can measure both cost and effectiveness at the same time, rather than a single clinical efficacy index; v. its study duration is long enough to include all impact of the intervention on costs and patients' outcomes; vi. In addition to clinical efficacy, it also needs to estimate the Type I and Type II errors of economic outcomes; and vii. statistical methods need to include costs. Due to the aforementioned differences, in order to achieve the same level of statistical significance and power as an RCT, the sample size for a pharmacoeconomic evaluation should be appropriately increased compared to the minimum sample size required for an RCT.

2.4 Time Horizon

2.4.1 In pharmacoeconomic evaluations, researchers should clearly justify the choice of the selected time horizon.

2.4.2 The time horizon should reasonably reflect a disease's natural progression and the duration should be long enough to include all impact of the intervention on costs and patients' outcomes. To ensure the consistency of the analysis, data collection regarding cost and effectiveness should use the same time horizon.

2.4.3 In a modeling study, when simulating long-term cost and effectiveness outcomes of the intervention, in addition to the time horizon for the modeling study and its justifications as well as the simulated results, the short-term simulated results based on the original data should also be provided.

Explanations

The time horizon refers to a time frame during which researchers need to observe or simulate cost and health outcomes of an intervention in a certain disease. The time horizon is determined by disease type, treatment goals, and expected output of the study, etc. Some diseases occur and develop in a short timeframe. For instance, acute diseases might lead to death or cure in a very short time. For these diseases, researchers can observe the whole process of disease occurrence, progression, treatment, and prognosis in a short period. Therefore, the time horizon will be short. For treatments of chronic diseases, the optimal time horizon is the lifespan of the patient, but this does not mean that cost and effectiveness occurring during the patient's entire life needs to be observed, which is neither convenient nor feasible. A commonly used method is modeling based on the cost and effectiveness data from short-term studies (clinical trial or observational studies) to extrapolate the long-term cost and effectiveness. In such case, researchers should list the time horizon and results of the short-term studies, while explaining the suitability of the extrapolation, including its causal relationship, study assumptions, and justifications for the extrapolation.

Chapter 3: Cost

The cost analysis in a pharmacoeconomic evaluation mainly includes cost identification, cost measurement, and cost valuation.

3.1 Cost Identification

3.1.1 Costs in pharmacoeconomic evaluations include direct costs, indirect costs, and intangible costs. Direct costs also include direct medical costs and direct non-medical costs.

3.1.2 The scope of cost identification should be consistent with the study perspective. From the societal perspective, all direct medical and non-medical costs, as well as indirect costs, should be included. All pharmacoeconomic evaluations that intends to inform public policy decision-making should provide evaluation results from a societal perspective. From the health care system perspective, all direct medical costs within the health care system should be included. From the payer's perspective, all direct medical costs within the insurance plan should be included. From the health care provider's perspective, direct medical costs and direct non-medical costs (if any) borne by the provider should be included. From the patient's perspective, all relevant direct medical costs, direct non-medical costs, and indirect costs should be included. Researchers have flexibility in handling intangible costs by either putting it at the end of cost or health outcomes; however double counting should be avoided. When the intangible cost is significantly large, it needs to be evaluated separately.

3.1.3 The scope of the cost identification should be consistent with the study duration. All current and future costs related to the implementation of interventions within the study duration should be included. If an intervention prolongs life, the cost analysis should include the disease-related costs and the intervention costs incurred during the extended lifespan. The costs unrelated to the target disease or the intervention can be excluded.

3.1.4 If there are adverse drug reactions (ADRs) associated with the medical intervention, the costs incurred while treating the ADRs should be identified, especially serious ADRs (e.g., grades 3–4 ADRs based on the WHO ADR classification). There are two main types of costs associated with ADRs (Li et al., 2009): (1) the cost incurred to prevent or to monitor ADRs; (2) the cost related to ADRs treatments.

3.1.5 When pharmacoeconomic evaluations use data collected from clinical trials, protocol-driven costs incurred during the clinical trial but not in the real-world clinical practice should be identified and excluded. If it is difficult to determine whether certain costs would be incurred in the real-world clinical practice, a sensitivity analysis can be performed by using the cost structure of similar evaluations that are not based on clinical trials.

Explanations

(1) Direct costs are the costs incurred directly during medical services, which include direct medical costs and direct non-medical costs. Direct medical costs include the costs for medical resources associated with a certain treatment strategy, such as the registration fee, pharmacy fee, surgery fee, consultation fee, treatment fee, nurse fee, monitoring fee, material fee, ward fee, examination fee, oxygen fee, and other medical costs. Direct non-medical cost is the costs for non-medical resources incurred by the patient to receive medical services, such as transportation expenses, accommodation expenses, and nutritional food expenses, etc. Normally, it is difficult to accurately calculate direct non-medical costs because conditions vary greatly. Therefore, the direct non-medical cost could be ignored in the study if it only accounts for a small proportion of the costs.

(2) Indirect costs, also known as productivity costs, refers to the working hours and productivity loss due to disease, disability or death, which includes the loss of salary for patients and their families caused by discontinuing school, sick leave, and early death, etc.

(3) Intangible costs are the physical and mental distress and discomfort, such as pain, anxiety, and tension, caused by the disease or medical services, such as intervention implementation or diagnosis. Intangible costs are not usually measured independently because (i) they are difficult to measure accurately and more difficult to convert to monetary units; (ii) intangible costs are usually included in the outcome measurement through the utility assessment; thus they should not be double counted.

(4) Researchers are allowed to use other methods for cost classifications other than those mentioned above. However, all resources related to the intervention should be included to prevent omission or double counting.

3.2 Cost Measurement

3.2.1 When estimating costs, we should first list the types of resource utilizations related to the intervention, define the unit for each type and then estimate the quantity of each type of resource utilization based on the defined unit. There are three main types of measuring units, the natural unit of health resource consumption, the standards set by relevant government agencies, and the measuring units defined based on research needs.

3.2.2 The unit for cost measurement can be broad, such as annual hospital visits, one hospitalization, and one outpatient visit, etc. It could also be at the micro level, such as one drug tablet, one injection, and one nurse visit, etc. The micro units should be used whenever possible.

There are several advantages to using the micro units. First, it allows detailed examination of the structure of the cost data and assessment for its justifications. Furthermore, using micro units, cost data from different regions can be compared through data adjustments, even if different treatment regimens are used and the costs vary across geographic regions.

3.2.3 When possible, the costs should be estimated based on the data from the Chinese populations. If such data is unavailable, the data from other countries should be adjusted to make it more suitable to Chinese populations.

Explanations

(1) If direct cost data from other countries is used for local adjustments, the adjustments need to be made in two areas. First, the treatment regimen should be localized. Second, the unit price should be localized. In this process, the cost items in a treatment regimen from other countries should be subdivided and recombined into the treatment plan commonly used in the clinical practice in China, and then multiply by the unit price in China to determine the costs in China.

(2) The frequency of resource use, the proportion of patients using each service and the service duration should be considered when estimating resource use. Rarely-used medical services and costs that are inconsequential to the result only need to be described in the report. No calculation is required (Drummond et al., 2005).

3.3 Cost Valuation

3.3.1 During cost valuation, the quantity of each type of resource utilization is multiplied by its unit price, and then the costs for all types are summed together to obtain the total cost of the study.

3.3.2 The unit price should be obtained based on the defined unit for resource utilization. If a broad measuring unit is used, there will be an average fee per hospital stay and a daily average fee per hospital stay. If a refined measuring unit is used, there will be a breakdown of prices for each medical service and drug.

3.3.3 It is recommended that unit prices be obtained from the latest pricing information published by the government or another authoritative source, such as the final price for provincial tendering or price determined during the national reimbursement negotiation. If a particular resource item has more than one price in the market and its market share distribution is known, then the average price weighted by the market share could be used. When the information for market share is unavailable, the median value of all known prices could be used.

For treatments, the prices of treatments with the same generic name and formulation but different specifications can be adjusted using the defined daily dose (DDD). The unit prices from multiple resources can be obtained and then the weighted or the median price could be calculated. If the indication and the efficacy of a treatment with different formulations are similar, a price conversion can be performed using the aforementioned method and an average price can then be calculated. If the drug has been launched in China, it is recommended to use the manufacturer's suggested price for analysis. Use of another pricing system should be clearly indicated, and justifications for such sources should be provided.

3.3.4 It is recommended to use the human capital approach (HCA) to perform calculations for the indirect cost incurred during treatment (Liljas, 1998). This approach uses the average salary from the labor market to estimate the productivity loss caused by the disease or early-death based on the assumption that all lost time will be used for production.

Explanations

(1) Theoretically, from a societal perspective, the cost in a pharmacoeconomic evaluation should be evaluated based on the opportunity cost (Luce et al., 1990). Considering its practical challenges, studies usually use the market price of a consumed resource as the standard for a cost calculation because the price tends to be the opportunity cost in a free market. Corresponding adjustments should be made and clarified if there is strong evidence suggesting that market prices have deviated from the valuation costs, for example, price controls or subsidies by the government.

(2) In classifying the cost and benefits of a treatment, double counting or omission should be avoided whenever possible. For example, whether “treatment-caused increase in working hours” should be classified as reduced cost or treatment benefit should be clarified.

Chapter 4: Discounting

Discounting is recommended for studies with a time horizon longer than one-year. Pharmacoeconomic evaluations should discount cost and health outcomes that occur in the future (i.e., converting future cost and health outcomes to the values at baseline).

The same discount rate is recommended for both cost and health outcomes (Smith et al., 2001). It is recommended 5% per year to be used as the discount rate for the base case and a minimum range of 0%–8% in the sensitivity analysis (Liu, 2015). Justifications should be provided if other discount rates are used.

Explanations

(1) The cost and health outcomes of healthcare services usually occur at different time points. In order to compare cost or health outcomes at the same time point, discounting is required. The concept of discounting is based on a population's preference over time, as it is thought that the current cost or health outcome will have greater value than those in the future.

(2) Internationally, applying discounting of health outcomes in non-monetary form remains controversial, which is mainly reflected in three areas. First, should discounting be performed? When quality-adjusted life year (QALY) is used as the index for health outcome, the time preference of the interviewee is already included, so it might lead to double discounting. If no discounting is performed for the health outcome, the results of the analysis will be biased toward alternative treatments with long-term effects, biasing against treatments with short-term effects. Second, should the discount rate for the health outcome remain identical to that of the cost outcome? Third, should the discount rate of the health outcome change with time? Currently, guidelines in many countries suggest that discounting should be performed for the health outcome and use the same rate as that of the cost (Table 1). If no discounting is performed for the health outcome, then a reasonable explanation for this decision should be provided or results with discounting for the health outcome, as well as without discounting for the health outcome, should be reported.

(3) The choice of a discount rate should be able to reflect factors such as varying rates of socioeconomic growth, fluctuations in prices, and the time preferences of consumers. The discount rates recommended by various countries are between 1.5%–5% while the range for the sensitivity analysis is usually between 0%–10% (Table 1). Appropriate adjustments for the discount rate should be made when there is significant inflation or when the increase in medical consumer price index is significantly higher than other consumer goods during the study period.

Table 1. Discount rate for cost and health outcomes, suggested by the pharmacoeconomic evaluation guidelines in selected countries

Country/Region	Cost	Health outcome	Sensitivity analysis
Canada (2017)	1.5%	1.5%	0%, 3%
Germany (2009)	3%	3%	0%, 5%, 7%, 10%
France (2004)	5%	5%	0%, 3%
United Kingdom (2013)	3.5%	3.5%	1.5%
South Korea (2009)	5%	5%	0%, 3%, 7.5%
Japan (2016)	2%	2%	0%-4%
Taiwan (2006)	3%	3%	0%-5%
Australia (2016)	5%	5%	0%, 3.5%
Brazil (2014)	5%	0%-10%	0%-10%

Chapter 5: Health Outcomes

The impact of diseases and interventions on patients can be divided into three categories – economic outcomes, clinical outcomes, and humanistic outcomes (Kozma et al., 1993). In pharmacoeconomic evaluations, economic outcomes are usually classified as costs, while clinical outcomes, defined as changes in clinical measurements, and humanistic outcomes, defined as as changes in patient subjective assessments, which mainly refers to the health-related quality of life (HRQoL), are classified as health outcomes. The broad definition of outcomes includes both costs and health outcomes, while the narrow definition only refers to health outcomes. Health outcomes can be estimated using three categories of measurements, efficacy/effectiveness, utility, and benefit.

5.1 Efficacy/Effectiveness

5.1.1 Efficacy should be based on the best available evidence (i.e., the best evidence among the clinical efficacy studies and effectiveness studies). For new drugs, if available and applicable, clinical efficacy data from a randomized controlled trial (RCT) is preferred; for drugs which have been in the market for years, when it is not possible to obtain new efficacy data or data is not applicable, effectiveness data from a real-world study should be used.

5.1.2 Clinical efficacy data obtained from a systematic review or a meta-analysis of an RCT has a higher level of clinical evidence in evidence-based medicine and is thus preferred. Data from a single clinical trial could be considered if the aforementioned data is unavailable. At the same time, it is necessary to evaluate whether the characteristics of the patient population from the literature are consistent with the characteristics of the population studied.

5.1.3 RCT data based on a Chinese population or international multi-centered RCTs with a Chinese population are preferred for clinical efficacy data. The characteristics of a Chinese subgroup should be described and analyzed where possible, if international multi-centered RCT based on a Chinese population is used. When RCT data based on a Chinese population is unavailable, RCT data based on populations from other countries or regions can be used. However, the rationale for using this data should be clearly explained, potential differences between populations should be emphasized, and a sensitivity analysis should be conducted for key parameters.

5.1.4 Efficacy data from RCTs of direct head-to-head comparisons between the invention group and the control group are preferred. When a direct comparison is lacking, an indirect comparison, or network meta-analysis (NMA) should be applied to indirectly compare the effectiveness of each intervention.

5.1.5 Studies including the final end-points are preferred for use in pharmacoeconomic

evaluations. In the absence of final end-points, key intermediate end-points can be used for analysis, but generally, the final end-point should be predicted by the relationship between the intermediate and final end-points reported in the published studies.

Explanations

(1) Efficacy refers to the performance of an intervention on patients under controlled, optimal conditions, often in the context of RCTs, while effectiveness refers to the performance of an intervention in the real-world clinical practice (i.e., under natural conditions) (Berger et al., 2012). Efficacy comes from a controlled clinical trial where professional researchers administer a drug and monitor results according to a rigorously designed plan. In addition, the study population includes patients who consent to join the trial after being screened using stringent inclusion and exclusion criteria, and remain highly compliant throughout the study. Under these conditions, the effects of confounding factors on efficacy measurements can be better controlled. On the other hand, effectiveness comes from real-world clinical practice. There is heterogeneity among patient populations, variations in patient knowledge about interventions and patient compliance. Moreover, there is heterogeneity among physicians in terms of their level of experience and degree of specialization. In addition, if patients suffer from other diseases or take other treatments at the same time, they will interfere the measurement of effectiveness. In principle, all possible effects of confounding factors should be minimized during the comparison of the intervention group with the control group in order to obtain the true difference in efficacy between the two groups. Therefore, if available, clinical efficacy data is preferred as the evidence for pharmacoeconomic evaluations.

(2) Data from RCTs might be unavailable under certain conditions even though they are preferred in a pharmacoeconomic evaluation. For example, in critical illnesses, rare diseases, and diseases without effective treatment available (e.g. oncology), an intervention may be conditionally approved based on single-arm clinical trials. Similarly, some Chinese traditional medicines on the market have not yet been studied in RCTs. In such cases, clinical effectiveness data can be used as an efficacy measurement in a pharmacoeconomic evaluation.

(3) A systematic review refers to defining the study question in a structured and clear way, using systematic and rigorous methods to search, screen, and evaluate studies, extracting data, analyzing the extracted data, and finally discussing and summarizing the descriptive results. There are quantitative and qualitative systematic reviews. Qualitative systematic reviews are also called systematic evaluation; quantitative systematic reviews are also called meta-analyses. Meta-analysis is a study approach using statistical methods to conduct

systematic, comprehensive, and quantitative analyses based on independent studies with the same objective (Borenstein et al., 2009). Systematic evaluations and meta-analyses both strictly follow the inclusion criteria to select all of the original studies meeting the criteria, summarize the results after evaluating the relevance and quality of each study. This approach reduces potential biases compared to individual clinical trials and is considered as the highest level of secondary evidence in evidence-based medicine (Jr et al., 2007). Therefore, researchers should prioritize evidence generated from a systematic review or meta-analysis that compares the clinical efficacy between the intervention group and the control group, if possible.

(4) A traditional meta-analysis is based on clinical studies including head-to-head comparisons. However, a traditional meta-analysis may not address the problem when one would like to compare efficacy of two interventions in the absence of a head-to-head comparison, or physicians or policy-makers need to select the most optimal regimen among multiple interventions to address a specific clinical problem. In such cases, an NMA may be used. An NMA is a method that expands on a traditional meta-analysis, which can only handle two interventions, to a method which can simultaneously compare multiple interventions and perform an integrated ranking. Its advantage is that it can quantitatively compare different interventions for the same disease and rank them based on an efficacy end-point to select the most optimal regimen (Mills et al., 2013). In the absence of clinical evidence comparing the intervention group and the comparator group directly, results from an NMA that simultaneously evaluates both the intervention group and the control group could be used as evidence for a pharmacoeconomic evaluation.

(5) The effectiveness end-points can be divided into two main categories – intermediate end-points and final end-points. Intermediate end-points are usually obtained from clinical tests, including biomarkers such as blood pressure, blood lipid level, and blood glucose level, etc. Final end-points usually reflect disease-related events that have happened or can be estimated, such as disease-related death or mortality, and life years, etc. Compared with intermediate end-points, the improvement in final end-points is the objective of interventions on health, and can more clearly demonstrate the final outcomes of a certain intervention.

5.2 Utility

5.2.1 Quality-adjusted life year (QALY) is recommended as the measurement incorporating utilities. Survival time and health utility values should be reported before QALY is presented.

5.2.2 The health utility measurements include direct and indirect measurements, with the latter preferred. A direct measurement can be performed when there is no applicable instrument for indirect measurement to obtain health utility values for certain diseases or symptoms. Commonly used health utility instruments in indirect measurement include the EQ-5D, and SF-6D, etc. For children, the EQ-5D-Y is recommended as it is a health utility instrument specifically for children. Commonly used direct measurements include the standard gamble (SG), time trade-off (TTO), and discrete choice experiment (DCE), etc. (Kopeck et al., 2003; Peter et al., 2017).

5.2.3 When an indirect measurement is used, generic utility instruments, such as the 3-level version EQ-5D (EQ-5D-3L), 5-level version EQ-5D (EQ-5D-5L), and SF-6D, etc., are preferred if there is evidence showing that these instruments have good reliability and validity for a target disease. On the other hand, if evidence shows that generic utility instruments are insufficient to assess the major characteristics of the patient population or disease symptoms, then disease-specific utility instruments can be used.

5.2.4 When an indirect measurement is used, it is recommended to measure patients' health-related quality of life directly from the patients. If this is not feasible, the measurement can be completed by the patients' formal caregiver or informal caregiver (e.g., family members), followed by a healthcare provider.

5.2.5 It is recommended that health utilities should be estimated using the value set based on the preference of the general population. If the health utility value set used is obtained based on patients' preference, justifications and possible influences on results of the evaluation must be reported in the study.

5.2.6 It is recommended that health utilities should be estimated using the tariff based on a Chinese population. The tariff for a population in countries or regions with a similar socio-cultural background, or the tariff widely recognized internationally, could be used when there is no tariff based on a Chinese population. However, the applicability of such tariff should be stated with justifications and a sensitivity analysis should be conducted.

5.2.7 Health utility values can only be obtained through direct measurement or a health utility instrument. Values obtained from the measurement using a non-utility instrument cannot be used directly as health utility values. If a mapping algorithm exists, the scores based on a non-utility instrument can be converted to a utility score using this algorithm.

5.2.8 If health utility values cannot be obtained from measurements, health utility values can be obtained from published studies through systematic literature reviews. However, in such cases, sensitivity analyses need to be carried out to compare potential impact of health utility values for the same disease or condition from different publications or different instrument

measurements.

5.2.9 If the disease or the intervention has a significant effect on the caregivers, the health-related quality of life and health utilities of the caregivers can be considered.

Explanations

(1) Utility in a pharmacoeconomic evaluation is based on a patient's or society's preference for the health outcomes resulting from an intervention. QALYs take into account both survival time and quality of life which is equal to the patient's survival time in a specific health state multiplied by the health utility value during that time (quality of life weight). Among these two, survival time is easier to obtain, so the key step in calculating QALY is the measurement of the health utility value. The health utility value is usually a value between 0 and 1, where 0 represents death and 1 represents perfect health. There is also a health utility value for conditions worse than death, which theoretically has no minimum limit. However, in order to avoid the large impact that a negative value has on the average health utility value calculation, it is usually converted to a value between -1 and 0, which is symmetrical to the value range for health conditions better than death. The main advantage of QALY is that it provides a standard metric to compare the health outcomes of different interventions, which can support decision-making among different diseases and different interventions (Neumann et al., 2017). For example, health loss caused by myocardial infarction and that by pneumonia could be compared directly by calculating their respective QALYs.

(2) Indirect measurements of utility refers to the preference-based multi-attribute health state classification system, which includes a health-related quality of life instrument and its corresponding tariff (Neumann et al., 2017). A health-related quality of life instrument can only be used after proved reliability and validity. When the corresponding Chinese version is not available, the instrument should be translated and tested cross-culturally, including forward translation, back translation, cultural adaptation, and cognitive debriefing, etc. (Wild et al., 2005). The Chinese versions of EQ-5D-3L and EQ-5D-5L instruments are available, and the corresponding tariffs for the general population of mainland China have been established (Liu et al., 2014; Luo et al., 2017). The SF-6D includes two versions—V1 and V2. SF-6D V1 only has the Hong Kong version of the instrument and its corresponding tariff (Lam et al., 2008). There is currently not a version for mainland China. SF-6D V2 has a version for mainland China, and its corresponding tariff is being set up.

(3) Direct measurement is a method to directly measure the utility values of individuals

in a particular health state. It is a basic method to measure health utility values and also a tool to obtain the utility value set. Commonly used techniques include SG, TTO, DCE, and the recently developed best-worst scaling (BWS). The SG approach requires an interviewee to make a choice between two hypothetical scenarios: (i) the interviewee is in a confirmed health state; and (ii) the interviewee has a certain probability of being in a relatively good health state (e.g., perfect health) and with a probability of being in a worse health state (e.g., death). The investigator will constantly change the probabilities of the two health states and allow the interviewee to make a choice between the two until the interviewee considers no difference between the two, at which time the interview ends (Dolan et al., 2000). The TTO approach asks the interviewee to make a choice between two scenarios: (i) the interviewee lives in a disease state for a certain period of time; and (ii) the interviewee lives in a better health state (e.g., perfect health) but with a shorter survival time. The investigator will constantly change the survival time in scenario (ii) and allow the interviewee to make a choice between the two scenarios, until the interviewee considers no difference between the two scenarios, at which time the interview ends (Torrance et al., 1972). The DCE approach provides the interviewee with a choice of two different health states and the interviewee can choose one based on his/her preference. However, as the utility value obtained with the DCE approach is not between 1 and 0 (complete health and death), a latent utility conversion needs to be carried out (Lancsar et al., 2008). The BWS approach allows the interviewee to choose both the most and the least important dimensions, levels, or health states from the available options and carry out a utility analysis using the largest difference in preference for each dimension or level. However, this approach has limited applicability as its methodology is not mature (Mühlbacher et al., 2016).

(4) Generic utility instruments, such as EQ-5D and SF-6D, can be used to measure the health utility values of all populations, including healthy populations and patient populations with various diseases. Disease-specific utility instruments are used to measure the health utility values of patients with a specific disease, which include Quality of Life Utility Measure-Core 10 dimensions (QLU-C10D) used to measure the health utility values of oncology patients (King et al., 2018) and the Health Utility for Glaucoma (HUG-5) used in glaucoma, an ophthalmologic disease (Muratov et al., 2018). Generic utility instruments have a broader application compared to disease-specific utility instruments, as the utility values of different populations or health states can be compared directly. However, the sensitivity of a generic utility instrument might be low when measuring health utility of certain diseases due to its non-disease specific characteristics. Therefore, the changes in signs or physical functions related to specific diseases could not be identified and quantified by a generic utility instrument. In such cases, a disease-specific utility instrument should be

used, if feasible.

(5) Preference-based quality of life instruments (e.g., EQ-5D and SF-6D) are designed based on the economic measurement theory whose utility value set can be used to calculate cardinal. Non-utility instruments, such as the MOS Item Short Form the Health Survey (SF-36) and the COPD assessment test (CAT) (Rowen et al., 2009; Hoyle et al., 2015), are designed based on a psychological measurement theory and the values are not ordered, which cannot be used directly used as health utility values. In such cases, mapping can be used to establish the algorithm linking the non-utility instrument and the utility instrument, and thus derive the utility values based on the non-utility instruments. Mapping is a method that uses quantitative modeling to estimate the relationship between the results obtained from non-preference-based measurements and preference-based measurements. A regression model with the result of a non-preference-based instrument as independent variables and the result of a preference-based instrument as dependent variables is built to estimate the health utility values and test the goodness of fit of the regression model (Brazier et al., 2010).

5.3 Benefit

5.3.1 Benefit is the quantification of health outcomes using monetary units. The benefits of a treatment regimen include direct benefit, indirect benefit, and intangible benefit. Direct benefit quantifies the gains from the actual monetary exchanges resulting from an intervention. Indirect benefit quantifies the gains from patients' increased healthy time, or their recovered productivity resulting from an intervention; and intangible benefit quantifies the alleviated or prevented physical and mental distress, as well as the comfort and pleasure post-recovery resulting from an intervention.

5.3.2 Direct benefit quantifies gains obtained from the actual monetary exchanges resulting from an intervention. Double counting should be avoided when measuring the direct benefit, (i.e., avoiding counting the changes of healthcare resources in both cost and health outcomes).

5.3.3 Indirect benefit and intangible benefit quantifies gains for which no actual monetary exchanges occur. Therefore, a calculation usually has to be performed using methods such as human capital approach (HCA) or willingness to pay (WTP). It is preferred to use HCA to calculate the indirect benefit. When the WTP method is used, the assumptions, the survey method, the scope of the benefit measurement, and the description of the questions, etc. in the study should be specified, and the derivation of the monetary values needs to be included in the

report or manuscript.

Explanations

HCA method refers to the use of wages from patients' increased healthy time to represent health outcomes. WTP method values a health state through the quantification of monetary value that patients are willing to sacrifice, under the condition that the overall individual utility value remains constant. The source of preference for WTP should change with the study perspective. Commonly used methods for WTP include contingent valuation and DCE (Neumanm et al., 2017).

Chapter 6: Evaluation Techniques

Pharmacoeconomic evaluation techniques include cost-minimization analysis (CMA), cost-effectiveness analysis (CEA), cost-utility analysis (CUA), and cost-benefit analysis CBA, etc.

6.1 Types of Evaluations

6.1.1 Researchers should choose an appropriate evaluation technique according to the characteristics of the disease and the intervention, the availability of data, the objectives, and the requirements of the study. If possible, a CUA should be conducted. Other techniques, such as a CEA, a CMA, or a CBA, etc. can also be used with the justifications provided.

6.1.2 Researchers can perform evaluations using two or more techniques. They can also use one technique as the main method but also include other techniques, and then compare and analyze the differences in the results from different evaluation techniques.

6.1.3 In order to comprehensively and transparently assess the information related to the research questions, in addition to the main results of CUA, CEA, CMA, or CBA, researchers should also describe other factors that may impact results but have not been systematically included in the analysis or model.

Explanations

(1) CUA is the most commonly used technique in pharmacoeconomic evaluations (Neumann et al., 2015), and it is the preferred method in the Guidelines. The health outcome of CUA is quality-adjusted life year (QALY), which is a standardized and commonly used health outcome. QALY takes into account the effect of the treatment regimen on the patient's survival time and his/her quality of life. It provides a more comprehensive evaluation of health outcomes compared to other outcome measures. CUA can perform comparative analyses of different treatment regimens regardless of whether the clinical outcomes are the same. It should be noted that using this method, different quality of life measurements, instruments, and tariffs will have an impact on the utility values. Therefore, the method for the utility value measurement should be explicitly stated. In some literature and textbooks, CUA is considered to be a subtype of CEA (Gold et al., 1996; Neumann et al., 2016). This is just the difference in terminology. The present Guidelines consider CUA to be an evaluation approach independent of CEA.

(2) CEA is generally used to compare regimens with the same clinical outcome. The measurement unit is usually a physical or natural unit, such as life years (LYs), or symptom-free days. CEA can be used if the outcome of a treatment regimen is only reflected in or

mainly reflected in a particular clinical outcome. The main limitation of a CEA is that comparisons between treatment groups are difficult to carry out if different health outcomes are used. In addition, it is usually difficult to conduct a comprehensive evaluation if the intervention has multiple health outcomes. Furthermore, there is no established threshold value to decide whether an intervention is cost-effective if the CEA shows better effectiveness but higher costs of the intervention group compared to the control group. Therefore, it is not easy for policy-makers to make decisions based on the results.

(3) Theoretically, a CBA is built directly on the theory of welfare economics, so the result can directly support relevant healthcare policies by the policy-makers and can be applied more broadly (Olsen et al., 2001). However, the monetization of health outcomes in a CBA study is mainly achieved using the willingness to pay (WTP) approach. The application of this approach in the health industry is still under development and its methodology has not yet reached a broad consensus (Ziekenfondsraad, 1999; Bala et al., 1999; Olsen et al., 2001). It is best to report the result from the CBA in the form of a net benefit. All steps and methodologies employed in converting health outcomes into monetary values should be explained in the analysis and report, and the major assumptions should be verified with a sensitivity analysis.

(4) CMA is used if evidence shows that important clinical outcomes (e.g., efficacy and safety) of both the intervention group and the control group of a drug are identical or do not have clinical difference. Statistical insignificance or clinical insignificance is acceptable as evidence to show that there is no difference between the clinical outcomes of two treatment regimens.

(5) Just reporting the results of a CUA, CEA, CMA, or CBA in a pharmacoeconomic evaluation may overlook other effect of a treatment on patients. Therefore, the effects on other relevant outcomes should be described after reporting the main results so that important information can be presented more comprehensively and transparently. For example, when evaluating the treatments for non-small cell lung cancer, researchers should also report other relevant clinical information, such as convenience of use for patients and patient compliance that resulted from the variation in the drug delivery route and the frequency of drug administration, in addition to important health outcomes such as QALYs and LYs.

6.2 Incremental Analysis

6.2.1 In CUA and CEA, the decision-making is based on an incremental analysis.

6.2.2 An incremental analysis is the comparison of costs and outcomes between the

intervention and the comparator. If the intervention has a lower cost and better outcome compared to the comparator, it will be the strictly dominant regimen. In contrast, if the intervention has a higher cost and a worse outcome compared to the comparator, it will be the strictly dominated regimen. If the intervention regimen has both a higher cost and a better outcome compared to the comparator, the incremental cost effectiveness ratio (ICER), i.e., the ratio of the difference in costs to the difference in outcomes between the two regimens, needs to be calculated. If the ICER is smaller or equal to the threshold value, the intervention is cost-effective than the comparator; if the ICER is larger than the threshold value, the intervention is not cost-effective compared to the comparator.

6.2.3 In incremental analysis, the WTP threshold value for QALY is recommended to be 1–3 times of the gross domestic product (GDP) per capita.

Explanations

(1) In economic theories, marginal analysis is usually used to guide economic policy. A pharmacoeconomic evaluation is also a sub-discipline built on the basis of welfare economics. To align with its theory, results of a marginal analysis (i.e., the results of an incremental analysis) should be reported in a pharmacoeconomic evaluation.

(2) In evaluating the economy of multiple (three and above) treatment regimens, a pairwise incremental analysis should be conducted for each pair of regimens from low to high costs. The specific procedures are: first, all regimens are ranked according to their cost from low to high; second, the strictly dominated regimen, which is the regimen having a higher cost and worse effectiveness compared to other regimens, is eliminated from the ranking; third, for the remaining regimens, an incremental analysis is sequentially performed for adjacent regimens, and in each analysis, the more cost-effective regimen is retained and used for an incremental analysis with the next regimen in sequence; and finally, the last regimen retained from all incremental analyses is the most cost-effective one among all regimens.

(3) There is no established standard in China for the value of QALY in an incremental analysis. The World Health Organization (WHO) provides a suggestion for economic evaluation using the disability-adjusted life year (DALY) as its outcome index – if $ICER < GDP$ per capita, then the incremental cost is worthwhile; if GDP per capita $< ICER < 3$ times of GDP per capita, then the incremental cost is acceptable; and if $ICER > 3$ times of GDP per capita, then the incremental cost is not worthwhile (WHO, 2010). QALY and DALY are conceptually similar as they both take the length of a patient's survival time and the state of survival under the diseased condition into account (QALY uses health utility value while DALY uses disability index for measurement). As there is currently no suggested

threshold value for QALY in China or by the WHO, the suggestion by the WHO on DALY could be referred for QALY.

(4) The threshold value of QALY may vary across different diseases or health states. For example, the weights of health preference of patients at the end of the life are greater than those of other populations, and the value of time among these patients is also higher than the value of time in another period in the life cycle (Round, 2012). As the research and development (R & D) cost is high but the patient population is small for orphan drugs which can only be used specifically to treat rare diseases, such R & D is only encouraged when the QALY threshold value is set to be higher than treatments for other diseases. Therefore, researchers are encouraged to use a higher threshold value range to perform a sensitivity analysis in the study report.

(5) The economic development across China is unbalanced and the variations in GDP per capita across regions are large. If a pharmacoeconomic evaluation is conducted to serve nation-wide healthcare policies, it is best to use GDP per capita at a national level. On the other hand, if a pharmacoeconomic evaluation is conducted to serve healthcare policies in a particular region, the local GDP per capita should be used.

Chapter 7: Model Analysis

In pharmacoeconomic evaluations, models are often used to compare the economic value of different interventions. Pharmacoeconomic evaluation models usually use methods, such as graph structure, equations, etc. to abstractly simulate the natural progression of a disease and the effect of interventions on its progression. It focuses on the interventions and important clinical events as well as the health state changes and health resource use incurred during the process, all of which will be used in the economic value comparisons of alternative interventions. A typical pharmacoeconomic model usually includes the following information, the incidence or prevalence rate of a disease or health condition, diagnosis and treatments, impact of interventions on the risk of clinical events, patient's survival and quality of life, as well as the costs of interventions.

Explanations

(1) A pharmacoeconomic evaluation model is used to: i. provide a clear framework for decision analysis; ii. integrate data on health outcomes and costs from different sources; iii. conduct the uncertainty analysis; iv. quantify the economic value of the intervention; and v. determine the role of a pharmacotherapy or non-pharmacotherapy in the management of the disease.

(2) In a pharmacoeconomic evaluation, a modeling analysis is necessary under the following conditions (Drummond et al., 2015): i. the economic value of multiple interventions needs to be compared but no randomized controlled trial (RCT) could include all interventions; ii. as the main data source, the RCTs did not collect all data required for a pharmacoeconomic evaluation (e.g., health utility, cost, etc.) or the data is incomplete (e.g., rate of adverse drug reaction [ADR], etc.) and needs to be supplemented with other data sources (e.g., observational study, expert opinions, etc.); iii. the RCTs only include the intermediate end-point/surrogate end-points (e.g., blood pressure, HbA1c, LDL-C, etc.) but not a final end-point (e.g., the incidence rate of a final event, such as acute myocardial infarction, stroke, diabetic retinopathy, diabetic foot disease, etc.); iv. the RCTs have a limited follow-up period that does not allow observations of the long-term effects of the intervention on patients' health outcomes and costs, especially for chronic diseases (e.g., type 2 diabetes, schizophrenia, etc.); v. the procedures and definitions used in the RCTs are significantly different from the ones used in the real-world practice and require adjustment ; for example, the requirement of routine follow-ups in RCTs in order to meet the study objectives, treatment cross-over, which is rare in real-world clinical practice, or the effective or recurrence rate in the RCT defined differently from that in real-world clinical practice; and vi. when there are variations in the data collected across different regions, different clinical practices or patient subgroups of different characteristics, where a

variability analysis or adjustment needs to be performed.

7.1 Model Questions

The description of questions for decision-making is the starting point for model construction and analysis. The description of questions for decision-making in the model should be clear and comprehensive. The model construction should reflect the current clinical practice to ensure that the model has a good external validity. It is suggested to provide clear information on relevant diseases, model objectives, target population, interventions, perspectives, scope of simulation, health and other outcomes, time horizon, etc.

7.2 Choice of Modeling Techniques

7.2.1 Analyses can be performed using different modeling techniques in a pharmacoeconomic evaluation. The choice of modeling technique should be decided based on the characteristics of the questions related to decision-making and researchers should provide the rationale for each choice.

7.2.2 Before choosing the modeling technique, researchers should first understand the clinical characteristics of the disease, such as its onset and symptoms, the characteristics of the disease progression, and the characteristics of clinical treatment pathway, etc. Moreover, the modeling technique should be selected according to the data availability and data requirements by different model types.

Explanations

Many modeling techniques are used in pharmacoeconomic evaluations, and some are still under development. Among these, the decision tree model, the Markov model, the DES, the PSM and the dynamic transmission models are more commonly used.

(1) The decision tree model is a static model to simulate the effect of an intervention on a disease, which usually has a visible tree structure. Its components often include the model structure, inputs, and assumptions. Among these, the model structure is defined by the health state and various nodes. Commonly used nodes include decision nodes, chance nodes, and terminal nodes. Model inputs mainly include probabilities, costs and effectiveness, etc. A decision tree model is applicable for pharmacoeconomic evaluation with a short time horizon, such as acute infection.

(2) The Markov model is a special cyclic decision tree model. It is a dynamic model

that incorporates the timing of clinical events and the implementation of relevant interventions into simulation. The Markov model is a simplified simulation of the continuous changes in a patient's health state in the real world, which is a discrete time point state transition model. In this modeling technique, the time horizon is divided into cycles with equal cycle length. Patients in this model are assumed to be in a finite number of health states/Markov states and each patient in the simulation can only be in one of the states in one given cycle. The initial probabilities are used to define the distribution of a group of patients in different health states at the beginning of the simulation, and the probability of one patient transitioning from one state to another within each cycle is defined by the transition probability matrix. The accumulated total costs and total outcomes during the time horizon are calculated by defining the cost and outcome of each state in each cycle. The Markov model is applicable for the simulation of disease progression among a finite number of health states. Under appropriate assumptions, it is convenient to simulate patients' long-term health state change for chronic disease.

(3) The DES model is a modeling method that could be used to show interactions between individual behaviors, between individuals, between individuals and populations, and between individuals and their environment. The DES model is fairly flexible. The core components of the DES include entity, attribute, event, resource, queue, and time. Entity refers to a subject with specific attributes that will experience events, consume resources and enter the queue. In health care modeling, entities usually refer to a patient with a specific disease. Attribute refers to the characteristics of each individual in the model, such as age, sex, race, health state, events experienced, quality of life, and accumulated cost, etc. Event refers to an incident that may happen to an individual or in an environment, such as the incidence and progression of a disease (e.g., disease onset, occurrence of ADRs, disease progression, etc.); the resource utilization (e.g., hospitalization); the clinical decision (e.g., dose change); and things happening outside of the health care system (e.g., taking leave). Resource refers to the medical services provided for individuals and usually includes physicians, drugs, and surgeries. A queue is formed by individuals when the resources required by the individuals are occupied. Resource and queue only need to be used in a model with limited resources, not in a model with unlimited resources. Compared to a Markov model, a DES is an individual simulation model, without a fixed time point of an event's occurrence, that can remember all clinical events experienced by the simulated individuals, and therefore has more flexibility. However, compared to the Markov model, the DES also requires more detailed, rich and high-quality clinical data to obtain the probability density function to estimate the occurrence time of each discrete event (Zhou et al., 2012). DES is applicable for model construction in the following situations: resource is

limited; there are interactions between individuals or between individuals and their environment; the time of event is not fixed; individual characteristics have a considerable impact on the simulation process; researchers are interested in an individual's experience (Caro et al., 2012).

(4) The PSM is a commonly used model in cost-effective analysis (CEA). Conceptually, this model is similar to the state transition model and uses survival curves to define a series of different health states to estimate costs and outcomes. For example, in the field of economic evaluation of malignant tumor, two curves – the progression-free survival (PFS) curve and the overall survival (OS) curve – commonly reported in clinical trials of malignant tumors are often used to divide patients' health states into pre-progression, post-progression, and death. The model calculates the proportion of patients in each state at a certain time point using the PFS and OS curves. The health outcomes and costs produced within the simulated time range are calculated based there proportions. The PSM is applicable for pharmacoeconomic evaluations of diseases that can be divided into a finite number of health states and require long-term simulation.

(5) The key property of the dynamic transmission model is dynamic. In such modeling technique, the risk of infection at a given time point of the simulated individual is usually not constant, but a function of the number of infected individuals in the population (or environment). If an intervention reduces the number of infected individuals, the risk of uninfected susceptible individuals will also be reduced. Dynamic infectious disease model is mainly used to simulate the occurrence and development of diseases that can be transmitted in the population, and also to simulate the direct and indirect impact of an infectious disease control program on the disease transmission process. (Pitman, 2012)

7.3 Model Development

7.3.1 A model's conceptualization should be based on the disease course and the effect of an intervention on disease progression. Researchers can refer to the existing models for similar diseases in the literature as the basis for the model construction. They should modify and optimize the model structure according to the current healthcare system and the characteristics of the intervention, and provide rationale for the differences from the existing models.

7.3.2 The model structure is normally simplified compared to the actual disease course. Researchers should include in the model structure the events that have considerable variations in patients' health and costs under different interventions. The model structure should not be decided based on data availability; instead, appropriate adjustments should be made based on

the availability of specific data after the initial construction of the model structure.

7.3.3 The description of the model structure should be clear. Use of the model structure diagram to illustrate its description is recommended. Researchers should validate the model structure and state the process and results of the validation.

7.3.4 In the Markov model, the time horizon should be long enough to reflect the overall effect of the intervention on patients' costs and health outcomes; and the model cycle length should be short enough to simulate the occurrence rate of related events more accurately.

7.3.5 Conducting a half cycle correction for the Markov model is recommended especially for models with a long cycle length and a smaller number of cycles.

Explanations

(1) Model structure refers to the logical framework structure that consists of disease progression, clinical treatment pathway, relevant clinical events, and the causal relationship, etc. In a model-based pharmacoeconomic evaluation, the model structure decides the inputs and the level of precision required by the simulation. Different researchers may use different model structures for the same evaluation, which will impact the final results of the model analysis.

(2) Modeling is the moderate simplification of a disease progression and the effect of an intervention in the real-world practice (Caro et al., 2012). The model structure should avoid being too complex. Otherwise, the required inputs might be difficult to obtain and the model transparency will be reduced. The overly complex model makes it difficult for readers and reviewers to understand the entire analysis process of the model, thus not to trust the results. The model structure should also not be oversimplified, either, which may cause the simulation results not to represent the real-world practice. There are no established standards in the moderate simplification, but it is one of the most important skills for researchers in modeling. Researchers should decide whether certain simplifications will omit important costs and health outcomes. For example, there is not an established standard yet on whether an ADR should be considered in the evaluation model of a treatment for a malignant tumor. In such case, the following questions should be considered: what is the occurrence rate of the ADR? Does the ADR require prevention, monitoring and treatment which will lead to extra costs? What are the extra costs caused by the ADR? Will the ADR lead to a significant reduction in patients' quality of life? Will the ADR lead to a significant increase in patients' mortality? In summary, will ignoring the ADR in the model possibly lead to significant changes in the study results?

(3) The construction of a model structure should be conducted before collecting the

model inputs. (Caro et al., 2012). Researchers should not construct a model only based on available data, as this might lead to an unjustified model structure driven by incomplete data. Researchers can make appropriate adjustments to the model structure based on available data from various sources, if it will not lead to significant biases in the results.

(4) To clearly present the model structure, use of a structure diagram is recommended. For example, a tree diagram or a bubble diagram could be displayed for the Markov model. The tree diagram is more comprehensive if the model structure is not too complicated. In the case of a more complicated model, the bubble diagram is recommended to illustrate the main structure, and secondary structures can be omitted in the illustration.

7.4 Data Sources

7.4.1 Researchers should systematically identify, collect, and evaluate the data used in the model, and describe the sources of all inputs in the model and the rationale for these sources.

7.4.2 When there are multiple sources for model inputs, different factors should be considered when choosing the appropriate source, such as the quality of inputs, characteristics of the population in the data source, the country or region for data collection, the practice setting for data collection, the duration of data collection, etc. To the extent possible, these factors should be consistent between the source studies and the model. If necessary, clinical experts should be consulted, and a sensitivity analysis or variability analysis should be conducted.

7.4.3 Clinical data sources should be comparable among different treatment arms in a model. Results from head-to-head clinical trials of the interventions are the preferred data source. When there is no clinical trial for direct comparison, results from an indirect meta-analysis or network meta-analysis of interventions with a common comparator are preferred. When clinical trials with a common comparator are not available, study characteristics of different data sources, such as patient demographic and socioeconomic characteristics at baseline, disease characteristics, treatment setting, treatment duration, type of study design, etc., should be compared and the source studies can be used only when these characteristics are similar.

Explanations

(1) The sources of various inputs in a model.

i. Efficacy data. RCTs are usually the main sources for efficacy data. The clinical trials that are included or excluded for efficacy inputs in the model should be described and explained. The effectiveness data from real-world studies can be used if the data are

applicable for the model being constructed.

ii. Safety data. ADR data can come from RCTs and observational studies after a treatment enters the market. Costs related to the prevention, monitoring, or treatment of ADRs should be included in the model calculation. Some severe ADRs with low incidence rates in RCTs are difficult to accurately estimate due to the limited sample size, so a broader range of data sources should be considered.

iii. Cost data. The cost data in the model should come from the local medical and healthcare system. Commonly used sources for cost data include claims databases, disease registries, hospital medical record, published literature, and estimated costs of standard treatments, etc. Cost data varies among different regions and healthcare systems, so the model should have some flexibility to allow for the adjustment and analysis of this variability.

iv. Health utility value. Health utility value should come from the literature or patient assessment conducted directly by researchers. The results of the preference based measurement of local populations (patient or general population) are the preferred health utility values. The measurement results from non-preference-based measures (e.g., SF-36, QLQ-C30, etc.) cannot be converted into a 'utility value' through the percentile conversion, which should be converted using mapping, as needed.

(2) Surrogate markers. Surrogate markers are also called intermediate end-points, which are usually disease-specific biochemical testing results. When these surrogate markers are used to extrapolate long-term health outcomes and costs in the model, evidence should be provided to establish the relevance and validity of the model or method used for the extrapolation.

(3) Expert opinions. The input values obtained from expert opinions are usually not preferred. Expert opinions should only be considered if literature, databases, and/or original medical records, cannot address the study's needs. Moreover, expert opinions cannot be used to obtain values of key inputs, such as the treatment efficacy, the rate of ADRs, etc. If some input values in the model are obtained from expert opinions, details of the questionnaire for expert survey, such as a survey outline, sampling method, number of physicians surveyed, survey method, and analytical method for the survey data, should be clearly described. It is best to perform a sensitivity analysis for the quantitative inputs obtained from expert opinions. Apart from obtaining certain input values, expert opinions may be used to support the rationale behind the model structure or assumptions, or to validate the face validity of the model.

(4) Quality of input sources. Model inputs can come from various sources, such as systematic literature reviews, meta-analyses, RCTs, observational studies, databases, medical records, expert opinions, unit prices of medical services, and researchers' assumptions or specifications in guidelines (e.g., discount rates), etc. For clinical efficacy and safety inputs, data sources with higher levels of evidence based on the evidence-based medicine classification system are usually selected (Li, 2008). Results from systematic literature reviews/meta-analyses are of the highest level of evidence. However, a pharmacoeconomic model is used to simulate the economic value of an intervention in a real-world setting, so the results of systematic literature reviews/meta-analyses on RCTs are not always the best sources, and the results from high-quality real-world studies with large sample sizes might be more suitable. Therefore, there is no absolute quality level for data sources in pharmacoeconomic models. Researchers should conduct a comprehensive assessment and clearly provide the available data sources as well as the rationale for the selection in the report. When there are multiple data sources with relatively high quality, it is recommended to conduct sensitivity analysis for each source, compare the results, and explain the possible reasons for variations.

(5) Population characteristics of the data source. The population characteristics of the data source in the model should be consistent with the pharmacoeconomic model, including age, standard for disease diagnosis, disease classification and staging, comorbidities/complications, prior treatments, etc. When there are variations in the population characteristics between the data source and the model, a variability analysis can be performed, or the effect of this variability on the final results could be discussed qualitatively.

(6) Country or region of data collection. Local data sources should be selected for model inputs whenever possible. When there is no local data source, whether the data source from other countries and regions can be used depends on the specific circumstances. For clinical efficacy and safety data, if there is only data from clinical studies based on populations from other countries/regions, the data can only be used when there is no obvious or significant variations in genetics, type of microbial infection, lifestyle, etc. Moreover, it is generally recommended to consult with clinical experts to confirm the applicability of such data. If there are data from global multi-center clinical studies, data from Chinese or Asian patients should be used whenever possible, and a sensitivity analysis should be conducted using global multi-center clinical data.

(7) Practice setting where data is collected. There might be variations in many parameters from different practice settings. For example, there might be variations in the

prices of the same medical service or drug, the level of medical skills, the physicians' prescribing habits, the infection control level within the hospital, etc. among medical institutions with different grades, types, and regions. Researchers should describe in detail the characteristics of the practice setting where data is collected, and conduct a variability analysis when necessary.

(8) Direct comparison and indirect comparison. In pharmacoeconomic models, the choice of different data sources has a significant impact on the results and the quality of the study. In clinical trials of direct comparisons, the patients in the intervention group and the control group are subject to the same inclusion and exclusion criteria, the same treatment setting, and similar treatment strategies. There are also randomizations for the treatment groups, so the clinical efficacy, safety, and medical costs are comparable between the intervention group and the comparator group. If there are only clinical trials with a common comparator (e.g., a placebo group) for two regimens, after comparing patient inclusion and exclusion criteria, trial setting, and treatment duration between clinical trials, an indirect meta-analysis or network meta-analysis could be conducted using the common comparator as the bridge which can provide the relative efficacy and safety inputs between different interventions. Sometimes, there are no clinical trials with the common comparator for interventions included in the model. For example, there are only results from single-arm studies or observational studies for one or more interventions. In such cases, cautions should be exercised when using modeling approach for evaluation. Variations in the patient demographics and socioeconomic status at baseline, the disease characteristics, the treatment setting, the treatment duration, the type of study design, etc., among studies might lead to loss of comparability of study results. Significant biases may result if these studies are applied in the same model.

7.5 Model Assumptions

7.5.1 Researchers should describe and explain the assumptions regarding causality, generalizability, scope, structure, and data, etc. in the model. Uncertainty analysis should be performed to assess the key assumptions.

7.5.2 Data extrapolation should be conducted based on valid techniques, which could provide scientific and appropriate evidence and should be tested with a sensitivity analysis.

7.5.3 It is recommended that researchers clearly list all key model assumptions in the study report.

7.6 Model Validation

7.6.1 Researchers should validate the model, including the model's face validity, internal validity, cross validity, external validity, and predictive validity, etc.

7.6.2 All models should be validated for face validity, internal validity and external validity. Assessment of face validity include validating specific questions simulated, structures, inputs, and results with expert opinions. If reviewers think that the face validity of the model is not optimal, the related reasons and subsequent model modifications should be recorded.

7.6.3 Assessment of internal validity ensures that the source, the method for parameter estimation, and the setting of each input are correct and reasonable, and every equation and programming code are complete, correct, and logical.

7.6.4 When assessing external validity, researchers should choose the most appropriate data source to conduct an external validation, and explain the reason for the validation choice.

7.6.5 Researchers should conduct literature searches and comparisons for similar models after constructing their model, and perform a cross validation if possible.

7.6.6 Predictive validation could be performed if researchers can obtain the long-term observational results from the relevant patients in the near future.

7.6.7 Information on model validation methods, process, results, explanation for sub-optimal validity, and adjustment made based on the validation results, should be recorded clearly in the model-based pharmacoeconomic evaluation report.

Explanations

(1) Model validation is an approach to assess the accuracy of the simulation or prediction results by a set of methods (Eddy et al., 2012). Model validation provides important information on whether the analysis and the predicted results in the model are valid and accurate, and is an indispensable step in the model construction and analysis. Model users and decision-makers can fully trust the analytical results of the model only when the model has been validated.

(2) Face validity refers to the extent to which a model, its assumptions, and applications are consistent with the current science and evidence. It is usually a qualitative validation determined subjectively by experts in that field. The contents of face validity mainly include four areas, specific questions, structures, inputs, and results. The validation of specific questions is mainly conducted by clinical or health economics experts with professional experience to determine whether the simulation setting, patients, interventions, outcomes, assumptions, and time horizon in the model are consistent with the questions to

be answered in the pharmacoeconomic evaluation. The validation of model structures is mainly the assessment of whether the model includes all the important effects of the disease progression and the intervention on the patients, whether the relationship between the disease course and the clinical pathway is reasonable, and whether the intervention is feasible or in line with clinical practice. The validation of model inputs is to determine whether the most appropriate input with the highest quality is used. The validation of a model results is to determine whether there is a major conflict between the analytical results from the model and the experience and common sense of clinical experts, and whether the conflict, if any, can be explained.

(3) Internal validity is also called technical validity. It refers to whether the operation of every part of the model is consistent with the researcher's expectations, and whether the calculations in the model are correct. Internal validation is usually a rigorous quantitative test. It mainly includes checking whether the data source for each parameter, parameter estimation method, and parameter setting are correct, and whether every equation or programming code is complete and correct, to ensure that there are no errors in the calculation and that the model is as accurate as possible. The specific methods and approaches of internal validation depend on the model used. The validation method may include data source check for inputs, sensitivity analyses, extreme value analyses, and trace analyses. The validation methods include: i. complete inspection of the inputs, equations, and programming codes in the model by researchers; ii. complete inspection of the model by another researcher; iii. independent model construction and comparisons of the model's results by two researchers; and iv. model construction and model results comparison using two different types of software (e.g., TreeAge Pro and Microsoft Excel).

(4) External validity refers to the comparison between the simulation results of the model and the observed results, which is the direct validation of the model results, and thus the most important type of validity. Three main steps are usually included in an external validation: choosing the data source, model simulation, and comparing results. Data sources used in the external validation could be demographic data, epidemiologic study results, observational study results, RCT results, claims data, electronic medical record data, disease management data, etc. Sometimes, the study uses the data source for key inputs as the source for validation, which is called dependent validation. Validation using other external data sources is an independent validation. When choosing the data source for validation, sources with the same or a similar simulation setting, disease, intervention, follow-up schedule, and outcome definition as those in the model should be used. Usually, the data source used in an external validation only covers part of the simulated results of the model, so the results of the model can only be partially validated. For example, PFS and

OS are often validated in an oncology model.

(5) Cross validity is also sometimes known as external consistency, comparative modeling, external convergence testing, convergent validity, or model corroboration. It refers to the comparison between the results of the current model vs other existing models addressing the same or similar questions. The results of cross validation are influenced by a number of factors, including model structure, model assumption, simulated clinical setting, data sources for model inputs, etc. Sometimes it is difficult to determine the reason for large variations between the model results, and this will reduce the value of cross validation. However, the reliability of the model could increase if two high-quality models that address the same or similar question produce similar simulation results.

(6) Predictive validity is a method comparing the long-term observed results in the study with the simulated long-term results predicted by the model. Predictive validity is the most effective and simple method of model validation, but its biggest challenge is that researchers will not always follow up with patients for a long term in a clinical trial or an observational study.

(7) Evaluation of the model validation results. Model validation is a complex task, and there is no standardized minimum requirement for a model validity. Therefore, researchers should evaluate comprehensively whether the model is valid enough based on their experience and expert opinions.

7.7 Model Transparency

7.7.1 A model should have sufficient transparency in order to allow readers, reviewers, and healthcare policy-makers to evaluate the credibility of the model and determine whether the model results are suitable to inform the decision-making process given the specific setting faced by the decision-makers.

7.7.2 Researchers should use common and simple software for model construction and analysis, such as TreeAge Pro or Microsoft Excel.

7.7.3 Unless necessary, the technical complexity of the model should be minimized. Variables, equations, or programming codes that are not visible or modifiable should be minimized so that readers and users can easily understand the model.

7.7.4 The methods to increase model transparency mainly include two components. One is that researchers should describe the simulated questions, model structure, model assumptions, data sources and estimation methods, analysis methods, analysis results, etc. clearly to increase

model transparency; the second is that a standalone model specification document should be generated for each model.

Explanations

(1) Transparency refers to the extent to which others can review the structure, equations, parameter values, and assumptions of a constructed model.

(2) There are two main reasons to increase model transparency. First is to allow non-professional readers to understand the research questions of the simulation, the logic, and the operation of the model. Second is to allow professional model reviewers or peers to reproduce the model's results according to the report or model documentation, or further adjust or adapt the model.

(3) There are some debates on the issue of improving model transparency. On one hand, a model will only obtain high recognition when it is recognized by other peers or reviewers, or when the scientific results can be reproduced, so increased transparency is required. On the other hand, the construction of a disease model requires a large investment of time, human resources, and other resources. Therefore, the modeler has certain expectations on the intellectual property rights of the model and may not want all model documentation to be publicly available. To balance these two needs, researchers should develop two model specifications. One specification should be a non-technical specification that describes general information, such as the modeling technique, objective, source of funding, model structure, model assumptions, model inputs, model validation methods and results, model analytical methods and results, etc. from a non-technical perspective. This documentation should be publicly available and can help all readers understand the basic characteristics of the model from its general content and logic. The other documentation should be a technical specification, describing the entire process of the model construction and details of the analysis, that can help peers or reviewers with the appropriate expertise to evaluate the model or reproduce/update the results. This documentation does not have to be publicly available, but could be provided under the researcher's agreement or by signing a confidentiality agreement.

7.8 Model Localization or Adaptation

7.8.1 Before model localization and adaptation, researchers should first evaluate the validity of the original model. If the original model is not validated, researchers should at the minimum invite experts to assess face validity of the model. Only models with face validity can be localized or adapted.

7.8.2 In order to conduct model localization or adaptation, researchers should try to obtain complete documentation, including specifications of the original model.

7.8.3 Researchers need to conduct localization and adaptation in areas that include perspective of the model, locally available treatments, patients' characteristics, cost data, guidelines for clinical treatment and clinical practice, epidemiological data, health utility value, clinical efficacy and safety data, discount rate, and quality-adjusted life year (QALY) threshold value, etc., according to the specific study questions, systematic literature searches, and local pharmacoeconomic evaluation guidelines.

Explanations

(1) The main problem to be solved during model localization or adaptation is the generalizability of the model, i.e., the modification of the model using cross-regional clinical data and cost data, etc. to allow the study results of the model to become applicable for local healthcare decisions (Mullins et al., 2014)

(2) During model localization and adaptation, attention should be paid to whether the original model's healthcare setting, patient characteristics, and modeling methods are consistent with relevant local characteristics, including epidemiological data, death rate, disease severity, demographic and socioeconomic status, risk factor, existing treatment regimens, discount rate, price of health care resource, health utility value, and clinical pathways, etc.

Chapter 8: Variability and Uncertainty

8.1 Variability Analysis

Variability refers to the variations in parameters that are confirmed to affect the evaluation results and are related to the variation in treatment setting. Variability cannot be eliminated completely. The causes of variability may be from the variations across regions and settings (e.g., variations in treatment regimens, clinical practice, and payment method, etc.), or heterogeneity among different patient subgroups.

Explanations

A sensitivity analysis or scenario analysis can be conducted to assess the variability resulting from differences in regions and settings. Variability caused by patient heterogeneity should be handled by dividing patients into smaller but more homogenous subgroups, as the cost-effectiveness evaluation may vary by subgroups. If it is a modeling study, different subgroups should be analyzed separately in a model or different models should be constructed for different subgroups (CADTH, 2017).

8.2 Subjects of Uncertainty Analysis

8.2.1 Researchers should conduct a comprehensive analysis of different types of uncertainties in a pharmacoeconomic evaluation, including uncertainties in methodology, inputs, and model, etc.

8.2.2 Theoretically, all inputs and assumptions in pharmacoeconomic evaluations should be included as candidate variables in a sensitivity analysis.

Explanations

(1) Source of uncertainty. The uncertainty in a pharmacoeconomic evaluation can be attributed to three sources: methodology, inputs, and the model. First, many methodological aspects in a pharmacoeconomic evaluation have not been standardized (e.g., study design, study perspective, measurement and estimation of cost and treatment effectiveness, discounting, statistical analysis, presentation of the results, etc.); second, there is substantial uncertainty in inputs, which is usually caused by sampling errors, such as sample size, representativeness of the sample, etc.; and lastly, there is uncertainty in the analytical method of the model, the model structure, the model assumption, the data source, and the extrapolation of the evaluation results to a broader population, etc., and subjectivity in reporting and interpreting the results from an evaluation (Briggs et al., 2001; Drummond et al., 2015).

(2) There is uncertainty in every stage of a pharmacoeconomic evaluation. According

to its property, uncertainty can be categorized to that related to data and that related to the evaluation process. Uncertainty in data is usually caused by sampling errors. That is, uncertainty in an estimation, based on the sample population, is associated with the level of uncertainty in the sample size. Uncertainty in the evaluation process can be further categorized into three types: uncertainty in the extrapolation of evaluation results, such as extrapolating a clinical result (e.g., decrease in cholesterol level) to a health outcome (e.g., decrease in morbidity or mortality rate); uncertainty in the generalization of evaluation results, such as generalizing the evaluation results from a specific setting to other study settings and populations; and uncertainty in the choice of analytical method, for example, whether indirect cost should be included in the analysis and assumptions for a pharmacoeconomic model.

8.3 Methods of Uncertainty Analysis

8.3.1 Uncertainties in methodology and modeling are often assessed with a scenario analysis. Researchers should clearly define the analytical methods and assumptions in different scenarios and explain the differences in results of different scenarios with justifications.

8.3.2 Uncertainty in inputs can be assessed using a deterministic sensitivity analysis (DSA), such as a one-way sensitivity analysis, multi-way sensitivity analysis, and extreme value analysis, or probabilistic sensitivity analysis (PSA) using Monte Carlo simulation. Results from both DSA and PSA should be reported in pharmacoeconomic evaluations.

8.3.3 In a DSA, a one-way sensitivity analysis is often conducted.

8.3.4 In a DSA, sufficient rationale should be provided to determine the ranges of variation for inputs. Commonly used references include 95% confidence intervals, maximum and minimum of input estimates reported in the literature, or high and low estimates from other similar studies. Some input ranges may be generated from different regions or hospitals. For example, the bidding price of a drug has high and low values across different regions in a country. If there is no available reference, the range could be set arbitrarily but its limitations and future improvements should be clarified.

8.3.5 In a PSA, it is recommended to refer to the variation of an input of the same or similar kind in the literature when only the point estimate of an input is available from the literature without information on its distribution or range.

8.3.6 In a PSA, researchers should include as many parameters as possible. The probability distribution, the distribution parameters, and the number of Monte Carlo iterations should be described and justified.

8.3.7 In a pharmacoeconomic evaluation alongside clinical trials or an observational study, individual patient-level data can be obtained. In such cases, a non-parametric bootstrapping method can be used to analyze sampling uncertainty.

Explanations

(1) A sensitivity analysis is the primary approach to handle uncertainties in a pharmacoeconomic evaluation and is used to determine the sensitivity of a system to changes in one or more specific input values, such as drug price, length of stay, response rate, and discount rate, etc. According to how the input value is determined, the sensitivity analysis can be divided into DSA and PSA (Drummond et al., 2015). The DSA includes one-way sensitivity analysis, multi-way sensitivity analysis, threshold analysis, extreme value analysis, and scenario analysis, etc.

Table 8.1 Recommended methods to handle variability and uncertainty

Type	Reason	Method
Variability	Variation in treatment approaches in different regions and settings	Sensitivity Analysis, Scenario Analysis
	Patient heterogeneity	Stratified Analysis/Subgroup Analysis
Uncertainty	Uncertainty in methodology and model: analytical method, model structure, assumption, data source	DSA: One-way Sensitivity Analysis, Multi-way Sensitivity Analysis, Threshold Analysis, Extreme Value Analysis, Scenario Analysis, and Model Validity Analysis
	Inputs uncertainty	DSA: One-way Sensitivity Analysis, Multi-way Sensitivity Analysis, Threshold Analysis, and Extreme Value Analysis; PSA: Monte Carlo Simulation, and Non-parametric Bootstrapping method

(2) In general, both DSA and PSA should be conducted in a pharmacoeconomic evaluation. Usually, a DSA can only simultaneously analyze the impact of a limited number of input values on results, while a PSA can analyze the impact from all model inputs together. There are many ways of presenting the results of a PSA. The commonly used and easy to

understand ways include cost-effectiveness scatter plot, a cost-effectiveness acceptability curve (CEAC), and cost-effectiveness acceptability frontier (CEAF), etc. (Fenwick et al., 2001; Drummond et al., 2015; Briggs et al., 2006).

(3) There are many ways to minimize uncertainty when designing a study. Analytical methods used to handle uncertainties. For example, calculating 95% confidence intervals to address sampling errors; conducting sensitivity analysis by addressing uncertainty in assumptions and data collection (Wakker et al., 1995).

(4) In a pharmacoeconomic evaluation or an observational study alongside clinical trials, a bootstrapping approach is the most commonly used approach to estimate the confidence interval of an incremental cost effectiveness ratio (ICER). As the distribution for cost and cost-effectiveness data is usually skewed, calculating the 95% confidence interval directly will lead to bias. There are at least three purposes for using a confidence interval to analyze uncertainties, i.e., to assess the directions and values of cost, effectiveness and ICER, and to inform the policy-makers of the reliability of the evaluation results.

8.4 Presentation and Interpretation

8.4.1 When there are multiple uncertain factors, a tornado diagram can be used to present the results of a deterministic one-way sensitivity analysis, which can clearly illustrate the impact of each uncertain factor on the result. It should be noted that information such as the outcome measure used in the axis, analytical variables and changes in the optimal treatment option in the tornado diagram should be explained as necessary.

8.4.2 Use of CEAC, or the cost-effectiveness scatter plot is recommended when presenting the results of a PSA. In a pharmacoeconomic evaluation alongside clinical trials or an observational study, it is best to use results derived from a non-parametric bootstrapping approach to draw the CEAC or the cost-effectiveness scatter plot.

8.4.3 When using the 95% confidence interval to indicate uncertainty for the ICER, attention should be given to the problem of the discontinuous distribution of ICER.

8.4.4 It should be noted that a sensitivity analysis results should be interpreted together with the base-case results. Both results from the sensitivity analysis and the base-case analysis are equally important; thus researchers should avoid drawing conclusions based mainly on the base-case results. Even though a positive result is obtained from the base-case analysis, it may be reversed in a sensitivity analysis.

Explanations

(1) The ICER is a ratio and has a discontinuous distribution. Therefore, when using a 95% confidence interval to indicate the uncertainty around ICER, sometimes the upper limit can be a negative value while the mean and the lower limit are positive values, which differs from estimating the range for continuous values (Wu et al., 2006). In such cases, researchers should explain the differences to avoid misleading information.

Chapter 9: Equity

In any pharmacoeconomic evaluation that could influence resource allocation, equity is a problem that policy-makers frequently face. When possible, equity in the evaluation results from a base-case analysis should be assessed. Globally, there are variations among countries and institutions regarding the definition of equity. In pharmacoeconomic evaluations, equity means that the values of all lives, life years (LYs) and quality-adjusted life years (QALYs) affected by an intervention are (assumed) to be equivalent, regardless of age, sex, or social status of the individuals in the target population (Lakdawalla et al., 2018; Round et al., 2018).

There are two methods to address equity issues. The first method is to perform a sensitivity analysis to illustrate the effect of equity assumption on the result (Woodley et al., 2015). The second method is to perform a subgroup analysis using pre-specified factors in order to compare equity-related characteristics between subgroups that benefit more versus less from the intervention, such as age, sex, race/ethnicity, region, socioeconomic status, health status, and other population characteristics. (Sun et al., 2010). When the effectiveness varies among subgroups, and it is possible to selectively implement the intervention in different subgroups, the cost-effectiveness results should be reported for each subgroup.

Explanations

(1) Equity can be divided into equity related to need and equity related to the access to services (Hauck et al., 2004). For example, the following areas might influence the equity of the study results: i. the equity assumptions in the study design, such as the target population, the comparators, the time horizon, the discounting, the key clinical assumptions, and the expert opinions; ii. the equity assumptions in the analytical method, such as the assumptions of inputs in model design; iii. the equity assumptions related to accessibility and availability of medical services. A comprehensive discussion of equity is beyond the scope of the pharmacoeconomics guidelines. However, the importance of equity encourages researchers and policy-makers to consider the timing of the implementation of a pharmacoeconomic evaluation when deciding the allocation of medical resources, and to separate an equity analysis from an analysis of cost-effectiveness.

(2) Researchers should clearly understand the implicit equity assumptions in the study. For example, one LY of each patient is assumed to be given the same weight, regardless of the patients' age, sex, and socioeconomic status. As another example, a large population receiving a slight increase in QALYs is assumed to have the same expected value as a small population receiving a large increase in QALYs.

(3) When researchers need to perform a sensitivity analysis, they should choose inputs for the sensitivity analysis according to the explicit or implicit assumptions. For example, in

evaluating the productivity gain from a new treatment for ovarian cancer, if the human capital approach is used, the value derived from using the wage rate of females will be lower than that derived by using the overall average wage rate including males. Therefore, the overall wage rate should be used in the sensitivity analysis. Additionally, using a high discount rate could derive a lower value for future time, so a lower discount rate should be used for the sensitivity analysis.

(4) When researchers conduct a subgroup analysis for equity, several subgroup assumptions to be analyzed should be determined a priori (Sun et al., 2010; Sun et al., 2014). The results of a subgroup analysis should be assessed for interaction effects. If the interaction effect is statistically significant, the cost-effectiveness estimates and the associated confidence intervals in different subgroups should be reported separately. The subgroup results should also be included in the executive summary, results, and discussion sections of the report, and the impacts on equity should be discussed.

Chapter 10: Generalizability

In a pharmacoeconomic evaluation, generalizability refers to “whether the results from one setting or population can be applied or extrapolated to another setting or population (Willke, 2003)”. It is the question that researchers must consider when interpreting and generalizing the study results. When data (economic, clinical, and humanistic data) is generated based on other healthcare settings (including other countries, regions or healthcare systems), researchers need to assess its suitability for the healthcare setting in the current study. If data adjustment according to the current healthcare setting is required, the method used for adjustment should be described and its suitability should be explained.

10.1 Epidemiology data often has geographic variations. When only national epidemiology data is available, researchers should assess whether applying that data may lead to bias in the current study. If bias exists, researchers should quantify the bias to the extent possible.

10.2 The applicability of clinical data is an important consideration when researchers apply or generalize the study results. Researchers should clarify the differences between efficacy and effectiveness, especially when a pharmacoeconomic evaluation is conducted using efficacy data from phase III clinical trials. If a pharmacoeconomic evaluation uses data from international multi-center studies, it should be first considered whether to use pooled data from multiple countries, or to use data from a particular region or country that best fits the setting of the decision-maker.

10.3 When applying cost data for interventions obtained from a certain country or region, researchers need to pay attention to the variations in cost data across different countries or regions, including medical resource use patterns, unit costs, economic factors, and other factors that may lead to variations across regions and settings.

10.4 The locations and grades (e.g., tertiary, secondary and primary) of healthcare institutes or related organizations will impact the generalizability of the data. Prices and clinical practice in different countries, regions or healthcare institutes (including medical practitioners as individuals, healthcare systems as a whole) vary considerably, which will impact the generalizability of the study.

Explanations

(1) Generalizability, also known as transferability, transportability, external validity, relevance, or applicability, is a challenge in pharmacoeconomic evaluations. The major issue regarding generalizability is whether costs and effectiveness of interventions vary across different settings or populations, for example, different countries or regions. The discussion of generalizability often relates to two levels: i. transferability, which means

application of original study data (e.g., cost and effectiveness, etc.) obtained from a certain setting or population into a study of another setting or population; and ii. generalizability, which means the extension of the pharmacoeconomic evaluation results of a certain country or region into another country or region. Given the limited availability of local inputs for pharmacoeconomic research in China, the current status of applying relevant inputs from other countries in local studies may continue for a while. Therefore, it is necessary to consider “generalizability” in pharmacoeconomic evaluations in China.

(2) The other major issue regarding generalizability is whether the efficacy data from clinical trials can reflect the therapeutic effects of the intervention in the real-world practice. Since the sample inclusion and exclusion criteria for clinical trials are stringent and specific, the generalizability of results will be affected to certain extent. Unlike efficacy, effectiveness is the treatment outcomes of an intervention used in clinical practice (Berger et al., 2012). That is to say, the outcomes are obtained from the treatments prescribed by physicians in various specialties based on individual patients’ needs in a relatively large patient population with considerable variations in disease severity and characteristics. Most researchers believe that clinical applicability/comparability is the primary consideration in study generalization.

(3) Cost data varies across countries or regions, reflecting differences in medical resource use patterns and related unit costs. For example, in chemotherapy for malignant tumors, some regions provide the treatment mainly in inpatient settings while other regions, mainly in outpatient setting. This lead to differences in the cost structure. In addition to the direct cost, the indirect cost varies with the economic level in different countries. Especially when the human capital approach is used, the gross domestic product per capita in developed countries can be several times or even tens of times higher compared to that in developing countries, which will lead to large differences in indirect costs among different countries.

(4) If study uses data from an international multi-center study, i.e., data is collected from multiple countries, how to integrate data from different countries in the analysis will become a key issue. Although differences in treatments across countries may be avoided by practicing a uniform intervention among different study centers, differences in other factors (e.g., race, treatment history, etc.) will still lead to in the question about the comparability of data obtained from different countries. Variability among data obtained from different countries may be hidden, if they are standardized directly. Therefore, when integrating data from multi-center studies, one should not only transfer data from one country to another, but also reflect the variability among data from different countries. Generally speaking, outcome data from different countries has good comparability, while cost data varies

greatly. Currently, researchers are trying to establish methods for handling issues related to economic data from multi-country multi-center trials, including multi-level modeling, trial-based Bayesian model, multi-parametric regression analysis, and net benefit regression analysis.

Chapter 11: Budget Impact Analysis

Budget impact analysis (BIA) is the evaluation of the impact on expenditure of a healthcare system after a new intervention enters a healthcare system (e.g., the reimbursement list). The results are generally calculated by comparing two possible scenarios, with and without entry. A BIA provides important evidence to decision makers during the market access of a new intervention. As evidence submitted to the decision-makers or for publication, BIA can be a stand-alone document or study, or as a supplemental evidence to the pharmacoeconomic evaluation. In addition, BIA plays an important role in price negotiation, volume-based procurement, and risk sharing agreement, etc.

Explanations

(1) A BIA aims to estimate the impact of including a new intervention on a healthcare system's expenditures (Sullivan et al., 2014). In general, the narrowly defined pharmacoeconomic evaluation refers to the assessment of the difference in economic efficiency between different interventions, that is, cost-effectiveness. A broadly defined pharmacoeconomic evaluation can include the BIA, which estimates the affordability of the budget after the inclusion of a new intervention. In addition, the BIA is also referred as Assessing Resource Impact in some guidelines and literature (NICE, 2017).

(2) Differences and connections between a pharmacoeconomic evaluation and a BIA. Both can serve as separate components of a comprehensive pharmacoeconomic evaluation of an intervention. A pharmacoeconomic evaluation determines the e intervention with economic efficiency by measuring its costs and outcomes, while the BIA determines whether the intervention can be included in the reimbursement list or whether adjustment in reimbursement list is needed by estimating the budget from the decision maker's perspective and evaluate the affordability for the new intervention based on the estimated budget. In addition, a pharmacoeconomic evaluation and a BIA have different requirements in study design, including study perspective, target population, study type, time horizon, cost estimation, etc.

(3) Application of BIA in the reimbursement access of interventions. Generally, if an intervention is considered as not cost-effective based on a pharmacoeconomic evaluation, no BIA is needed, and the corresponding intervention should not be included in the reimbursement list. If a pharmacoeconomic evaluation shows that a new intervention is cost-effective, and the BIA considers it affordable, the decision-makers should consider including the new intervention in the reimbursement list. If a pharmacoeconomic evaluation shows that a new intervention is cost-effective, but the BIA shows that it is not affordable, the decision-makers need to discuss the approaches of including it into the reimbursement list. For example, requesting the suppliers to reduce price through price negotiations or "volume-

based pricing” mechanism, or provide a safety net to the insurance fund through risk-sharing agreements.

(4) The healthcare-related policies in China are relatively complex. Therefore, in a BIA, certain characteristics should be clarified, which include but are not limited to the following: i. actual reimbursement rate, ii. payment methods, e.g., fee-for-service, per capita, or bundled payment by DRG, iii. method for global payment, iv. deductibles and reimbursement cap, v. restrictions on proportion of drug expenditures, quota for drugs selected by central volume-based procurement, and other relevant policies.

(5) A BIA can set up a decision threshold or define the ranges of different types of expenditure in order to assess the budget impact of a new intervention is considered as cost saving, minor impact, medium impact or significant impact. The decision makers can set up or adjust the decision threshold based on their own budget and thus improve the transparency of the decisions.

11.1 Perspective

The perspective of a BIA is normally a budget holder’s perspective. Depending on the need from the decision-makers, the perspective can be defined as different levels of government payers from national to local level, commercial insurance organizations, or a medical institution of a certain type in a certain area. Selection of a study perspective will affect the range of cost estimation. Other than the budget holder’s perspective, researchers can present the results from other broader perspectives. When designing a BIA study, the characteristics of the healthcare system, in which the budget and payment decisions will be made, should be considered first to ensure that the study content and results meet the requirements in the decision-making practice.

11.2 Target Population

A BIA should clearly define the target population. The target population size should be estimated based on the inclusion and exclusion criteria and other applicable patient characteristics. The estimation should use a step by step approach based on the corresponding epidemiology data. Estimation of a target population size should be precisely estimated by considering patient access in a healthcare system. For example, based on relevant healthcare insurance policies, only eligible patients can get their drug costs reimbursed, thereby affecting the expenditure of the healthcare insurance fund.

Explanations

(1) Target population refers to all patients who are eligible for the evaluated intervention under the corresponding reimbursement policy during a specified time period. The target population of a BIA is not a static group, but a dynamic population that varies with incidence, cure, prognosis, and death. In addition, the adoption and costs of the evaluated intervention may be affected by patient characteristics such as disease severity, disease stage, complications, age, sex, and ethnicity, etc. Researchers should consider whether it is necessary to conduct subgroup analyses.

(2) Incidence rate and prevalence rate. Prevalence rate is the proportion of cases (including new and existing) in the total population during a certain period. Incidence rate is the frequency of new cases in a population during a certain time period. When used to estimate the target population size, the two definitions should be differentiated. There are differences in various BIAs in terms of using primarily incidence or prevalence rate to estimate the target population size. The decision should be made on a case to case basis.

11.3 Scenarios

A BIA will, at a minimum, analyze two scenarios, without entry scenario and with entry scenario, respectively. Both scenarios should consider expected market changes, including entries of other interventions into the market, withdrawal of similar drugs from the market, and possible replacement treatments.

Explanations

Scenarios in BIAs refer to the market compositions corresponding to the market access status of the evaluated intervention. Generally, the “without entry scenario” is the situation where the new intervention is not included in the reimbursement list of budget holders, while the “with entry scenario” is the situation where the new intervention is included in the reimbursement list of budget holders.

11.4 Time Horizon

The time horizon of a BIA is usually 3 to 5 years. Study results should be reported for each year according to the budget cycle and, if necessary, as the total impact within the time horizon. Moreover, in a BIA, because the period when future costs are estimated corresponds to the budget cycle, discounting is not recommended.

11.5 Market Share

In a BIA, researchers should report market shares in the two scenarios, i.e., without entry scenario and with entry scenario. The entry of a new intervention may result in three types of market share changes: (1) substitution, where the new intervention replaces the shares of one or more existing interventions at a certain proportion; (2) combination, where the new intervention is used in combination with existing interventions; (3) expansion, where the new intervention covers patients who previously have no effective treatments or have already stopped using existing interventions, resulting in a net growth of the treated population. Forecasting changes in market shares is a very important and challenging component when different scenarios are compared. Researchers should ensure the transparency of the prediction method and should describe in detail the assumptions, the reference data, and the selected prediction model.

Explanations

(1) Market share refers to the proportion of patients using each intervention in the target population.

(2) The market shares in the “without entry scenario” refers to the proportion of patients using each existing intervention in the target population, which is generally obtained from real-world studies.

(3) The market shares in the “with entry scenario” predict the market share of the new intervention in the target population, and the market shares of all interventions in the target population based on specific assumptions.

(5) Off-label use. Even if the evaluated intervention has not been granted access, off-label use of the intervention is a possibility in the current market. In such case, when estimating the market shares in the “without entry scenario”, the off-label in the target population should be included. The intention is not to encourage the off-label use but more objectively reflect the current market status. On the other hand, when estimating the market shares in the “with entry scenario”, off-label use of a new intervention is normally not included unless specially requested by the decision makers.

11.6 Costs

Costs in a BIA include two parts. The first part is the costs of the intervention itself, which can be calculated based on the unit price of the intervention and the amount of use in the target population under different scenarios. The second part is the impact of an intervention on other

costs, which consists of two components: condition-related costs and indirect costs. In a BIA, selection of the scope of cost estimation should strictly follow the perspectives of decision makers, especially the impact of the intervention on other costs. Among them, indirect costs are usually not relevant from the perspectives of decision makers, and thus recommended not to include in the costs in a BIA. If necessary, consultations with decision makers should be conducted to clarify the scope for costs.

Explanations

(1) Condition-related costs. The entry of a new intervention may lead to health condition changes like disease symptoms, duration of disease, disease outcomes, or the rate of disease progression, which in turn affect the use of related health services, including costs of monitoring, costs of disease progression, costs of adverse drug reactions, and costs of nursing care.

(2) Indirect costs. The costs are derived as a results of the impact of a new intervention entry on productivity, social services, and other costs.

11.7 Computing Framework

The computing framework for a BIA is generally presented in the form of a spreadsheet. To the extent possible, researchers should provide a cost calculator to decision makers, clearly listing each cost component. In some complex situations, for example, when important variables, such as the target population size, the composition of patients with different disease severity, the combination of interventions, cannot be directly calculated, it is often necessary to introduce some modeling methods based on cohort simulation or individual simulation to calculate the values of relevant important variables in different scenarios. The principle of model selection can refer to the recommended principles in the modeling section of this Guidelines.

Explanations

When describing the computing framework and inputs, researchers should ensure transparency and reliability of data, e.g., presenting the data using Microsoft Excel.

11.8 Uncertainty and Scenario Analyses

Uncertainty of the BIA includes the uncertainty of the model structure and the uncertainty of the input parameters. The uncertainty of the model structure mainly comes from the

uncertainty of the study method and the modeling assumptions (Shiroiwa et al., 2017), and the uncertainty of the input parameters results from the limitation of data availability, the variability of data and the limitations of data sources. The non-negligible uncertainty should be quantified. Researchers should record and describe the decisions related to the selection of model structure and underlying assumptions, assess uncertainty through scenario analyses (by changing structural assumptions) and one-way/multi-way sensitivity analyses (by changing selected parameter input values), and, if necessary, conduct probabilistic sensitivity analyses (Foroutan et al., 2019).

11.9 Validation

Key analytical process and inputs should also be validated. The validation consists of three main components: (1) face validity assessment, in which consultation with relevant decision-makers is conducted to ensure that the computing framework, the content, and relevant decision requirements are properly reflected; (2) technical validation, which validates whether the model operates as expected and whether the logic and operation are correctly implemented; (3) external validation, which validates whether the model correctly reflects the real world and has the ability to replicate reality within the defined scope (IQWiG, 2009). In addition, if possible, the actual payment amount should be compared with the estimate for the starting year in the BIA. After the new intervention is granted entry, it is also recommended to continue collecting data and compare it with the estimates obtained from the BIA. Although the comparison cannot affect the entry decision for the current intervention, it provides important reference for future decision-making and studies.

11.10 Data Sources and Hierarchy

Besides a proper study design, whether a BIA can effectively support decision-making largely depends on the quality of data. Possible data sources include real-world data that is consistent with the perspective of decision makers, clinical trials, reference data from other countries or regions, health statistics provided by the government, market research data and expert interviews or survey data.

Data sources, should be selected based on the perspective of the decision-makers and the data should be best suited to address the questions relevant to decision making. It is recommended to prioritize data with the highest quality in the same region and the same population. Real-world data should be prioritized over clinical trial data. If data in the same region and the same population is unavailable, data from similar regions or populations should be used or supplemented from expert surveys. When different sources are available, appropriate data source should be selected after the evaluation of applicability of the data. If the data quality of different sources is comparable, results from the scenario analysis should be presented.

Explanations

(1) Real-world data with consistent perspective as the one of decision-makers can provide utilization and costs data close to real-world practice in a BIA. These studies can be registry analyses or database analyses;

(2) When data from real-world studies are not available, extrapolation can be considered based on data from clinical trials in populations consistent with the target population that decision-makers are interested in. For example, in oncology, the progression-free survival period can be used as a proxy for the annual average days on treatment, while in chronic diseases, the average treatment time can be used as a proxy.

(3) Reference data from other countries or regions refer to the data generated for similar populations or similar treatment patterns in other countries and regions, e.g., usage rate of an intervention, quantity of usage and compliance, etc.

(4) Official health statistics data generally include health statistics from multiple sources, such as census and surveys collected by the government. Such data may provide information about demographics (e.g., age and sex) and health-related behaviors and risk factors (e.g., weight and smoking status). When health statistics is used, the data source should be clarified, and the quality and relevance of the data need to be appraised to ensure the rationality, timeliness and completeness of the data (IQWiG, 2009).

(5) Market research data is an important data source for BIA, including the distribution of competing products, early intervention options, and changes in treatment patterns.

(6) Expert opinion or survey is generally used as qualitative assessment. For example, confirmation of treatment pattern, list of competing products, etc. When other data sources are not available, expert interviews or surveys can be used to estimate certain inputs. However, there is a higher uncertainty in the input value from such data sources and thus sufficient sensitivity analysis should be conducted.

(7) Evacuation of data applicability can follow the following steps: (1) assessing the quality of the available data and its relevance to specific study objectives; (2) identifying whether discrepancies exist between the available data and the ideal data, and whether the difference, if any, can be adjusted by reasonable assumptions; and (3) appropriately processing the available data to improve the reliability of the model (IQWiG, 2009).

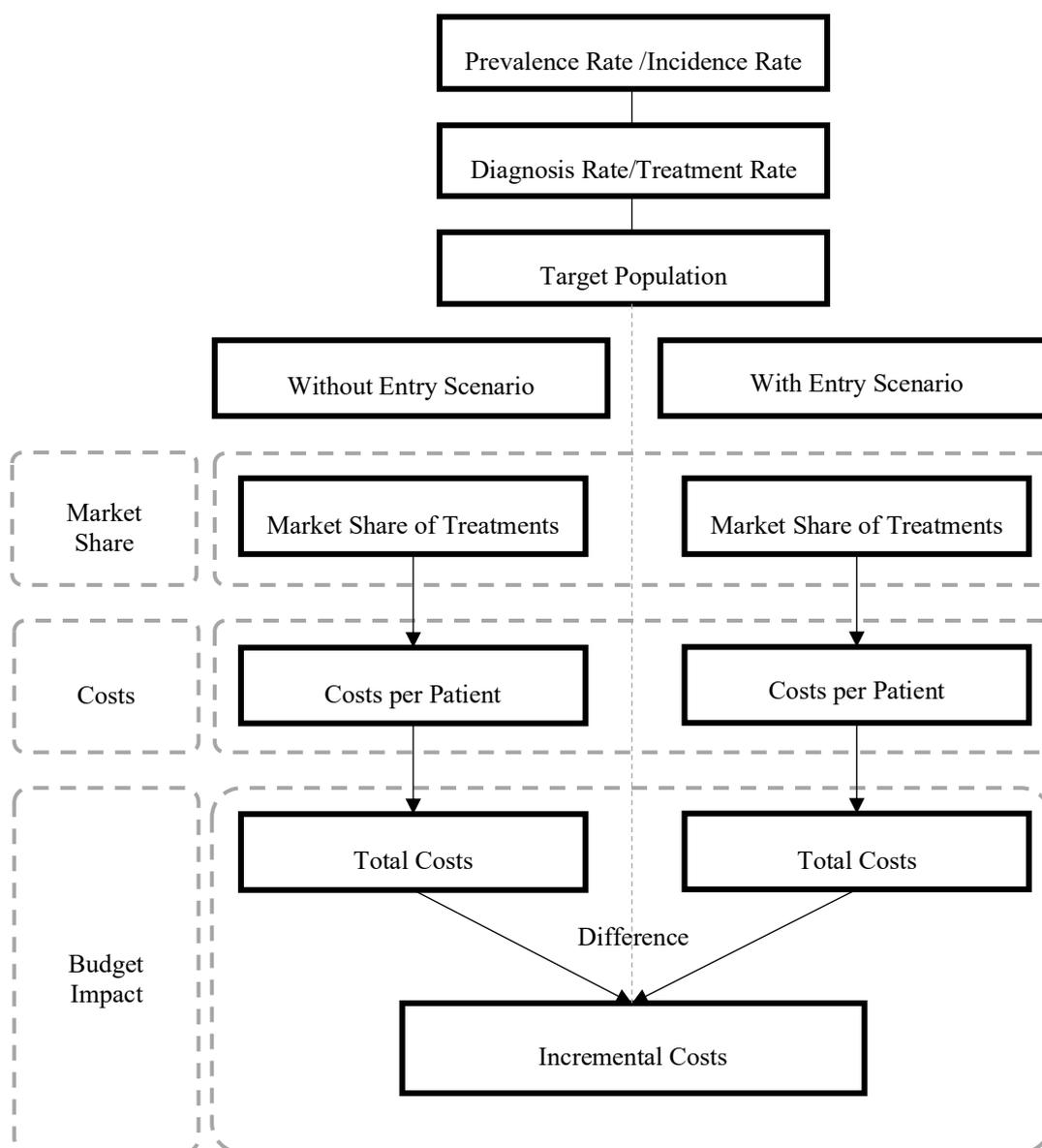


Figure 1. Flow chart of budget impact analysis

References

- Backhouse ME. Use of randomized controlled trials for producing cost-effectiveness evidence: Potential impact of design choices on sample size and study duration [J]. *Pharmacoeconomics* 2002; 20 (15): 1061-77.
- Bae EY, Lee EK. Pharmacoeconomic Guidelines and Their Implementation in the Positive List System in South Korea [J]. *Value Health* 2009; 12(3):36-41.
- Bala MV, Mauskopf JA, Wood LL. Willingness to pay as a measure of health benefits[J]. *Pharmacoeconomics* 1999; 15(1):9-18.
- Berger M L, Dreyer N, Anderson F, et al. Prospective observational studies to assess comparative effectiveness: the ISPOR good research practices task force report [J]. *Value Health* 2012; 15(2):217-30.
- Borenstein M, Hedges LV, Higgins JPT, et al. *Introduction to Meta-Analysis*[M]. 2009.
- Brazier J E, Yang Y, Tsuchiya A, et al. A review of studies mapping (or cross walking) non-preference based measures of health to generic preference-based measures[J]. *Eur J Health Econ* 2010; 11(2):215-25.
- Brazil Ministry of Health. *Methodological Guidelines: Economic Evaluation of Health Technologies* [M]. 2nd Edition. 2014.
- Bress AP, Bellows BK, King JB, et al. Cost-Effectiveness of Intensive versus Standard Blood-Pressure Control[J]. *N Engl J Med*. 2017; 377(8): 745–55
- Briggs AH, Sculpher M, Claxton K. *Decision Modelling for Health Economic Evaluation* [M]. Oxford: Oxford University Press. 2006.
- Briggs AH. Handling uncertainty in cost-effectiveness models [J]. *Pharmacoeconomics* 2000; 17 (5): 479-500.
- Canadian Agency for Drugs and Technologies in Health (CADTH). *Guidelines for the Economic Evaluation of Health Technologies: Canada* [M]. 4th Edition. 2017.
- Caro JJ, Briggs AH, Siebert U, et al. Modeling good research practices -overview: A report of the ISPOR-SMDM modeling good research practices task force-1[J]. *Value Health* 2012; 15: 796-803.
- Dolan P. Chapter 32 The measurement of health-related quality of life for use in resource allocation decisions in health care[M]. *Handbook of Health Economics*. Elsevier B.V. 2000:1723-1760.
- Drummond MF, Sculpher MJ, Claxton K, Stoddart GL, Torrance GW. *Methods for the economic evaluation of healthcare programmes* [M]. Fourth edition. Oxford: Oxford University Press. 2015.
- Drummond MF, Stoddart GL, Torrance GW. *Methods for the economic evaluation of healthcare programmes* [M]. Third edition. Oxford: Oxford University Press. 2005.
- Eddy DM, Hollingworth W, Caro JJ, et al. Model transparency and validation: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force-7[J]. *Value Health* 2012; 15: 843-50.
- Fenwick E, Claxton K, Sculpher M. Representing uncertainty: the role of cost-effectiveness acceptability curves [J]. *Health Econ* 2001; 10: 779-87.
- Foroutan N, Tarride JE, Xie F, et al. A Comparison of Pharmaceutical Budget Impact Analysis (BIA) Recommendations Amongst the Canadian Patented Medicine Prices Review Board (PMPRB), Public and Private Payers. *Pharmacoecon Open* 2019.
- Gerduso LA, Kotstanos JG. Review of health economic guidelines in the form of regulations, principles, policies, and positions [J]. *Drug Inf J* 1996; 30 (4): 1003-16. Gold MR, Siegel JE, Russell LB, Weinstein M. *Cost-effectiveness in health and medicine*[B]. New York: Oxford University Press, 1996.

Guideline for Preparing Cost-Effectiveness Evaluation to the Central Social Insurance Medical Council. Tokyo,2016.

Guidelines of Methodological Standards for Pharmacoeconomic Evaluations in Taiwan. Taiwan ,2006.

Hauck K, Smith PC, Goddard M. 2004. The economics of priority setting for health care: a literature review. HNP discussion paper series. Washington, DC: World Bank. Available: <http://documents.worldbank.org/curated/en/896691468780616417/The-economics-of-priority-setting-for-health-care-a-literature-review>.

Health Insurance Council (Ziekenfondsraad). Dutch guidelines for pharmacoeconomic research [R], 1999.

Hoyle C K, Tabberer M, Brooks J.Mapping the COPD Assessment Test onto EQ-5D[J].Value Health 2016; 19:469-477.

Institute for Quality and Efficiency in Health Care (IQWiG). General Methods for the Assessment of the Relation of Benefits to Costs [M].1st Edition. 2009.

Institute for Quality and Efficiency in Health Care (IQWiG). Working Paper on Modelling in health economic evaluations.2009.

ISPOR. 2019: <https://tools.ispor.org/peguidelines/>

Jr G L, Neumann P J, Erickson P, et al. Using real-world data for coverage and payment decisions: the ISPOR Real-World Data Task Force report.[J]. Value Health 2007; 10(5):326-35.

King M T, Viney R, Simon A P, et al. Australian Utility Weights for the EORTC QLU-C10D, a Multi-Attribute Utility Instrument Derived from the Cancer-Specific Quality of Life Questionnaire, EORTC QLQ-C30.[J]. Pharmacoeconomics 2018; 36(2):225-38.

Knies S., Severens JL, Ament HAJA, et al. The Transferability of Valuing Lost Productivity across Jurisdictions. Differences between National Pharmacoeconomic Guidelines [J]. Value Health 2010; 13 (5): 519-27.

Kopec J A, Willison K D. A comparative review of four preference-weighted measures of health-related quality of life[J]. J CLIN EPIDEMIOL 2003; 56(4): 317-25.

Kozma CM, Reeder CE, Schulz RM. Economic, clinical, and humanistic outcomes: a planning model for pharmacoeconomic research [J]. Clin Ther 1993; 15 (6): 1121-32.

Lakdawalla DN, Doshi JA, Garrison LP, Jr., et al. Defining Elements of Value in Health Care-A Health Economics Approach: An ISPOR Special Task Force Report [3]. Value Health 2018; 21(2): 131-9.

Lam CLK, Brazier J, McGhee SM. Valuation of the SF-6D health states is feasible, acceptable, reliable, and valid in a Chinese population [J]. Value Health 2008; 11 (2): 295-303.

Lancsar E, Louviere J. Conducting Discrete Choice Experiments to Inform Healthcare Decision Making[J]. Pharmacoeconomics 2008; 26(8):661-77.

Liljas B. How to calculate indirect costs in economic evaluations [J]. Pharmacoeconomics 1998; 13 (1): 1-7.

Liu G G, Wu H, Li M, et al. Chinese time trade-off values for EQ-5D health states [J]. Value Health 2014; 17(5):597-604.

Luce BR, Elixhauser A. Estimating costs in the economic evaluation of medical technologies [J]. Int J Technol Assess Health Care 1990; 6 (1): 57-75.

Luo N, Liu G, Li M, et al. Estimating an EQ-5D-5L Value Set for China [J]. Value Health 2017; 20(4):662-9.

MacPherson H. Pragmatic clinical trials [J]. Complement Ther Med 2004; 12: 36-40.

Mauskopf JA, Sullivan SD, AnnemansL, et al. Principles of good practice for budget impact analysis: report of the ISPOR Task Force on Good Research Practices—Budget Impact Analysis. Value

- Health 2007;10(5):336–47.
- Mills E J, Thorlund K, Ioannidis J P. Demystifying trial networks and network meta-analysis[J]. *BMJ* 2013; 346:f2914.
- Mühlbacher AC, Kaczynski A, Zweifel P, et al. Experimental measurement of preferences in health and healthcare using best-worst scaling: an overview[J]. *Health Econ Rev* 2016; 6(1):2.
- Mullins CD, Onwudiwe NC, Brancode GT, et al. Guidance Document: Global Pharmacoeconomic Model Adaption Strategies[J]. *Value Health Reg Issue* 2014; 5:7-13.
- Muratov S, Podbielski WD, Kennedy K et al. Preference-based Glaucoma-specific Health-related Quality of Life Instrument: Development of the Health Utility for Glaucoma[J]. *J Glaucoma* 2018; 27(7): 585-91.
- National Institute for Health and Clinical Excellence (NICE). Guide to the methods of technology appraisal 2013 [M]. 2013.
- Neumann PJ, Sanders GD, Russell LB, et al. Cost- Effectiveness in Health and Medicine (2nd edition) [M]. New York: Oxford University Press, 2017.
- Neumann PJ, Sanders GD, Russell LB, Siegel JE, Ganiats TG. Cost- Effectiveness in Health and Medicine[M]. 2nd Edition. New York: Oxford University Press.2016.
- Neumann PJ, Thorat T, Shi J, Saret CJ, Cohen JT. The changing face of the cost-utility literature, 1990-2012[J]. *Value Health* 2015; 18(2):271-7.
- Neumann P J , Willke R J , Garrison L P. A Health Economics Approach to US Value Assessment Frameworks—Introduction: An ISPOR Special Task Force Report [1][J]. *Value Health* 2018; 21(2):119-23.
- NICE. Assessing resource impact process manual: technology appraisals and highly specialised technologies. April 2017.
- Olsen JA, Smith RD. Theory versus practice: a review of 'willingness-to-pay' in health and health care[J]. *Health Econ* 2001; 10(1):39-52.
- Oxman AD, Guyatt GH. A consumer's guide to subgroup analyses [J]. *Ann Intern Med* 1992; 116 (1): 78-84.
- Round J. Is a QALY still a QALY at the end of life?[J]. *J Health Econ*, 2012, 31(3): 521- 527.
- WHO. Choosing Interventions that are Cost Effective (WHO-CHOICE), Threshold values for intervention cost-effectiveness by Region. http://www.who.int/entity/choice/costs/CER_thresholds_regions.xls.
- Round J, Paulden M. Incorporating equity in economic evaluations: a multi-attribute equity state approach. *Eur J Health Econ*, 2018; 19(4): 489-98.
- Rowen D, Brazier J, Roberts J. Mapping SF-36 onto the EQ-5D index: how reliable is the relationship?[J]. *Health Qual Life Outcomes* 2009; 7(1):27.
- Sculpher M J, Pang F S, Manaca A. 2004. Generalisability in economic evaluation studies in healthcare: a review and case studies [J]. *Health Technol Assess*2004; 8(49): 1-192.
- Sculpher MJ, Fenwick E, Claxton K. Assessing quality in decision analytic cost-effectiveness models: a suggested framework and example of application [J]. *Pharmacoeconomics* 2000; 17 (5): 461-7.
- Shiroiwa T , Fukuda T , Ikeda S , et al. Development of an Official Guideline for the Economic Evaluation of Drugs/Medical Devices in Japan[J]. *Value Health* 2017; 20(3):372-78.
- Smith DH, Gravelle H. The practice of discounting in economic evaluations of healthcare interventions [J]. *Int J Technol Assess Health Care* 2001; 17 (2): 236-43.
- Sullivan SD, Mauskopf JA, Augustovski F, et al. Budget Impact Analysis—Principles of Good Practice: Report of the ISPOR 2012 Budget Impact Analysis Good Practice II Task Force. *Value Health* 2014;17(1):5–14.

Sun X, Briel M, Walter SD, et al. Is a subgroup effect believable? Updating criteria to evaluate the credibility of subgroup analyses. *BMJ* 2010; 340: c117.

Sun X, Ioannidis JP, Agoritsas T, et al. How to use a subgroup analysis: users' guide to the medical literature. *JAMA* 2014; 311(4): 405-11.

Takashi Fukuda. Guideline for preparing cost-effectiveness evaluation to the central social insurance medical council[M]. 2016.

The Collège des Économistes de la Santé (the French Health Economists Association). French Guidelines for the Economic Evaluation of Health Care Technologies[M]. 2004.

The Pharmaceutical Benefits Advisory Committee (PBAC). Guidelines for preparing a submission to the Pharmaceutical Benefits Advisory Committee[M].5th Edition. 2016.

Thomas Butt, Gordon G.Liu, David D, et al. Taking stock of cost effectiveness analysis of healthcare in China[J]. *BMJ Global Health* 2019; (4):1136-418.

Torrance GW, Thomas WH, Sackett DL. A utility maximization model for evaluation of health care programs. *Health Serv Res.* Summer 1972;7(2):118-33.

Wakker P, Klaassen M. Confidence intervals for cost-effectiveness ratio [J]. *Health Econ* 1995; 4(5): 373-82.

WHO. World Health Statistics 2018.

https://www.who.int/gho/publications/world_health_statistics/2018/en/

Wild D, Grove A, Martin M, et al. Principles of Good Practice for the Translation and Cultural Adaptation Process for Patient-Reported Outcomes (PRO) Measures: report of the ISPOR Task Force for Translation and Cultural Adaptation [J]. *Value Health* 2005; 8(2):94-104.

Willke R J. 2003. Tailor-made or off-the-rack? The problem of transferability of health economic data[J]. *Expert Rev Pharmacoecon Outcomes Res* 2003; 3(1):1-4.

Woodley HJ, Bourdage JS, Ogunfowora B, et al. Examining Equity Sensitivity: An Investigation Using the Big Five and HEXACO Models of Personality. *Front Psychol* 2015; 6: 2000.

Appendix 1: Standard Reporting Format

Cover Page

- ◇ Title
- ◇ Author and Affiliation
- ◇ Author Contributions: List the contributions of each author in this report
- ◇ Funding/Support: List all sources of funding and financial support received in this report
- ◇ Acknowledgements

Abstract

- ◇ Background: State the importance and significance of the study questions
- ◇ Objectives: Specify the study questions to be addressed
- ◇ Methods: Describe the study design, sample population, and the primary end-points used for analyses
- ◇ Results: Report primary results and data generated from analyses
- ◇ Limitation: Discuss major limitations of the analytical methods and data used in the study
- ◇ Conclusions: Interpret the study results with basic findings

Text part

1 Introduction

1.1 Background

1.1.1 Condition

- ◇ Describe the severity, epidemiological characteristics, and treatment pattern of the condition
- ◇ Describe the societal burden of the condition, including economic burden and quality of life burden

1.1.2 Interventions

- ◇ Describe major interventions, treatment routes, dosing schedules, concomitant medications, requirements of the treatment setting, etc. for the condition
- ◇ Describe treatment efficacy, adverse events, cautions or warnings, costs, etc.
- ◇ Describe the features, target population, and current status of availability in China market and healthcare insurance coverage of the new intervention

1.2 Literature Review

- ◇ Systematically review local and global literature on the perception and relevant evidence of the new intervention
- ◇ State the differences, features and innovations of the study to be conducted compared with existing literature

1.3 Objectives

- ◇ Specify the primary and secondary objectives of the study
- ◇ Identify the target audience and potential audiences

2 Methods

2.1 Target Population

- ◇ Composition of the study sample: Describe the target population of the intervention/application, and the selection criteria of the control group with justifications
- ◇ State whether subgroup analyses will be conducted in the study and specify the criteria of subgroup classifications with justifications

2.2 Perspective

- ◇ Describe the selected perspective and reasons for the selection

2.3. Interventions and Comparators

- ◇ Describe the selected comparators and reasons for the selection
- ◇ Describe the interventions and comparators, including dosing schedules, treatment durations, concomitant examinations and treatments, management of adverse events and subsequent treatments after disease progression

2.4 Time Horizon

- ◇ Describe the selected time horizon and reasons for the selection

2.5 Methods

- ◇ Study design: The study design includes the modeling study, study based on individual patient data (including retrospective and prospective analyses), and others
 - For the modeling study, briefly describe the rationale for the selection of the modeling technique (e.g., a decision tree model, a Markov model, a partitioned survival model, etc.), the model structure, definition of health states, and analytical method.
 - For the study based on individual patient data, briefly describe the primary study design and definition of key parameters

2.6 Key Assumptions - *for modeling studies only*

- ◇ Describe key assumptions as well as the important features and justifications for the assumptions

2.7 Model Outputs - *for modeling studies only*

- ◇ Describe key model outputs

2.8 Model Inputs and Data Sources - *for modeling studies only*

2.8.1 Baseline Characteristics of Target Population

- Describe baseline characteristics of the target population in the model
- List sources of the input data with justifications

2.8.2 Efficacy Inputs

- Describe efficacy inputs and analytical methods
- List sources or assumptions of the input data with justifications

2.8.3 Safety

- Describe safety inputs and analytical methods
- List sources or assumptions of the input data with justifications

2.8.4 Resource Use and Cost Identification

- Identify cost components (e.g., direct cost, indirect cost, etc.) for analyses and provide reasons

- Describe, measure, and evaluate the resource use included in the analyses
- For each cost input, describe where it is obtained and how it is measured

2.8.5 Health Outcomes Identification

- Describe the selected key health outcome measures (e.g., utility), the measurement approach, and reasons for the selection

2.8.6 Discount Rate

- Describe whether discounting is applied and reasons
- Describe the selected discount rate and reasons for the selection

2.9 The analytical method of primary results - *for modeling studies only*

- ◇ Describe the specific method (including cost-effectiveness analysis, cost-utility analysis, cost-benefit analysis and cost-minimization analysis, etc.) with justifications
- ◇ Describe methods for the incremental analysis

2.10 Variability and Uncertainty - *for modeling studies only*

- ◇ Variability: Describe the stratification of the target population, indicate the possible subgroups (based on effectiveness, preference and cost), and state how variability of other forms is analyzed
- ◇ Uncertainty: Indicate the source of uncertainties, and methods to analyze uncertainties (i.e., deterministic sensitivity analysis, probabilistic sensitivity analysis)

2.11 Selection of Data and Sample - *for observational studies only*

- ◇ Describe the data sources
- ◇ Describe the sample selection process

2.12 Selection of Study Variables - *for observational studies only*

- ◇ Describe the selection and definitions of variables for the cost, effectiveness and incremental cost effectiveness ratio (ICER)

2.13 Statistical Analysis - *for observational studies only*

- ◇ Describe the methods for statistical analyses

2.14 Variability and Uncertainty - *for observational studies only*

- ◇ Variability: Describe the stratification of the target population, indicate the possible subgroups (based on effectiveness, preference and cost) and state how variability of other forms is analyzed
- ◇ Uncertainty: Indicate the source of uncertainties, and methods to analyze uncertainties

3 Results

3.1 Base-case Results

- ◇ Report the statistical distribution and characteristics of key variables across groups in the overall sample population (including major demographic characteristics, relevant clinical measures, cost measures, and health outcome measures, etc.)
- ◇ Report primary study results obtained from model estimation and systematic analyses, with the focus on the costs and the health outcomes of different interventions, and the ICER
- ◇ Report results first in natural units first, and then in selected units (e.g., quality-adjusted life years [QALYs] or monetary benefits) converted from the natural units
- ◇ If data permits, consider providing statistical analysis results of key subgroups
- ◇ Present results in tables and figures (graphical results are encouraged)

3.2 Sensitivity Analysis Results

- ◇ Report variation and distribution of the primary results after varying key assumptions or input values in the model
- ◇ Describe the sensitivity of primary results, and the major influential factors, sources or situations
- ◇ Report subgroup analysis results, demonstrate the impact of statistical distribution, and report results of other variability analyses

4 Discussions

4.1 Summary of Results

- ◇ Summarize primary modeling or statistical analysis results
- ◇ Compare the similarities and differences of results between this study and other relevant studies, and discuss the possible reasons
- ◇ Discuss uncertainties and key influential factors
- ◇ Discuss trade-offs among costs, benefits and harms

4.2 Generalizability

- ◇ Discuss the applicability and variability of the study results in terms of disease epidemiology, population characteristics, regional characteristics, clinical practice, resource use patterns, etc.

4.3 Limitations

- ◇ Indicate the major problems and limitations that may exist in the study methods or input data
- ◇ Discuss major problems that should be noted in further research and possible methods to address them
- ◇ Discuss biases that may be caused by the study methods and input data used

4.4 Impacts on Healthcare Service

- ◇ Comment on possible impacts on the medical service resources
- ◇ Discuss impacts on the budget caused by epidemiological factors and other intervening factors
- ◇ Provide reasonable suggestions on relevant healthcare policies based on the pros and cons of interventions and comparators

4.5 Future Study Directions

- ◇ Identify existing evidence gaps, and indicate directions and areas for further research

4.6 Equity

- ◇ If the situation allows, equity of the evaluation results should be evaluated by conducting sensitivity analyses or subgroup analyses

5 Conclusions

- ◇ State the study objectives and questions to be addressed

- ◇ Summarize the primary study results and relevant implications, the population and settings appropriate for the intervention, and uncertainties and considerations in the study

6 References

7 Appendices

- ◇ Present detailed tables and figures of key inputs, questionnaire design, data sources, the decision tree model or other explanatory materials necessary, etc. for readers' review and verification
- ◇ If a systematic review/meta-analysis is conducted, databases for the literature search, search strategy, literature screening process, basic information for the included studies, the standard and format of the extracted data, literature quality evaluation tools or standards, etc. should be presented

Appendix 2: Pharmacoeconomic Evaluation Quality Checklist

When evaluating the quality of a pharmacoeconomic research report, the following contents should be checked against the standard reporting format :

Standard reporting format	Checklist contents
Cover Page	
Title	<ul style="list-style-type: none"> • Whether it is a pharmacoeconomic research; the research may describe itself as an economic evaluation or cost-effectiveness analysis • Whether the intervention is listed • Whether the disease or health condition is described
Funding/Support	<ul style="list-style-type: none"> • Whether funding/support is listed • Whether the research is influenced by funding/support
Abstract	<ul style="list-style-type: none"> • Whether research background, objectives, methods, results, limitations, and conclusions are outlined
1 Introduction	
1.1 Background	<ul style="list-style-type: none"> • Whether the background information of the condition is described • Whether the basic information of the intervention is described
1.2 Literature review	<ul style="list-style-type: none"> • Whether a systematic literature review is conducted

	<ul style="list-style-type: none"> • Whether limitations of the existing literature are identified
1.3 Objectives	<ul style="list-style-type: none"> • Whether the research question is specified • Whether the study audience is specified
2 Method	
2.1 Target population	<ul style="list-style-type: none"> • Whether the characteristics of the target population are described
2.2 Research perspective	<ul style="list-style-type: none"> • Whether the research perspective is specified
2.3 Selection of comparators	<ul style="list-style-type: none"> • Whether reasons for the selection of the comparators are provided
2.4 Time horizon	<ul style="list-style-type: none"> • Whether the research time horizon is described
2.5 Methods	<ul style="list-style-type: none"> • Whether the research clearly describe itself as a modeling study or a study based on individual patient data
2.6 Key assumptions (<i>Modelling study</i>)	<ul style="list-style-type: none"> • Whether the research assumptions are described
2.7 Sample data (<i>Observational study</i>)	<ul style="list-style-type: none"> • Whether the sample selection process is specified • Whether the sample characteristics are described
2.8 Cost identification	<ul style="list-style-type: none"> • Whether the cost component is specified
2.9 Cost measurement	<ul style="list-style-type: none"> • Whether the sources of cost data are described (<i>Modelling study</i>) • Whether the methods of cost measurement are described (<i>Observational study</i>)
2.10 Outcomes identification	<ul style="list-style-type: none"> • Whether the type of outcomes is specified: effectiveness, utility, or benefit

2.11 Outcomes measurement	<ul style="list-style-type: none"> • Whether the sources of outcomes data are described (<i>Modelling study</i>) • Whether the methods of outcomes measurement are described (<i>Observational study</i>)
2.12 Discounting	<ul style="list-style-type: none"> • Whether the discount rate of cost is specified (<i>if needed</i>) • Whether the discount rate of outcomes is specified (<i>if needed</i>)
2.13 Model selection (<i>Modelling study</i>)	<ul style="list-style-type: none"> • Whether reasons for the model selection are provided • Whether the model structure diagram is presented
2.14 Statistical analysis (<i>Observational study</i>)	<ul style="list-style-type: none"> • Whether the statistical analysis method is described
2.15 Variability/Uncertainty	<ul style="list-style-type: none"> • Whether different subgroups are analysed • Whether uncertainties are analysed • Whether the probabilistic sensitivity analysis is included
3 Results	
3.1 Base-case results	<ul style="list-style-type: none"> • Whether the analysis results of key variables in the intervention and comparator groups are reported • Whether the analysis results of costs and outcomes in the intervention and comparator groups are reported
3.2 Incremental analysis	<ul style="list-style-type: none"> • Whether the incremental analysis is conducted

	<ul style="list-style-type: none"> • Whether the incremental cost-effectiveness ratio is reported
3.3 Sensitivity analysis results	<ul style="list-style-type: none"> • Whether the sensitivity analysis of primary results is reported
3.4 Equity	<ul style="list-style-type: none"> • Whether the equity impact of the study results is reported
4 Discussions	
4.1 Generalizability	<ul style="list-style-type: none"> • Whether the generalizability of study results is discussed
4.2 Limitations	<ul style="list-style-type: none"> • Whether the study limitations are discussed
5 Conclusions	<ul style="list-style-type: none"> • Whether the main findings are summarized
6 References	<ul style="list-style-type: none"> • Whether the references are listed
7 Appendices	<ul style="list-style-type: none"> • Whether the key data tables/figures are provided