
Multi-Criteria Decision Analysis in Healthcare

Its usefulness and limitations for
decision making

weber®

Multi-Criteria Decision Analysis in Healthcare

Its usefulness and limitations for
decision making

Editors:

Néboa Zozaya González

Juan Oliva Moreno

Álvaro Hidalgo Vega

Sponsored by:



weber®

Edited by:

© Fundación Weber

C/ Las Norias, 123

28221 Majadahonda - Madrid

Editorial coordination: iris.sanpedro@weber.org.es

Translation copyright © by Ana Moreno & Robin White

ISBN: 978-84-947703-8-8

D.L.: M-36667-2018

Printed in Madrid, December 2018

Free copy. Not available for sale.

AUTHORS

Chapter Néboa Zozaya González. Health Economics Area Manager. Weber.

1

Juan Oliva Moreno. Associate Professor in the Economics and Finance Department and senior research fellow of the Health and Economics Group. University of Castilla-La Mancha.

Álvaro Hidalgo Vega. Associate Professor in the Economics and Finance Department. University of Castilla-La Mancha. President. Weber Foundation.

INTRODUCTION

Chapter Antonio García Ruíz. Associate Professor. University of Malaga. Senior research fellow of Health Economics and Rational Use of Medication Unit.

2

Nuria García-Agua Soler. University of Malaga.

THE VALUE OF INNOVATION IN HEALTHCARE

Chapter Jaime Puig-Junoy. Associate Professor. Department of Economics and Business. Main researcher of the Centre for Research in Health and Economics (CRES). Universitat Pompeu Fabra.

3

MULTI-CRITERIA DECISION ANALYSIS: WHAT IS IT AND WHAT IS IT FOR?

Chapter Jaime Espín Balbino. Associate Professor. Andalusian School of Public Health.

4

SELECTION OF CRITERIA IN MULTI-CRITERIA DECISION ANALYSIS

Chapter Carlos Martín Saborido. Head of Postgraduate studies and research. San Rafael-Nebrija Health Sciences University Center.

5

TECHNIQUES OF WEIGHTING, SCORING, MODELLING AND MANAGEMENT OF UNCERTAINTY

Chapter Marta Trapero-Bertran. Associate Professor. Institut de Recerca en Avaluació i Politiques Públiques (IRAPP). Universitat Internacional de Catalunya.

6

REVIEW OF METHODOLOGICAL AND GOOD PRACTICE GUIDELINES IN MCDA

Chapter Javier Mar Medina. Head of the Health Management Unit. Alto Deba Hospital.

7

APPLICATION OF MULTI-CRITERIA DECISION ANALYSIS TO PREVALENT DISEASES

Chapter Carlos Campillo-Artero. Servei de Salut de les Illes Balears. Associate Researcher of the Centre for Research in Economics and Health (CRES). Pompeu Fabra University.

8

APPLICATION OF MULTI-CRITERIA DECISION ANALYSIS TO ORPHAN DRUGS

Chapter Xavier Badia Llach. Partner & CEO. Omakase Consulting.

9

Antoni Gilabert-Perramon. Director of Pharmacy and Medicines. Catalan Health and Social Care Consortium.

MCDA IN DECISION-MAKING: EXPERIENCES IN THE SPANISH HEALTHCARE FRAMEWORK

Chapter Bleric Alcalá Revilla. Health Economics Area Senior Consultant. Weber.

10

Néboa Zozaya González. Health Economics Area Manager. Weber.

A PRACTICAL CASE OF MCDA APPLIED TO RARE AND PREVALENT DISEASES

Chapter All authors

11

SUMMARY AND FINAL CONCLUSIONS

INDEX

| | |
|--|-----------|
| FOREWORD..... | 13 |
| 1. INTRODUCTION..... | 15 |
| References..... | 19 |
| 2. THE VALUE OF INNOVATION IN HEALTHCARE..... | 21 |
| 2.1. Introduction..... | 22 |
| 2.2. Definition of innovation..... | 23 |
| 2.3. The social value of the drug..... | 26 |
| 2.3.1. The value of medical innovation: saving lives, saving money..... | 28 |
| 2.4. Innovation and prevalence of disease | 30 |
| 2.4.1. Priority drugs | 33 |
| 2.5. Tools for measuring innovation | 37 |
| 2.6. Policies to promote innovation | 40 |
| 2.6.1 From the regulation of prices to the regulation of healthcare outcomes. Risk-Sharing Agreements (RSAs) as a case study..... | 41 |
| 2.7. Conclusions | 45 |
| References | 46 |
| 3. MULTI-CRITERIA DECISION ANALYSIS: WHAT IS IT AND WHAT IS IT FOR? | 49 |
| 3.1. Introduction..... | 50 |
| 3.2. Multiple criteria in healthcare decision-making..... | 50 |
| 3.3. What is MCDA?..... | 54 |
| 3.4. What can MCDA be used for? A simple illustration | 59 |
| 3.5. Good practices in MCDA: advantages and risks..... | 63 |
| 3.6. The ‘multi-criteria’ rationale versus the rationale of opportunity cost are they alternatives or complementary? | 65 |
| References | 68 |
| 4. SELECTION OF CRITERIA IN MULTI-CRITERIA DECISION ANALYSIS | 71 |
| 4.1. Introduction..... | 72 |
| 4.2. Selection of criteria in MCDA..... | 72 |
| 4.3. Experiences in the selection of criteria..... | 76 |
| 4.4. Conclusions | 79 |
| References | 81 |
| 5. TECHNIQUES OF WEIGHTING, SCORING, MODELLING AND MANAGEMENT OF UNCERTAINTY..... | 83 |
| 5.1. Introduction..... | 84 |
| 5.2. What does the weighting of criteria reflect?..... | 84 |
| 5.3. What is behind the scoring of the criteria? | 85 |
| 5.4. How are the criteria scored and weighted? | 86 |
| 5.4.1. Techniques of direct scoring | 86 |
| 5.4.2. Pair-wise comparison with ordinal scales (hierarchical technique) | 88 |
| 5.4.3. Discrete choice method | 90 |
| 5.4.4. Keeney-Raiffa method | 91 |

| | |
|--|------------|
| 5.5. How do we combine the weighting and the scoring?..... | 92 |
| 5.6. Discussion about scoring and weighting..... | 93 |
| 5.7. Assessment and management of uncertainty | 94 |
| 5.7.1. Structural uncertainty | 95 |
| 5.7.2. Uncertainty in the scoring..... | 96 |
| 5.7.3. Uncertainty in the weighting..... | 96 |
| 5.7.4. Sensitivity analysis in MCDA | 97 |
| 5.8. Conclusions | 97 |
| References | 99 |
| 6. REVIEW OF METHODOLOGICAL AND GOOD PRACTICE GUIDELINES IN MCDA | 101 |
| 6.1. Introduction..... | 102 |
| 6.2. Methods for the review of methodological and good practice guidelines in MCDA..... | 102 |
| 6.3. Results | 103 |
| 6.3.1. An overview of the steps necessary to conduct an MCDA | 104 |
| 6.3.2. Criteria most commonly used in MCDA | 105 |
| 6.3.3. Types of MCDA models..... | 107 |
| 6.3.4. Examples of healthcare decisions to which MCDA could be applied | 109 |
| 6.3.5. Checklist of good practice guidelines in MCDA..... | 111 |
| 6.4. Discussion and next steps in terms of methodology | 112 |
| 6.5. Conclusions | 114 |
| References | 114 |
| 7. APPLICATION OF MULTI-CRITERIA DECISION ANALYSIS TO PREVALENT DISEASES | 119 |
| 7.1. Introduction to MCDA in prevalent diseases | 120 |
| 7.2. Methods..... | 120 |
| 7.3. Results of the literature review about the use of MCDA in prevalent diseases | 121 |
| 7.3.1. MCDA in oncology | 121 |
| 7.3.2. MCDA in COPD | 123 |
| 7.3.3. MCDA in myocardial infarction | 125 |
| 7.3.4. MCDA in stroke | 125 |
| 7.3.5. MCDA in hypertension..... | 127 |
| 7.3.6. MCDA in diabetes | 127 |
| 7.3.7. Review of books..... | 127 |
| 7.4. Discussion..... | 129 |
| 7.5. Conclusions | 130 |
| References | 131 |
| 8. APPLICATION OF MULTI-CRITERIA DECISION ANALYSIS TO ORPHAN DRUGS | 135 |
| 8.1. Introduction..... | 136 |
| 8.2. Reasons for using MCDA in decisions about price and reimbursement of orphan drugs..... | 137 |
| 8.3. Models of combination of criteria applied to orphan drugs | 138 |
| 8.3.1. Theoretical MCDA models | 139 |
| 8.3.2. Application of MCDA to simulated orphan drugs..... | 142 |
| 8.3.3. Application of MCDA to real orphan drugs..... | 142 |
| 8.3.4. Other models of combination of multiple criteria | 144 |
| 8.4. The appraisal of health technologies and the MCDA with orphan drugs: practical cases | 146 |

| | |
|--|------------|
| 8.5. Recommendations and international consensus about assessment and appraisal of orphan drugs and the incorporation of MCDA..... | 147 |
| 8.6. Critical appraisal and lines of progress | 149 |
| 8.6.1. General comments..... | 149 |
| 8.6.2. Methodological elements | 149 |
| 8.6.3. HTA and MCDA | 152 |
| 8.6.4. Criticisms concerning the ineptitude of economic evaluation for the appraisal of orphan drugs..... | 153 |
| 8.6.5. Recommendations and international consensus | 155 |
| 8.7. Conclusions | 155 |
| References..... | 157 |
| 9. MCDA IN DECISION-MAKING: EXPERIENCES IN THE SPANISH HEALTHCARE FRAMEWORK | 161 |
| 9.1. The need to make better decisions in the selection, evaluation and positioning of drugs | 162 |
| 9.2. The multicriteria decision analysis and the need to make decisions based on a reflective framework..... | 163 |
| 9.2.1. The importance of the criteria used in the MCDA appraisal according to the decision-makers | 164 |
| 9.3. Evidence-based information: the reflective scoring matrix to inform the discussion of the decision | 168 |
| 9.4. The relation of the MCDA evaluation to decision-making..... | 168 |
| 9.5. Conclusions | 169 |
| References..... | 170 |
| 10. A PRACTICAL CASE OF MCDA APPLIED TO RARE AND PREVALENT DISEASES | 171 |
| 10.1. Introduction and objectives..... | 172 |
| 10.2. Methodology..... | 172 |
| 10.2.1. Selected diseases..... | 172 |
| 10.2.2. Steps followed to carry out the implementation of the MCDA | 174 |
| 10.3. Results..... | 180 |
| 10.3.1. Weighting of the criteria | 180 |
| 10.3.2. Scoring of the interventions..... | 183 |
| 10.3.3. Estimated total value of the interventions | 184 |
| 10.3.4. Managing uncertainty | 186 |
| 10.4. Discussion..... | 190 |
| 10.5. Conclusions | 192 |
| References..... | 193 |
| 11. SUMMARY AND FINAL CONCLUSIONS | 197 |
| 11.1. MCDA in rare versus prevalent diseases | 198 |
| 11.2. Strengths and limitations of MCDA..... | 199 |
| 11.3. Conclusions and final recommendations | 203 |
| References..... | 205 |

TABLE INDEX

| | | |
|-------------------|--|------------|
| TABLE 2.1. | DETERMINANTS AND THEIR CONTRIBUTION IN HEALTH | 22 |
| TABLE 2.2. | DEGREE OF INNOVATION | 24 |
| TABLE 2.3. | BURDEN OF THE MAIN DISEASES IN THE WORLD, EXPRESSED IN DISABILITY-ADJUSTED LIFE YEARS, 2016 | 32 |
| TABLE 2.4. | NEW DRUGS APPROVED BY THE EMA, 2016 | 35 |
| TABLE 2.5. | NET SOCIAL VALUE OF INNOVATION VERSUS ALTERNATIVE | 38 |
| TABLE 2.6. | ADVANTAGES, BARRIERS AND DIFFICULTIES IN RISK-SHARING AGREEMENTS | 43 |
| TABLE 3.1. | MAIN CRITERIA AND 'OTHER' CONSIDERATIONS USED INTERNATIONALLY TO PRIORITISE NEW HEALTH TECHNOLOGIES | 52 |
| TABLE 3.2. | COMPARISON OF MCDA METHODS | 58 |
| TABLE 3.3. | CHARACTERISTICS OF THE DRUGS IN THE APPRAISAL PROCESS | 60 |
| TABLE 3.4. | PERFORMANCE SCORES OF DRUGS | 60 |
| TABLE 3.5. | OUTRANKING RELATIONS AND WEIGHTS | 62 |
| TABLE 3.6. | ATTRIBUTES OF DRUGS, THE GOALS AND WEIGHTS AGAINST DIFFERENT CRITERIA | 62 |
| TABLE 3.7. | PRINCIPLES OF GOOD PRACTICE IN MCDA IN THE HEALTHCARE SECTOR | 64 |
| TABLE 3.8. | KEY ISSUES FOR THE USE OF MCDA IN HTA | 67 |
| TABLE 4.1. | CRITERIA USED TO ASSESS DRUGS FOR PRICE AND REIMBURSEMENT IN THE EUROPEAN CONTEXT | 72 |
| TABLE 4.2. | STEPS IN THE PROCESS OF MEASURING VALUE IN AN MCDA | 74 |
| TABLE 4.3. | APPROACHES FOR SELECTING OBJECTIVES AND CRITERIA | 76 |
| TABLE 4.4. | CRITERIA INCLUDED IN THE EVIDEM FRAMEWORK | 77 |
| TABLE 4.5. | KEY QUESTIONS AND POSSIBLE CRITERIA FOR A DRUG REIMBURSEMENT APPRAISAL PROCESS IN BELGIUM, IN ACCORDANCE WITH THE MCDA FRAMEWORK | 78 |
| TABLE 4.6. | SELECTED CRITERIA AND THEIR WEIGHTS | 78 |
| TABLE 5.1. | EXAMPLE OF MATRIX PROBLEM | 85 |
| TABLE 5.2. | EXAMPLE OF DIRECT ESTIMATES: EVIDEM FRAMEWORK | 87 |
| TABLE 5.3. | EXAMPLE OF THE DISCRETE CHOICE METHOD | 90 |
| TABLE 5.4. | SCORING AND AGGREGATION METHODS | 93 |
| TABLE 6.1. | EXAMPLES OF HEALTHCARE DECISIONS TO WHICH MCDA MIGHT BE APPLIED | 110 |
| TABLE 6.2. | ISPOR CHECKLIST OF GOOD PRACTICES FOR MCDA | 113 |
| TABLE 7.1. | RESULT OF SEARCHES FOR ARTICLES ABOUT PREVALENT DISEASES | 121 |
| TABLE 7.2. | CRITERIA APPLIED IN THE APPRAISAL OF TREATMENTS FOR COPD USING MCDA | 124 |

| | | |
|---------------------|---|------------|
| TABLE 7.3. | LIST OF TECHNOLOGIES ASSESSED IN LOMBARDY AND RESULT OF THE EVALUATION | 129 |
| TABLE 8.1. | ARGUMENTS ABOUT THE LIMITATIONS OF HTA AND ECONOMIC EVALUATION IN MAKING DECISIONS ABOUT THE VALUE OF ORPHAN DRUGS AND ABOUT THEIR COVERAGE AND REIMBURSEMENT | 138 |
| TABLE 8.2. | MAIN CRITERIA INCLUDED IN THE MCDA OF SOME OF THE STUDIES REVIEWED | 140 |
| TABLE 8.3. | RECOMMENDATIONS FOR INCREASING CONGRUENCE BETWEEN THE CRITERIA IN ASSESSMENT AND THOSE IN APPRAISAL OF ORPHAN DRUGS | 148 |
| TABLE 8.4. | REPLIES TO THE ARGUMENTS ABOUT THE LIMITATIONS OF THE HTA AND THE ECONOMIC APPRAISAL IN MAKING DECISIONS ABOUT THE VALUE OF ORPHAN DRUGS AND ABOUT THEIR COVERAGE AND REIMBURSEMENT | 153 |
| TABLE 9.1. | DATA RESULTING FROM THE WEIGHTING OF THE QUANTITATIVE CRITERIA OF THE EVIDEM FRAMEWORK BY THE EVALUATORS OF THE SPANISH AGENCY FOR MEDICINES AND HEALTH PRODUCTS | 165 |
| TABLE 9.2. | DATA RESULTING FROM THE WEIGHTING OF THE QUANTITATIVE CRITERIA OF THE EVIDEM FRAMEWORK AT THE REGIONAL LEVEL IN SPAIN | 166 |
| TABLE 10.1. | TYPES OF MPS | 174 |
| TABLE 10.2. | CURRENT SITUATION OF COPD, MPS I AND MPS VII IN THE 13 CRITERIA TO BE EVALUATED | 175 |
| TABLE 10.3. | EVIDENCE AND ASSUMPTIONS FOR THE APPRAISAL OF THE FICTITIOUS DRUG IN SEVERE COPD | 177 |
| TABLE 10.4. | EVIDENCE AND ASSUMPTIONS FOR THE APPRAISAL OF THE FICTITIOUS DRUG IN MPS-I (SCHEIE SYNDROME) | 178 |
| TABLE 10.5. | EVIDENCE AND ASSUMPTIONS FOR THE APPRAISAL OF THE FICTITIOUS DRUG IN MPS-VII (SLY SYNDROME) | 179 |
| TABLE 10.6. | BASE CASES CONSIDERED FOR THE FICTITIOUS DRUGS | 183 |
| TABLE 10.7. | ESTIMATED TOTAL VALUE OF THE INTERVENTIONS IN THE FOUR SCENARIOS OF THE SENSITIVITY ANALYSIS, USING THE TWO WEIGHTING METHODS | 187 |
| TABLE 10.8. | PERCENTAGE VARIATIONS IN THE TOTAL ESTIMATED VALUE COMPARED WITH THE BASE CASE, FOR THE FOUR SCENARIOS OF THE SENSITIVITY ANALYSIS, UNDER THE TWO WEIGHTING METHODS | 188 |
| TABLE 10.9. | COMPARISON OF THE TOTAL ESTIMATED VALUE IN THE TEST AND RE-TEST, AND INTRAClass CORRELATIONS | 189 |
| TABLE 10.10. | TOTAL ESTIMATED VALUE OBTAINED IN THE RE-TEST FOR SENSITIVITY ANALYSES | 190 |

FIGURE INDEX

| | |
|--|------------|
| FIGURE 2.1. HEALTH EXPECTANCY AT BIRTH | 22 |
| FIGURE 2.2. YEARS OF HEALTHY LIFE AT 65, 2014 | 23 |
| FIGURE 2.3. NEW DRUGS AND LIFE EXPECTANCY AT BIRTH | 27 |
| FIGURE 2.4. INNOVATIONS AND BURDEN OF DISEASE (DALYs), 1995-2009 | 33 |
| FIGURE 2.5. REQUESTS FOR TECHNICAL ADVICE FROM THE EMA, 2016 | 36 |
| FIGURE 2.6. DISEASE BURDEN IN DALYS PER 100,000 IN HABITANTS IN EUROPE, 2016 | 37 |
| FIGURE 2.7. DISEASE BURDEN MEASURED IN DALYs, SPAIN 2000-2016 | 39 |
| FIGURE 2.8. AGREEMENTS BASED ON RESULTS, BETWEEN FUNDERS AND INDUSTRY | 43 |
| FIGURE 2.9. DISTRIBUTION OF RISK-SHARING AGREEMENTS IN EUROPE, 2009-2012 | 44 |
| FIGURE 3.1. PROCESS OF PERFORMING AN MCDA | 56 |
| FIGURE 3.2. CLASSIFICATION OF MCDA MODELS | 56 |
| FIGURE 4.1. CRITERIA INCLUDED IN MCDA FOR INVESTMENT DECISIONS | 74 |
| FIGURE 4.2. CRITERIA AND CONDITIONS FOR REIMBURSEMENT IDENTIFIED BY THE CITIZENS PANEL | 80 |
| FIGURE 5.1. HIERARCHY OF CRITERIA FOR A HYPOTHETICAL PATIENT | 89 |
| FIGURE 5.2. EXAMPLE OF COMPARISON BY PAIRS TO DETERMINE THE WEIGHT OF THE CRITERIA IN A HIERARCHICAL ANALYSIS | 90 |
| FIGURE 5.3. EXAMPLE OF PARTIAL UTILITY FUNCTION | 91 |
| FIGURE 9.1. WEIGHTING OF QUANTITATIVE CRITERIA OF THE EVIDEM FRAMEWORK AT REGIONAL LEVEL IN SPAIN, ACCORDING TO THE EVALUATION COMMITTEE IN ANDALUSIA, THE BASQUE COUNTRY AND CATALONIA | 167 |
| FIGURE 9.2. IMPORTANCE OF THE CRITERIA FROM THE PERSPECTIVE OF THE PATIENT AND FROM THE PERSPECTIVE OF NATIONAL AND REGIONAL EVALUATORS IN SPAIN | 167 |
| FIGURE 10.1. STEPS CARRIED OUT FOR THE MCDA, THE CASE STUDY | 180 |
| FIGURE 10.2. AVERAGE WEIGHTS USING THE SCALE 1-5 METHOD | 181 |
| FIGURE 10.3. AVERAGE WEIGHTS USING THE 100-POINTS DISTRIBUTION METHOD (NORMALISED X 100) | 181 |
| FIGURE 10.4. COMPARISON OF RESULTS BETWEEN THE TWO WEIGHTING METHODS (NORMALISED AVERAGE WEIGHTS 0-1) | 182 |
| FIGURE 10.5. AVERAGE SCORES OF THE EVALUATED INTERVENTIONS (BASE CASES) | 184 |
| FIGURE 10.6. ESTIMATED TOTAL VALUE USING THE TWO WEIGHTING METHODS | 185 |
| FIGURE 11.1. STRENGTHS OF MULTI-CRITERIA DECISION ANALYSIS | 200 |
| FIGURE 11.2. LIMITATIONS OF MULTI-CRITERIA DECISION ANALYSIS | 202 |

ABBREVIATIONS

| | | | |
|-----------------|---|----------------|---|
| 6MWT | 6 Minutes Walk Test | IQWiG | Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen |
| AHRQ | Agency for Healthcare Research and Quality | ISPOR | International Society for Pharmacoeconomics and Outcomes Research |
| ASCO | American Society of Clinical Oncology | KCE | Belgian health care knowledge centre |
| AVF | Advanced Value Framework | LABA | Long-Acting β 2 Agonist |
| BRAT | Benefit-Risk Action Team | LAMA | Long-Acting Muscarinic Antagonist |
| CADTH | Canadian Agency for Drugs and Technologies for Health | MAUT | Multi-Attribute Utility Theory |
| CDC | Center for Disease Control and prevention | MAVT | Multi-Attribute Value Theory |
| CHAQ | Child Health Assessment Questionnaire | MCD | Multi-Criteria Decision Analysis |
| COMP | Committee for Orphan Medicinal Products | MEA | Managed Entry Agreements |
| COPD | Chronic Obstructive Pulmonary Disease | MPS | Mucopolysaccharidosis |
| DALY | Disability-Adjusted Life Year | MRP | Medication-Related Problems |
| EE | Economic Evaluation | mu | Monetary units |
| EMA | European Medicines Agency | NICE | National Institute for Health and Clinical Excellence |
| EU | European Union | OD | Orphan Drugs |
| EUnetHTA | European network for HTA | OECD | Organisation for Economic Co-operation and Development |
| EVIDEM | Evidence and Value: Impact on Decision-Making | PASFTAC | Programa d'Avaluació, Seguiment i Finançament dels Tractaments d'Alta Complexitat |
| FEV | Forced Expiratory Volume | PBAC | Pharmaceutical Benefits Advisory Committee |
| FVC | Forced Vital Capacity | PBMA | Programme Budgeting and Marginal Analysis |
| G-BA | Gemeinsame Bundesausschuss | PDA | Personal Development Analysis |
| HIV | Human Immunodeficiency Virus | PRO | Patient-Reported Outcomes |
| HTA | Health Technology Assessment | QALY | Quality-Adjusted Life-Years |
| IC | Inhaled Corticosteroids | RD | Rare Diseases |
| ICC | Intraclass Correlation Index | RSA | Risk-Sharing Agreement |
| ICER | Incremental Cost-Effectiveness Ratio | SD | Standard Deviation |
| ICVR | Incremental Cost-Value Ratio | SET | Substitute Enzyme Treatment |
| iFOBT | Immunochemical Fecal Occult Blood Test | TPR | Therapeutic Positioning Reports |
| IHME | Institute for Health Metrics and Evaluation | USA | United States of America |
| | | WHO | World Health Organization |

FOREWORD

I am very pleased to preface this book about multi-criteria decision analysis in healthcare, and I feel sure that it will become a basic reference in our field. Top-level experts in health economics in Spain have contributed their academic, clinical or management vision, giving shape to this work. The document is conceived as an interesting combination between a manual and a reference book. This is what enriches its content, which I foresee will prove to be a great contribution, not only to scientific dissemination but also to practical learning about the use of supporting tools when making health decisions.

The demographic changes that we have been experiencing during the last few decades generate changes in people's needs, due not only to ageing but also to a real change from the epidemiological pattern of non-communicable diseases, which generate chronic conditions of great impact on the quality of life of those who suffer from them and their families. Moreover, the great advances in diagnosis and treatment, and the associated cost of health care, push our healthcare system towards uncertainty about its sustainability if we do not adopt compromises and correct decisions.

In this context, in the current debate about how to combine innovation and sustainability with the needs of people, a tool such as multi-criteria decision analysis now appears.

It is not surprising that these methods are being increasingly used to inform decision-making about financing and prioritisation of techniques, technologies and procedures in the healthcare sector. Although their use began with orphan drugs, their application is essential for all types of pathologies, because it allows us to consider the holistic value of health technologies for society, and explains the importance of each element considered.

This work offers us a broad panorama of the usefulness of multi-criteria decision analysis, which positions it as a method that enriches the decision-making process by providing it with greater transparency, consistency and legitimacy.

Nevertheless, as with any other emerging tool, its use will improve over time, but first we must understand it well and understand its advantages and limitations, through practical examples. In this sense, this book is an undoubted help, by exposing the ideas on which the tool is based and by developing a very coherent perception of the applicability of multi-criteria analysis in the healthcare field. Something that infuses the text with practicality is the inclusion in some chapters of practical examples, applicable to different decision cases which we face in our daily healthcare management at different levels -macro, meso or micro- and which are cases through which the reader's understanding of the model is facilitated.

I convey my congratulations to the authors for the excellent work done, because they have managed to combine the theoretical aspects of the method with practical experiences of its use, both nationally and internationally.

The text facilitates the understanding of key issues that arise when performing or interpreting an analysis of this type, in which different analytical techniques are used to inform decision-making in contexts such as ours, in which there are multiple criteria which may be in conflict. This type of analysis is an opportunity to integrate, among other criteria, for example, the perspective of the patient, into a system that we want and believe to be centred on the person, and on their needs and expectations. Incorporating new information for the analysis that comes from people's experience, values and preferences helps to legitimise the decision-making.

The work ultimately leads the authors to reflect critically about multi-criteria decision analysis, and to develop a compendium of recommendations to enhance its usefulness.

At the Spanish Ministry of Health, Consumption and Social Welfare we are committed to the quality and equity of our healthcare system, to whose sustainability we are all committed. In this sense, multi-criteria decision analysis, complementing other techniques, can help us to make decisions about the prioritisation and use of resources based on evidence, and to improve commitment, transparency and accountability to citizens.

I encourage readers who would like to become more familiar with this topic, or to study it more deeply, to read this book carefully; it explains in a clear and readable way how a multi-criteria decision analysis should be carried out and what practical usefulness it can bring to the decision-making process which we face in healthcare management, with the ultimate aims of ensuring that these decisions serve to improve people's quality of life and contribute to the sustainability of the system.

Paloma Casado Durández

*Deputy Director-General of Quality and Innovation,
Spanish Ministry of Health, Consumption and Social Welfare*



INTRODUCTION

Néboa Zozaya González

Juan Oliva Moreno

Álvaro Hidalgo Vega

CHAPTER

1

The allocation of public healthcare resources – always limited – is a difficult but necessary task, which implies having to decide between options, and therefore having to establish priorities. Healthcare systems must combine making available to the public innovations in the therapeutic field with the protection of the financial sustainability of the system itself. And this, in an environment where demographic challenges (such as population ageing, increased life expectancy or the chronicity of diseases) and technological challenges (which increase treatment options, but at a higher cost) are evident, but not their consequences in the evolution of the systems in the medium and long term. To this, we must add budgetary constraints, always present, but much more explicit in the wake of the recent economic crisis, and the growing importance of clearly including social preferences throughout the evaluation and decision-making processes.

With these premises, healthcare decision-makers use a series of tools to guide their decision-making on planning, prioritisation and allocation of resources. Among the most used instruments are the economic evaluation and the budget impact analysis, but there are other criteria that decision-makers, in one way or another, also often take into account. Among these 'other criteria' are, for example, the severity of the disease, the availability of therapeutic alternatives, the population group affected, the quality of the available evidence, the degree of technological innovation and the value for public health (Golan et al., 2011; Baltussen y Niessen, 2006).

However, when decision-makers affirm that in their decisions they are considering 'other factors', there are immediately doubts about the type of criteria to which they refer, and about the weight which each of them exerts. Frequently, there is a lack of transparency in decisions, whether consciously or unconsciously. In some cases, such as the pricing of orphan medicines, the process has come to be called 'the black box' (Picavet et al., 2014).

Additionally, as citizens' representatives, public decision-makers are called upon to collect and apply social references to their decisions, and to ensure that these are taken in a rational, transparent and reasoned manner, so that society can consider them legitimate and compatible with their economic, moral, social, legal and political restrictions.

Given this scenario, the Multi-Criteria Decision Analysis (MCDA) is presented as an emergent tool to help decision-making. Although its use in the field of healthcare is relatively recent (Adunlin et al., 2015), these methods have been widely used in recent decades, in both the public and private sectors, in the fields of transport, energy, defence, environment, immigration and investments (Thokala et al., 2016).

The MCDA encompasses a set of techniques to help the deliberative decision-making processes, serving as a methodological tool to bring order to decisions. As we shall see in detail throughout the book, the MCDA is a support instrument that helps to specify and inform the preferences inherent in multi-criteria decisions, in a more consistent and transparent manner, but which in no case replaces the taking of decisions.

In the healthcare field, interest in the MCDA has experienced a notable boom in recent years, palpable in the growing number of scientific articles published in this regard. Since 2016, almost as many have been published as in all previous years put together. But this 'fashion' has not been left only in the theoretical framework, it has also begun to be implemented in real practice in many countries, including Spain (Drake et al., 2017; Gilabert-Perramon et al., 2017).

The **objective** of this 'White Paper' is to provide an exhaustive framework that condenses the existing knowledge about MCDA in the field of health. With it, we intend to help readers to understand the usefulness of MCDA, clarify often obscure methodological aspects and identify their strengths, but also their limitations. So the work not only delves into the theoretical part, but also aims to capture an applied vision of MCDA, reviewing how this technique has been used in practice. In the applied part, we analyse MCDA focused on rare diseases separately from those carried out on prevalent diseases, to see whether there are differences depending on the prevalence of the disease.

Another of the book's strengths is that it includes an *ad hoc* practical case, by way of example, where the authors of the work act as evaluators of an MCDA applied to a prevalent disease and a rare disease. The exercise, although applied to fictitious drugs, allows the authors to reflect on the methodology and usefulness of the MCDA in a more reasoned way, complementing the authors' review of the literature.

This work is aimed primarily at decision-makers, researchers and clinicians, and aims to be an intermediate product between a manual and a roadmap, which on the one hand synthesises, in simple language, the existing information and applies it to the Spanish context, and on the other hand, incorporates reflections and recommendations from a group of prestigious healthcare economists.

The **structure** of the book is as follows. After this introductory chapter, Chapter 2 (by Antonio García and Nuria García-Agua) contextualises the concept of the value of healthcare innovations, providing a general framework with which to understand the regulation of innovative medicines. The authors suggest that, considering that an innovation only makes sense if it contributes social value, it is worth asking how this value should be measured. And determining what is considered an innovative medicine depends on different factors. The authors review the regulatory tools and warn that, in Spain, we continue putting more emphasis on costs than on the value that medicines and / or diagnostic and therapeutic procedures can have, which can lead to policies that distort the incentives to innovate and are harmful both in terms of efficiency and in terms of equity.

In Chapter 3, Jaume Puig introduces the reader to the MCDA, presenting the basic concepts that govern its application in decisions about the evaluation and prioritisation of health interventions. The author explains what this type of analysis consists of, how it is done (through different illustrative examples for each method) and for what it can (and cannot) serve, detailing its necessity, scope, strengths and limitations. Thus, this chapter shows the main methods and stages of the MCDA, based on the principles of good practice, but also shows the conditions of its proper use. The chapter also reviews the debate about complementarity or substitutability between MCDA and the logic of opportunity cost.

The choice of criteria or attributes is a key aspect in ensuring that the final result of the MCDA adequately includes all the necessary dimensions. For this reason, in Chapter 4, Jaime Espín delves into this phase, explaining the process and the characteristics that the criteria must meet to avoid overlaps and other problems, and at the same time endow them with legitimacy. In addition, the author explains, through different real examples, some initiatives put in place in connection with the selection of criteria for MCDA in the healthcare field.

Chapter 5, perhaps the most technical in the book, explains in a simple but detailed way the different weighting, scoring and aggregating techniques used in MCDA. Carlos Martín provides the reader with the knowledge necessary to understand how these phases of the MCDA are carried out in practice, through a theoretical framework and examples illustrative of the different methods, which will give the reader an idea of their strengths and limitations. The author places special emphasis on the importance of identifying and adequately managing the different types of uncertainty inherent in any MCDA, in order to improve the validity of the model.

Once the MCDA has been contextualised, and once it has been explained in detail how to address its most important phases, Chapter 6, written by Marta Traperero, offers a review of methodological guides and good practices in MCDA published by different public and private organizations. The author details the main guidelines and recommendations, whose ultimate goal is to contribute to more informed, efficient and equitable healthcare decision-making, and she identifies the challenges still to be overcome in order to achieve greater standardisation and homogenisation of the method.

Chapters 7 and 8 focus on reviewing the practical applications of MCDA that have been made in prevalent and rare diseases, respectively, through individual reviews of the literature. Thus, in chapter 7, Javier Mar reviews the few MCDA that have been performed in the field of diseases such as cancer, chronic obstructive pulmonary disease (COPD), diabetes and cardiovascular diseases, highlighting the emerging use of this tool in this type of prevalent pathologies. For his part, in Chapter 8, Carlos Campillo investigates the use of MCDA in the evaluation of orphan drugs (both simulated and real), also assessing the development of this tool in decisions about price and reimbursement of interventions aimed at rare or infrequent diseases. In addition, the author presents and discusses the arguments about the potential complementarity versus substitution of the MCDA in the evaluation of healthcare technologies.

Next, Antoni Gilabert and Xavier Badía devote Chapter 9 of the book to a presentation of some recent experiences of incorporating MCDA into the making of decisions about the evaluation of drugs, with a special focus on the Spanish healthcare framework. Among other experiences, the authors report exercises with evaluators from the Spanish Agency for Medicines and Healthcare Products and from the Pharmacotherapy Committee of Catalunya, some of which they experienced at first hand. These initiatives serve as a basis for reflecting on the relationship between the MCDA evaluation and decision-making in real practice, and the validity of the methods implemented.

Chapter 10, by Bleric Alcalá and Néboa Zozaya, gathers the results of a practical case included specifically for this book, in which the authors of the previous chapters adopted the role of evaluators. In the exercise, different fictitious drugs aimed at treating a rare versus a prevalent pathology were evaluated, not so much with the objective of comparing the specific results obtained, but more for the purpose of helping us to reflect on the methodology and process of the MCDA in a more enlightened way.

Finally, in the conclusions section, the findings and reflections derived from each chapter are summarised, paying special attention to the fit of the MCDA with the traditional economic evaluation, and some recommendations are made for the future.

REFERENCES

- Adunlin G, Diaby V, Xiao H. Application of multicriteria decision analysis in health care: a systematic review and bibliometric analysis. *Health Expect*. 2015; 18(6): 1894–905. doi:10.1111/hex.12287.
-
- Baltussen R, Niessen L. Priority setting of health interventions: the need for multi-criteria decision analysis. *Cost Eff Resour Alloc*. 2006; 4(1): 14.
-
- Drake JI, de Hart JCT, Monleón C, Toro W, Valentim J. Utilization of multiple-criteria decision analysis (MCDA) to support healthcare decision-making FIFARMA, 2016. *J Mark Access Health Policy*. 2017; 5(1):1-10. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5645903/>.
-
- Gilbert-Perramon A, Torrent-Farnell J, Catalan A, Prat A, Fontanet M, Puig-Peiró R, et al. Drug evaluation and decision making in Catalonia: development and validation of a methodological framework based on multi-criteria decision analysis (MCDA) for orphan drugs. *Int J Technol Assess Health Care*. 2017: 1–10.
-
- Golan O, Hansen P, Kaplan G, Tal O. Health technology prioritization: Which criteria for prioritizing new technologies and what are their relative weights? *Health Policy*. 2011; 102(2–3):126–35.
-
- Picavet E, Morel T, Cassiman D, Simoens S. Shining a light in the black box of orphan drug pricing. *Orphanet J Rare Dis*. 2014; 9: 62.
-
- Thokala P, Devlin N, Marsh K, Baltussen R, Boysen M, Kalo Z, et al. Multiple Criteria Decision Analysis for Health Care Decision Making—An Introduction: Report 1 of the ISPOR MCDA Emerging Good Practices Task Force. *Value Health*. 2016; 19(1): 1–13.
-



THE VALUE OF INNOVATION IN HEALTHCARE

Antonio García Ruíz

Nuria García-Agua Soler

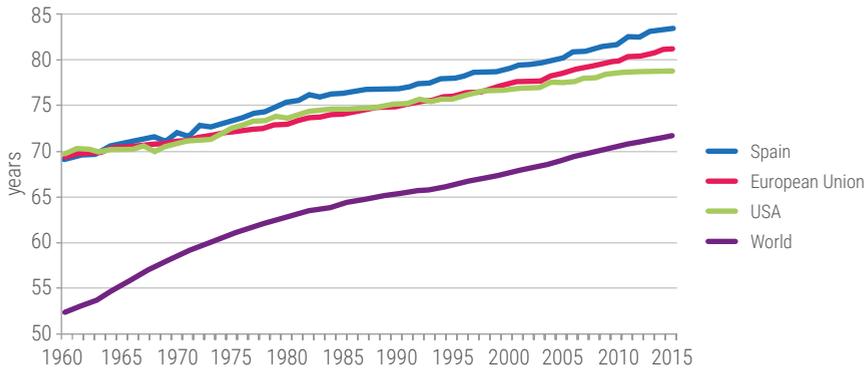
CHAPTER

2

2.1. INTRODUCTION

In the past seventy years there have been prodigious advances in medicine that have contributed to the longevity of the population, with an increase in life expectancy of more than 2.5 years per decade in some countries and an increase in the quality of life (Klenk et al., 2016) (Figure 2.1). The new technologies that were discovered have not been the only determining factor in healthcare outcomes, but they have been the fundamental element on which modern health systems have been built (each time with a less collective and more individualistic approach).

FIGURE 2.1. HEALTH EXPECTANCY AT BIRTH



Source: World Development Indicators (last update Aug 2017).

In fact, it is estimated that healthcare contributes between 10% and 20% to health outcomes, whereas it is other determinants, such as social, biological, environmental and educational ones, among others, that affect a multifactorial equation and can even account for 100% of the causes (McGovern et al., 2014; Remington et al., 2015) (Table 2.1).

TABLE 2.1. DETERMINANTS AND THEIR CONTRIBUTION IN HEALTH

| DETERMINANTS OF HEALTH | PERCENTAGE CONTRIBUTION TO HEALTHCARE OUTCOMES | FACTORS |
|-----------------------------|--|---|
| Healthy behaviours | 30% | Tobacco, physical exercise, diet, alcohol, drugs abuse, sexual activity |
| Healthcare | 20% | Access and quality |
| Economic and social factors | 40% | Education, employment, salary, family and social support, security |
| Physical environment | 10% | Quality of water and air, housing, workplace |

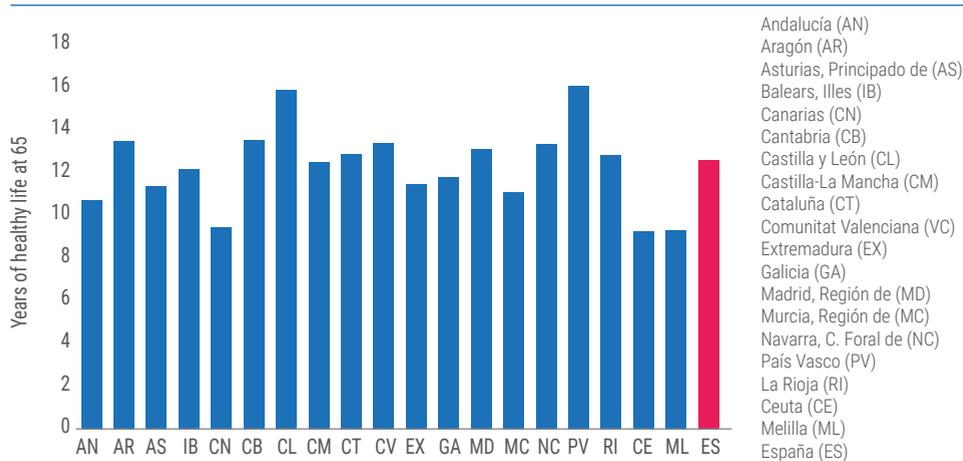
Source: Remington et al., 2015.

In our country, the improvement in life expectancy in the first two-thirds of the twentieth century was largely due to the improvement in living conditions in childhood, but in the last third of the century it seems to have been much more related to the improvement in health

care and its extension to the entire population, especially benefiting those over 65 (Goerlich et al., 2006), exceeding 2.8 years per decade in the last 50 years (Figure 2.1). Nevertheless, there are notable differences between autonomous communities in health outcomes such as years of life without disease after 65 years of age (Figure 2.2).

The effect of population ageing on the use of healthcare services will depend on the relationship between gains in longevity and the health of the population. It is not known whether the additional increases in life expectancy contribute to the improvement in health. Nevertheless, the results suggest that investing in healthy ageing can contribute to containing the growth of healthcare spending, since it can reduce morbidity (Wouterse et al., 2015).

FIGURE 2.2. YEARS OF HEALTHY LIFE AT 65, 2014



Source: INE, 2017.

2.2. DEFINITION OF INNOVATION

Now, we could define what we understand by innovation. According to the Oslo Manual (Guide to the collection and interpretation of data about innovations of the OECD and Eurostat), innovation is the conception and implementation of significant changes in a company's product, process, marketing or organisation for the purpose of improving results. The Frascati Manual (Frascati, 2002) defines it similarly: "technological innovation activities are all scientific, technological, organisational, financial and commercial steps, including investment in new knowledge, which lead to the implementation of technologically new or improved products and processes. R&D is just one of these activities and can take place at different stages of the innovation process. It can act not only as the original source of inventive ideas, but also as a means of solving problems that may arise at any point of the route that leads to innovation".

But this meaning at therapeutic level (with or without drugs) should be understood as something broader and not exempt from subjective appreciations by the various actors that come into play in relation to pharmacological therapy (patients, doctors, pharmacists, nurses, the pharmaceutical industry and healthcare authorities). Thus, a broad concept assumes that an advance in therapeutics would be new treatments that offer health benefits,

in terms of life expectancy and quality of life for patients, when compared with previously existing options.

The need to define innovation in relation to drugs is partly due to the desire to stimulate innovation, recognizing it and rewarding it appropriately. In fact, innovation is increasingly becoming an important consideration in the assessment of new drugs and health technologies – for example, with the introduction of the value-based pricing in the United Kingdom. Not all drugs are innovative, and when they are, they are not innovative to the same degree. We have to decide what is and what is not innovative, promote innovation and decide what is rewarding, and what innovation is beneficial for health and not merely expensive for the system. In this sense, authors such as Ferner (2010) classify the degree of innovation of drugs as high, moderate, slight or none, depending on the type of innovation achieved (Table 2.2).

TABLE 2.2. DEGREE OF INNOVATION

| DEGREE OF INNOVATIVENESS | | TYPE OF INNOVATION | EXAMPLES |
|--------------------------|--------------------|--|--|
| High | | New target or novel mechanism | Selective 5HTagonists (migraine) |
| | | Novel application | Aspirin (prevention of stroke) |
| | | Improved identification of those who are likely to benefit or be harmed (pharmacogenetics) | KRAS gene predicts efficacy (panitumumab, cetuximab). HLA B*5701 predicts adverse effect (abacavir) |
| Moderate | | New type of compound | Monoclonal antibodies |
| | | Fewer adverse effects or interactions | Ranitidine versus cimetidine |
| | | Novel structure | Low molecular weight heparins |
| Slight | Health related | Improved disposition (pharmacokinetics) | Short acting benzodiazepines |
| | | Improved delivery (formulation) | Modified-release formulations |
| | Non-health related | Improved production | Recombinant insulin |
| | | Novel structure | Meptazinol; esomeprazole versus racemic omeprazole |
| None | | Remarketing | Standard release oxycodone |

Source: author's preparation from Ferner et al., 2010.

In the opinion of the International Society of Drug Bulletins (ISDB, 2002), the term 'innovation' should include three concepts:

- **The commercial concept:** any newly marketed me-too product, new substance, new formulation and new treatment methods.

- **The technological concept:** any industrial innovation (including so-called copy medicines), such as the use of biotechnology, the development of new release systems (patches, sprays, etc.), or the selection of an isomer or a metabolite.
- **The concept of therapeutic advance:** a new treatment that offers health benefits for the patient when compared to previously existing options.

Is every innovation the same and should it be 'rewarded' in the same way? There are important factors that should be included in any definition of innovation (Aronson et al., 2012) in order to make a fair assessment of the value of the progression:

- The first aspect is **novelty**, which is not necessarily implied by a newness. For example, aspirin in the prevention of cardiovascular events. Nevertheless, we must be aware that deciding when the novelty is such requires a value judgement.
- The second aspect is **usefulness**. Regarding health, there are two general outcomes that are crucial: health-related quality of life, and survival, and both are a consequence of the clinical benefits and risks (damage):
 - A possible range of usefulness that reflects the probability of these results, from highest to lowest, is the following: 1st) the benefit in a pathology without existing effective treatment; 2nd) the improvement in the treatment of a condition that does not have a sufficiently effective treatment; 3rd) a safer treatment (for example, due to a smaller number of adverse reactions or of interactions between drugs); 4th) a more profitable treatment; and 5th) a more convenient treatment (for example, oral versus intravenous administration).
 - Non-health outcomes should also be considered (improvement of employment, profitability of companies and increase in national wealth-GDP). The different agents involved (patients, pharmaceutical companies, healthcare systems and governments) will surely have different opinions about the order of priorities of this list since the benefit derived from an innovation can affect not only individuals but also society (for example, vaccination and immunisation). Within this section, the cost of innovation must also be of interest, since the value (incremental, as opposed to the best option) must compensate for the (incremental) cost.
 - Innovation and usefulness should be considered separately if we want to identify rewarding innovations: there are important advances that bring a clinical benefit that is worthwhile and there are others that also add a social benefit (decrease in sick leave, increase in productivity, etc.).

A third aspect is how the innovation arises: it can be **revolutionarily or evolutionarily** (also called sustained or evolving). Disruptive innovation occurs when a product emerges drastically in a way that the market does not expect or that exceeds the market's expectations, whereas sustaining or incremental innovation produces improvements in an expected manner and there is no prospect of its altering existing markets. Evolutionary innovations in themselves are not harmful to the market and can be truly innovative, but there is a thin and delicate line between truly innovative drugs and what are known as 'me-too drugs', which are

new, but not innovative. In fact, the small pharmacological differences between successive drugs can eventually lead to a significant difference that can be considered innovative in a certain sense, for example, angiotensin receptor antagonists (ARAs) would not have existed without the prior synthesis of inhibitors of the angiotensin convertase (ACEs).

2.3. THE SOCIAL VALUE OF THE DRUG

True innovations in therapeutics are valuable inventions that usually generate positive externalities, that is, the benefits that they produce go beyond direct users and companies which market the product, and are transferred to society as a whole (Ibern, 2002), and the world is full of good examples (antiretrovirals and HIV, and, more recently, hepatitis C and new therapeutic targets). But the social value of drugs must be measured by the opportunity cost, taking into account both the desired and sought effects (benefits for the health of individuals) and the unwanted effects (medication-related problems [MRP], including adverse drug reactions [ADR], lack of adherence to treatment by patients or therapeutic non-compliance, interactions between drugs, and cost of treatment).

In general, drugs are one more input into the health production function, forming part of a treatment that occurs in a given context and environment, provided and dispensed by qualified professionals, after a medical diagnosis after consultation or diagnostic test, etc. That is to say, it is the system as a whole that contributes the value, the drug being a part (an important one) of this mechanism.

The value of a new drug will come from its ability to improve health, not just from its contribution to reducing or increasing the costs of health care. It will depend on its marginal productivity in the function of health production and on the cost-effectiveness / utility ratio, which varies between patients both in the numerator (costs) and especially in the denominator (effectiveness/utility) (Puig-Junoy, 2002).

But also, as we have said, the value of a new drug (or medical innovation) can have different meanings and scope for the different agents in the system (patients, clinicians, decision-makers, the general population, among others). A drug can be innovative even if it does not save lives: it can improve the self-perceived quality of life, increase the convenience of the patient, improve the interaction with other drugs, free up resources for other uses, etc. There is also no doubt that currently the results reported by the patient (patient-reported outcomes [PRO]) seem to be more important than other results, whether clinical, physiological or reported by any health professional. But we must also recognise that not every innovation in health has to have social value.

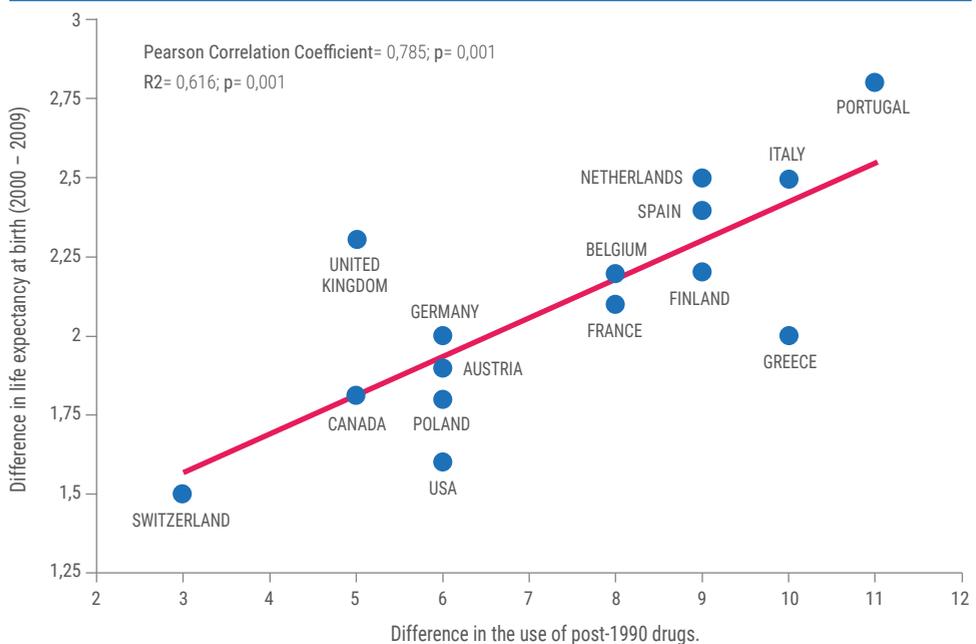
The impact of drugs on patients' quality of life is not always perceived in all its dimensions (sometimes due to a lack of effectiveness of the drug – including PROs). Nevertheless, there are numerous cases in which drugs are the 'culprits' of true quantitative and qualitative changes in the health of individuals.

During the last century, the life sciences and health care eradicated some of the most feared diseases in the world, such as polio and smallpox. In the new millennium, other diseases such as breast cancer, HIV/AIDS, heart diseases and lung cancer (large cells) are no longer the death sentences they once were.

There is no doubt that the pharmaceutical, biotechnology, medical technology, devices and diagnostics companies have helped people to live longer, with less pain and a better quality of life. In some developed countries, medical advances have helped to increase life expectancy by 5 years and to reduce deaths from heart diseases, strokes and breast cancer by more than half in the past 30 years (NCHS, 2013). These advances in medicine and health not only affect drugs, but also include novel surgical procedures or techniques such as those in plastic and reconstructive surgery (Wang et al., 2013), in neurosurgery (Marcus et al., 2015), in asthma and COPD, with the evaluation of new parameters of effectiveness (Bousquet et al., 2016), or in advanced Parkinson's disease with deep brain stimulation, infusion of apomorphine or intestinal infusion of levodopa-carbidopa (Kulisevsky et al., 2013).

Nevertheless, the most commonly used therapeutic resources within the fields of health care and healthcare are drugs. Collectively, the new therapies are the biggest contributors to the increase in life expectancy (Figure 2.3). Between 1960 and 1997 new therapies accounted for 45% of the increase in life expectancy in 30 developing and high-income countries. Between 2000 and 2009, new therapies accounted for 73% of the increase in life expectancy in those countries (Lichtenberg, 2012).

FIGURE 2.3. NEW DRUGS AND LIFE EXPECTANCY AT BIRTH



Source: author's preparation from Lichtenberg, 2014.

From the position of health economics, the distortion that the drug can produce in terms of contributing to equity, accessibility, quality, efficacy and efficiency in the healthcare sector is well known. Moreover, there may not be any other procedure within health care and its problems that has so much final impact, and on which there have been more regulatory and intervening measures, as that involving drugs. So both the healthcare au-

thorities and the real actors of the sector (in which patients should be included) need to focus their attention on this, recognising that the drug, from the perspective of health economics, is a commodity of social interest, which serves to guarantee the right to individual and collective health.

We healthcare professionals can have different opinions about the value of the various diagnostic and treatment techniques for the same pathology. These differences can be fundamentally due to the presence of uncertainty (when there is no scientific evidence about the results of possible alternative treatments or about the value of certain diagnostic tests in specific situations) or ignorance (when there is scientific evidence about the value of the tests or treatments, but the doctor is not aware of it or, even being aware of it, uses other guidelines). So the observed high variability in medical practice is usually a symptom of this strong uncertainty about the effectiveness of treatments, although at other times it may be a symptom of lack of knowledge of the value provided not only by the drug, but also by any diagnostic test or medical procedure, and this, for example, in the emergency services conditions the defensive application of disproportionate technologies (not by the level of severity but by the different diagnostic possibilities).

That is why doubts arise about how drugs or therapeutic interventions really affect the function of health production, in other words, does our health improve, and if so, by how much? Is marginal utility greater with therapeutic innovations? Does this gain in health compensate for the price that we pay for it? Would it be better to invest in healthcare, or in drugs?

Nevertheless, it is not necessary to answer these questions exhaustively: if the value contributed by drugs (in terms of health) increases more than the expenditure generated by their use (including inappropriateness – infra/overuse, non-compliance, lack of adherence, adverse reactions, interactions and variations due to uncertainty, ignorance, organisational problems, etc.), this would improve the efficiency of the healthcare system in a remarkable way. Thus, the growth of health and pharmaceutical expenditure would be compatible with a price reduction if the value contributed increases more than the expenditure (Puig-Junoy, 2002), that is, it would be desirable to increase spending on healthcare only if the social value exceeds the cost of the drug.

In our country, we continue to place more emphasis on costs than on the value that drugs and/or diagnostic and therapeutic procedures can have, which can lead to inefficient policies (Puig-Junoy, 2001). In fact, and despite the time that has elapsed, our drug policies have changed little or almost not at all: the same price-control system, reference prices, co-payments, agreements with laboratories, etc.

2.3.1. The value of medical innovation: saving lives, saving money

Medical innovations, past and future, also have a benefit often not contemplated by healthcare authorities (including politicians): the incalculable savings to patients, families, insurers, employers, governments and hospitals in avoided medical expenses associated with maintaining healthy people or controlling risk factors, especially in chronic conditions (hypertension, diabetes, arthritis, etc.).

Certainly, drugs, therapies, medical technologies, devices and diagnostic tools should serve to keep people healthier, limiting the need for frequent visits to the doctor and avoiding costly hospitalisations, and can help patients avoid costly surgery.

Instead, we hear frequent reports about the high cost of medicines or about new technologies or diagnostic tools which are considered too expensive or unnecessary. We hear that medical innovation involves a high cost and not a saving of costs.

Reality is not shown only in binary format. Drugs, therapies and medical technologies and devices can not only 'save lives', but can also serve to save resources, and these have a measurable economic value. By eradicating or preventing a disease or simply reducing its morbidity, healthcare systems save costs, not only in pharmacological treatments, but also through better management and prevention of more serious complications in an existing disease. When discovering a new treatment, the costs that have been incurred in the treatment of the patient's ongoing medical problems can be almost completely avoided.

A clear example has emerged today (McCombie et al., 2017) with a new category in the prognosis of a chronic disease: 'diabetes in remission'. The authors propose that recognising and accurately coding the reversal of type 2 diabetes is crucial in improving health outcomes and reducing the costs of medical care. The remission of diabetes (which relates to people who no longer have 'diabetes', at least for a period) is clearly achievable for some patients, possibly many, but the concept currently has difficulty in being recognised, even by professionals. Another example appeared with the new treatments used in hepatitis C, which have represented a real breakthrough in therapeutics, with more than 95% of cure rates in routine clinical practice (sustained viral response) (Pinazo, 2017), and although the savings in costs have not yet been evaluated in our country, avoiding a transplant (with all the cost which it generates) will certainly speak in favour of the good cost-effectiveness of these new drugs. Another example is that of sanitary accessories for patients with digestive ostomies of elimination and bag wearers who can return to their job maintaining or even improving their quality of life before surgery, leaving behind the concept of 'ostomate person= retired person' (Montesinos, 2017).

It follows that the development of new treatments and health technologies is one of the most important steps we can take as a society—not only to save lives and improve the quality of life but also to avoid the cost of healthcare.

How many savings does innovation in health produce? There is no simple answer to that question. Nevertheless, there are numerous statistics (academic and governmental) that point to the economic benefits of innovation in the healthcare market.

In an article published by the *Journal of Political Economy* (Murphy and Topel, 2006), it was estimated that, in the last 50 years, medical innovation had been the source of almost half of all economic growth in the United States (USA). Other estimates indicate that, in the Medicare population, reducing the age of the drugs used reduces non-pharmacological spending by 7.2 times more than the spending on drugs. For example, it was estimated that reducing the average age of drugs used to treat a condition increases the spending on prescription drugs by \$18, but reduces other medical expenses by \$129, which repre-

sents a net reduction of \$111 in total spending on health. Most of the savings are due to reductions in hospital expenses (72%) and expenses of medical consultations (22%). So, for every dollar spent on innovative drugs, the total cost of medical care is reduced by \$7.20 (Lichtenberg, 2007). In another study (Civan and Koksall, 2008), also with data from the USA, it is shown that newer drugs increase expenditure on prescription drugs, since they tend to be more expensive than their predecessors. Nevertheless, the demand for other types of medical services decreases, which means that total spending decreases. A one-year decrease in the average age of prescription drugs causes per capita health expenses to decrease by \$31.92. The biggest decrease occurs in hospital spending and health care due to new drugs.

Nevertheless, in the literature we also find contrary arguments. Some argue that the data about the adoption of new drugs saving money, increasing life expectancy and increasing productivity are unreliable and should not be taken into account in healthcare policy decisions (Baker and Fugh-Berman, 2009). They argue that, for example, the ALLHAT study, federally funded, found that chlorothiazide, an older and cheaper diuretic, was superior to new drugs for treating hypertension. The CATIE study, also funded by the government, found that older antipsychotics are as effective as new antipsychotics in treating schizophrenia. New drugs have more risks than older drugs, since problems associated with long-term use or in special populations (for example, the elderly) are revealed only after widespread use.

2.4. INNOVATION AND PREVALENCE OF DISEASE

Since 2004, for the first time, there have been more people in the countries of the European Union (EU) who are over 65 years of age than people who are under 15 years of age. In many of the countries of the world, including low- and middle-income countries, something similar is happening. As a result, health systems in Europe and other parts of the world are facing an ageing population and an increase in chronic, non-communicable diseases associated with economic development and changes in lifestyle.

The World Health Organization (WHO) is therefore asking pharmaceutical researchers to adjust their research and development to take this change into account (Kaplan et al., 2013). Thus we can identify key areas of priority research for pharmaceutical innovation aimed at meeting public health needs, mentioning that this change in the countries of the EU is a symbol for the rest of the world, as more and more people everywhere will grow old and will face similar health problems in the future.

For some time now there has been development of methods that help us to measure the burden of disease. One of them (perhaps the most used) is Disability-Adjusted Life Years (DALYs) as a unique integrated measure of mortality and disability due to a particular disease or condition. One DALY represents one lost year of healthy life, and the burden of disease is a measure of the gap between current health status and an ideal situation in which everyone lives in old age free of disease and disability.

Mortality can also be used as a measure of the burden of disease since it is easy to understand. Nevertheless, it is not able to reflect the burden of pain and suffering experienced by patients with chronic diseases (osteoarthritis, diabetes...).

The greater the burden of disease, the higher the cost of disease for society, and the greater should be the need for research (basic, clinical and therapeutic). Many chronic non-communicable diseases contribute substantially to the burden of morbidity (disability and mortality - DALYs) both in Europe and in the world (Table 2.3):

- Ischemic heart disease and stroke are the biggest contributors to the burden of disease in Europe and worldwide. There are effective drugs to treat cardiovascular diseases, which reduce the incidence of recurrent heart attacks and strokes. Nevertheless, these drugs are not used properly for secondary prevention. It is necessary to investigate how to optimise second prevention treatment through the use of existing drugs.
- In depression, the priority areas for research continue to be appropriate treatment for adolescents and the elderly, reducing side effects and identifying the best treatment strategy for different populations and age groups.
- Stroke, osteoarthritis, Alzheimer's disease, hearing loss, lower back pain, chronic obstructive pulmonary disease and alcoholic liver disease are conditions that represent a great burden of disease, particularly in Europe, for which the currently available treatments are inadequate to reverse or stop their progression. An important challenge for many of these diseases is the absence of specific biomarkers that could be used to identify new pharmaceutical products, to diagnose and evaluate the progression of the disease, and the effect of the treatment. Continued support for basic research into these pathological conditions is therefore necessary.
- Bacterial resistance and some viral pandemics (influenza) continue to be threats to global public health that require a coordinated international effort. The research priorities are the development of new, rapid diagnostic tests, new business models for the R & D of new drugs and vaccines, and the prevention of infections through vaccination, infection control and other environmental measures.
- Malaria and tuberculosis, especially in low- and middle-income countries, are among the high-burden diseases. Tuberculosis is also an important disease in some European countries.
- Diarrhoea, pneumonia, neonatal conditions and maternal mortality are the main factors contributing to the global burden of morbidity. Some existing therapies are often not available in low- and middle-income countries due to health system limitations such as those in healthcare management and accessibility, and other barriers. Meanwhile, the lack of diagnoses and medical attention points create problems in the management of cases.
- For rare diseases, new mechanisms to promote the translation of basic research into clinically important products remain a priority.
- Smoking, alcohol abuse and obesity are risk factors that underlie many of the most common serious diseases that affect both Europe and the world. While prevention efforts should have priority, it is necessary to investigate new drugs to address these risk factors and the pathologies which they exacerbate (for example, COPD, various kinds of cancer, alcoholic liver disease, osteoarthritis and diabetes).

TABLE 2.3. BURDEN OF THE MAIN DISEASES IN THE WORLD, EXPRESSED IN DISABILITY-ADJUSTED LIFE YEARS, 2016

| 20 MOST FREQUENT CAUSES | RATE ADJUSTED FOR AGE, IN 2016 | |
|--|---|-----------------------|
| | DALYs (PER 100,000 POP) | % CHANGE 2006-2016 |
| All causes | 33.641,0 | -16,9 |
| Ischemic heart disease | 2.562,8 | -11,9 |
| Stroke | 1.711,2 | -19,2 |
| Ischemic stroke | 787,5 | -15,8 |
| Intracerebral hemorrhage | 923,6 | -21,8 |
| Lower respiratory infections | 1.326,7 | -34,4 |
| Low back and neck pain | 1.182,7 | -1,4 |
| Diarrhoea | 1.063,1 | -39,9 |
| Sense organ diseases (glaucoma, catarat, macular degeneration, refraction and accommodation pathologies, age-related and hearing loss, others) | 959,3 | -1,9 |
| Traffic accidents (cars, pedestrians, cyclists, motorcycles, others) | 954,5 | -14,7 |
| COPD (chronic obstructive pulmonary disease) | 945,3 | -17,9 |
| Neonatal preterm birth | 892,7 | -24,1 |
| Diabetes | 814,2 | -1,6 |
| Skin diseases | 781,3 | 1,4 |
| Malaria | 794,7 | -30,7 |
| HIV/AIDS | 762,1 | -50,0 |
| Congenital defects | 716,0 | -16,1 |
| Neonatal encephalopathies | 682,2 | -22,0 |
| Migraine | 598,6 | 0,1 |
| Depressive disorders | 597,9 | -3,6 |
| Tuberculosis | 593,1 | -35,3 |
| Lung cancer | 526,1 | -11,7 |
| Falls | 506,5 | -4,6 |
| Hepatitis C | 1,1 (infection) 47,9 (Ca hepatic) 133,5 (cirrhosis) | -22,3 -8,2 -9,4 |
| Adverse effects of drugs | 69,7 | -13,2 |

DALYs: Disability-Adjusted Life Years (years of potential life lost + years lived with disability).

Source: author's preparation from GBD 2016 DALYs and HALE collaborators, 2017.

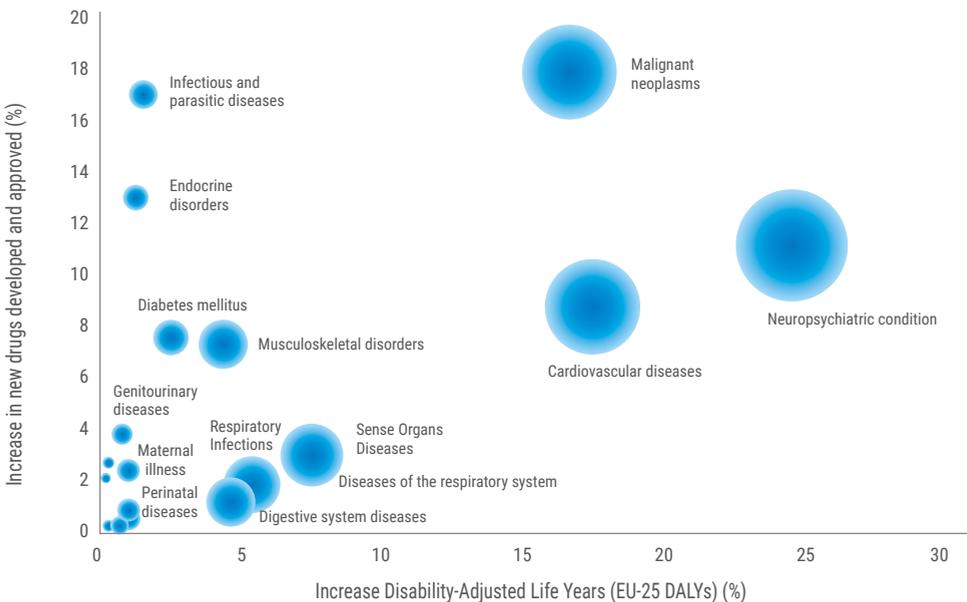
2.4.1. Priority drugs

While prevention remains an important pillar in many of the diseases, more research into the development of new drugs and the improvement of existing ones will benefit everyone. There are still therapeutic gaps, where the treatments for a disease are not sufficiently effective, or may soon become ineffective, or are not suitable for a certain group of patients, or do not exist or present too many or serious adverse reactions.

One study (Catalá-López et al., 2010) analysed whether the development of innovative drugs in Europe focused on the most important diseases from a global public healthcare perspective. The authors reviewed information about new drugs approved by the EU centralised procedure from 1995 to 2009 and assessed the association between authorised drugs and the burden of disease (DALYs), in the EU and throughout the world. Authorisations were considered for the marketing of 520 drugs and 338 active substances. The authors found a positive correlation between DALYs and the development of new drugs ($r= 0.619$, $p= 0.005$) in the EU, and a moderate correlation for low -and middle- income countries ($r= 0.497$, $p= 0.030$) and worldwide ($r= 0.490$, $p= 0.033$).

As shown in the following double-figure (Figure 2.4), the development of new drugs is unequal depending on the pathologies. So the pharmaceutical industry and policy-makers should consider the existing imbalance, establishing priorities from the perspective of public healthcare.

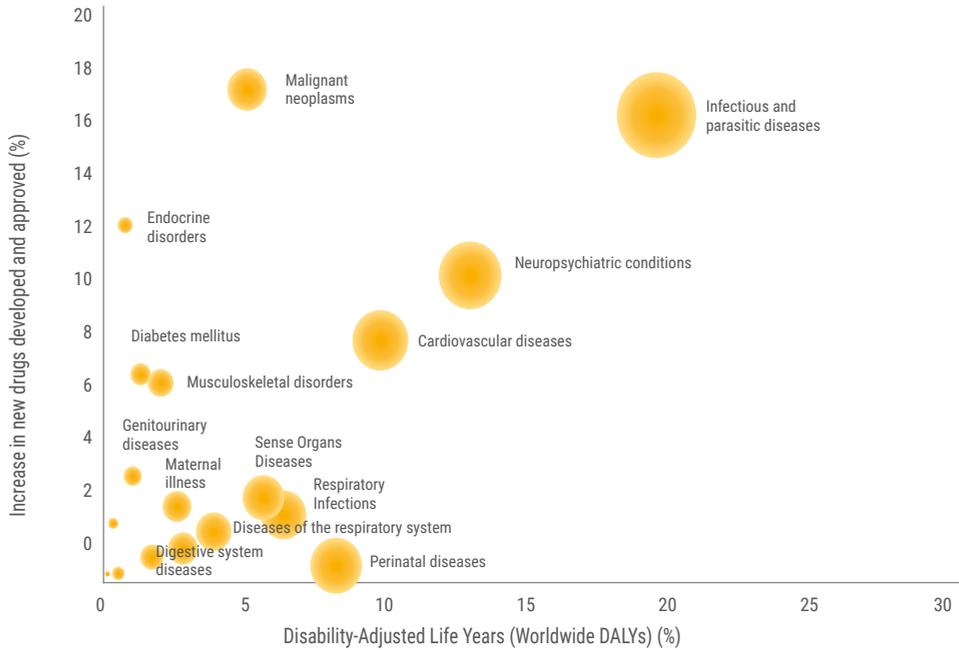
FIGURE 2.4. INNOVATIONS AND BURDEN OF DISEASE (DALYs), 1995-2009



Source: author's preparation from Catalá-López et al., 2010.

Continue

Continued



Source: author's preparation from Catalá-López et al., 2010.

The most neglected diseases (based on years of life adjusted for disability) were neuropsychiatric, cardiovascular and respiratory diseases of the sensory organs, perinatal conditions, respiratory infections and digestive diseases.

We should therefore establish priorities in the R&D of therapeutic innovations based on:

- Drugs to treat conditions for which there is little or no effective treatment, or for which the available drugs are of very limited efficacy or effectiveness. These include rare diseases, which are undoubtedly an important public health problem because of unmet needs. Not only is innovation necessary in this area, but patients do not have universal access to new drugs, since in addition to therapeutic gaps there is difficulty in making decisions about price and reimbursement in rare diseases. So far, there has been little consensus about the most appropriate evaluation criteria, perspective or evaluation process.
- Drugs which have not yet been developed but which are necessary for diseases and conditions that are going to become important public health problems in Europe and in the rest of the world (diseases and conditions associated with ageing, such as heart disease, stroke, cancer, diabetes, osteoarthritis, lumbago, hearing loss and Alzheimer's disease).
- Drugs needed for special groups/situations such as the elderly, children, pregnant women, etc.

Table 2.4 shows the drugs approved by the European Medicines Agency (EMA) during 2016 and which were therapeutic innovations. Figure 2.5 shows the pharmaceutical industry's requests for technical advice from the EMA about drugs in development (EMA, 2017), classified

by therapeutic area. Figure 2.6 shows the burden of disease in Western Europe. As can be observed, there is little alignment between the real innovations and the burden of disease represented by the main illnesses that affect the population in European countries.

TABLE 2.4. NEW DRUGS APPROVED BY THE EMA, 2016

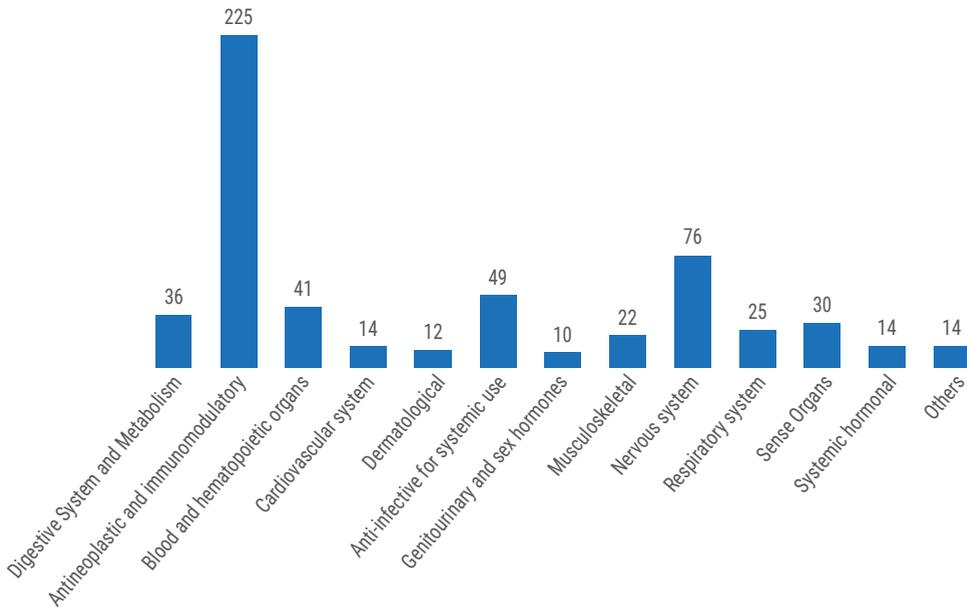
| DISEASE | TRADE NAME | ACTIVE SUBSTANCE | THERAPEUTIC AREA |
|-------------------------------|------------|---|---|
| Cancer | Alecensa | Alectinib | Metastatic non-small cell lung cancer |
| | Darzalex | Daratumumab | Relapsed and refractory multiple myeloma |
| | Empliciti | Elotuzumab | Multiple myeloma |
| | Ibrance | Palbociclib | HER2 metastatic breast cancer |
| | Lartruvo | Olaratumab | Soft tissue sarcoma |
| | Lonsurf | Trifluridine/Tipiracil | Metastatic colorectal cancer |
| | Ninlaro | Ixazomib | Multiple myeloma |
| | Venclyxto | Venetoclax | Chronic lymphocytic leukaemia |
| Infections | Epclusa | Sofosbuvir/Velpatasir | Hepatitis C |
| | Zavicefta | Ceftazidime/Avibactam | Complicated intra-abdominal infection, urinary tract infection, hospital-acquired pneumonia |
| | Zepatier | Elbasvir/Grazoprevir | Hepatitis C |
| | Zinplava | Bezlotoxumab | Clostridium difficile infection |
| Cardiovascular | Uptravi | Selexipag | Pulmonary arterial hypertension |
| Rheumatology | Olumiant | Baricitinib | Rheumatoid arthritis |
| | Truberzi | Eluxadoline | Irritable bowel syndrome |
| Metabolism | Galafold | Migalastat | Fabry disease |
| Haematology | Afstyla | Lonococog alpha | Treatment haemophilia A |
| | Alprolix | Eftrenonacog alpha | Treatment haemophilia B |
| | Idelvion | Albutrepenonacog alpha | Treatment haemophilia B |
| | Zalmoxis | Allogeneic T cells genetically modified | Haploidentical haematopoietic stem cell transplantation |
| Hepatology & Gastroenterology | Ocaliva | Obeticholic Acid | Primary biliary cholangitis |
| Neurology | Ongentys | Opicapone | Adjunctive therapy Parkinson disease |
| Neumology / Allergy | Cinquaero | Reslizumab | Severe eosinophilic asthma |
| Endocrinology | Parsabiv | Etelcalcetide | Secondary hyperparathyroidism |
| Immunology | Strimvelis | CD34+ CELLS | Severe combined immunodeficiency due to adenosine deaminase deficiency |
| Dermatology | Taltz | Ixekizumab | Plaque psoriasis |

Source: author's preparation from EMA, 2017.

There are other factors that are not particularly 'innovo-dependent', such as the polypill in elderly patients, who often require drugs for multiple chronic diseases, but research and treatment guidelines tend to focus more on the disease than on the patient. Or drugs which do

not require cold storage, such as heat-stable insulin and oxytocin for childbirth. This would provide an important benefit to healthcare services in countries that do not have constant access to refrigeration.

FIGURE 2.5. REQUESTS FOR TECHNICAL ADVICE FROM THE EMA, 2016



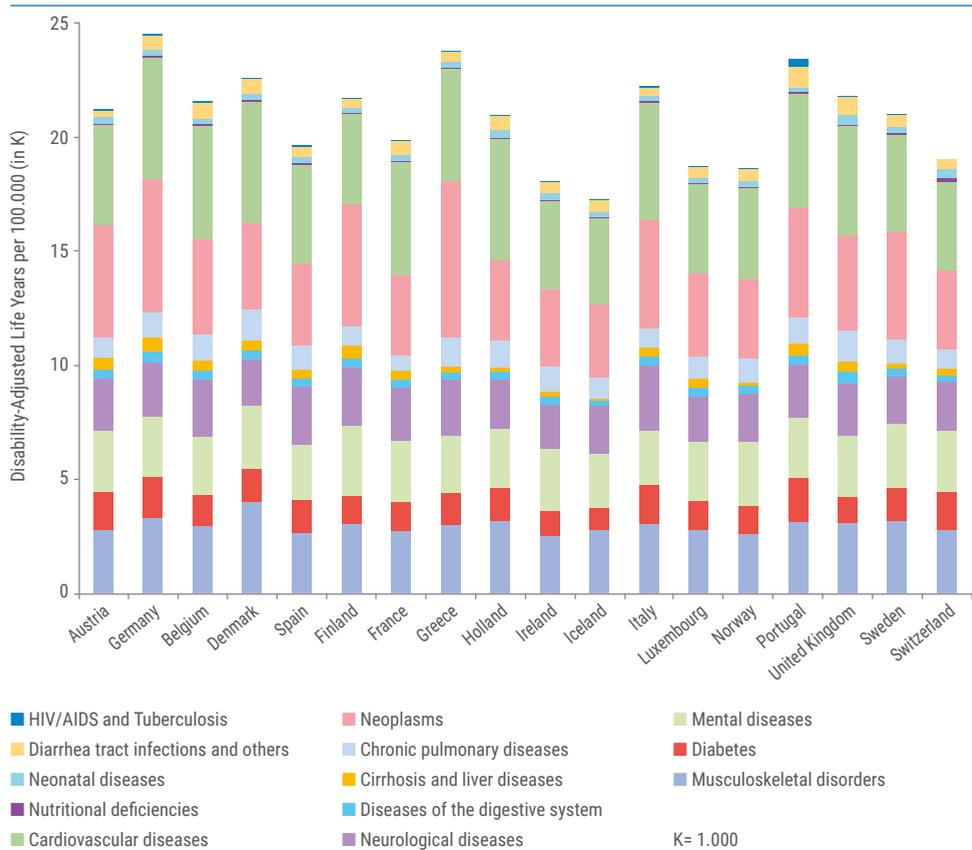
Source: author's preparation from EMA, 2017.

Other additional factors that have an impact on pharmacological innovation could be the following (Wirtz, 2015):

- Innovation and optimisation in the regulatory systems for authorisation in the market;
- The adoption of effective pricing and reimbursement policies to create incentives and take advantage of existing electronic health records to obtain valuable data to improve the safety and effectiveness of medicines (new and existing);
- Adaptive licensing (multiple market authorisations and reimbursement decisions may be required over time instead of a single decision at any one time);
- Prices based on values (outcomes) that can change access to new drugs and the incentives for them. To further develop value-based price-fixing and adaptive licensing. It is necessary to develop new methods for generating evidence, benefit/risk assessment and dialogue with the regulator.
- Need for meaningful participation of patients and citizens in pharmaceutical innovation and access.

We must ensure that the industry develops safe, effective, affordable and adequate drugs to meet future health needs.

FIGURE 2.6. DISEASE BURDEN IN DALYs PER 100,000 IN HABITANTS IN EUROPE, 2016



DALYs: Disability Adjusted Life Years (years of potential life lost + Years lived with disability).

Source: author's preparation from Institute for Health Metrics and Evaluation, 2017.

2.5. TOOLS FOR MEASURING INNOVATION

The increase in the rate of biomedical research that is important for clinical innovation has been a growing concern among researchers and politicians. In response, some of these actors have recently promoted a variety of approaches which they call translational research and translational medicine. Its supporters argue that the productivity of biomedical innovation systems can be reinforced by means of (Vignola-Gagné et al., 2013):

- The extension of large-scale development collaborations;
- The strengthening of clinical experimental platforms;
- Training and support for human capital dedicated to R&D;
- The achievement of greater collective coordination of the research teams.

A medical innovation only makes sense if it provides social value, so we could now ask ourselves: what is the social value of innovation? How can we measure it?

To effectively promote social objectives we should be clear as a society about what those socially desirable goals are, and this would also include our politicians. While we will agree that the appropriate objectives in healthcare policy should be to maximise the net social value of activities in health care, there are authors who apparently consider it sufficient to know the amount of the expenses deriving from health care, while for others this amount should be translated into the value per euro spent (as we have seen previously).

The social value of health care would be the difference between the social value of health improvements resulting from care (social benefits) and the social cost of providing that care. Thus, there may be innovations with a positive social value (socially desirable) and activities that have a negative social value (that is, those that are not worth their social costs), being reducers of value or socially undesirable. It is clear that in order to obtain the social benefit of innovation it is necessary to invest, which will often increase the current social costs. Sometimes, too, the value of innovative activity, in incremental terms, is difficult to observe, because our healthcare system is too complicated to achieve a totally optimal allocation of health resources (whatever the social objective). Moreover, the effects of innovation on health are uncertain, and we assume that the assessments of the social value are based on the values expected from the information available at the time when the assessments were made. So the health outcomes and associated costs occur at different times, with the result that when comparing we have to discount the present value.

A homogeneous way of measuring health outcomes, perhaps the most internationally accepted one, is through the aggregation of quality-adjusted life-years [QALYs]. This implies that social benefit can be obtained by different people equally, regardless of their socio-demographic determinants (income, wealth, state of health, etc.). In order to know the social value of innovation we can use two fundamental tools from pharmacoeconomics: net health benefit and net monetary benefit, and knowing our willingness to pay for a QALY, the two can indicate a priori whether we should adopt innovation or not (Table 2.5).

TABLE 2.5. NET SOCIAL VALUE OF INNOVATION VERSUS ALTERNATIVE

| INCREMENTAL COST | INCREMENTAL QALY | NET BENEFIT OF THE TREATMENT | HEALTHCARE POLICY TO BE IMPLEMENTED |
|------------------|------------------|------------------------------|-------------------------------------|
| C= 1000 mu | X= 0.1 | $(WTP \cdot X - C) > 0$ | Increase use |
| | Y= 0.04 | $(WTP \cdot Y - C) < 0$ | Reduce use |
| | Z= -0.001 | $(WTP \cdot Z - C) < 0$ | Reduce use |

Note: Willingness to Pay= €20.000/QALY.

mu: monetary unit. WTP* Willingness to Pay.

Source: author's preparation.

Nevertheless, the assessment and appraisal of new medical technologies is a debate about political priorities, the specificities of the healthcare system and the expectations of society. So if we want to include the three parts in the decision-making process, in addition to the economic evaluation and the budget impact analysis, studies of the socio-economic burden of

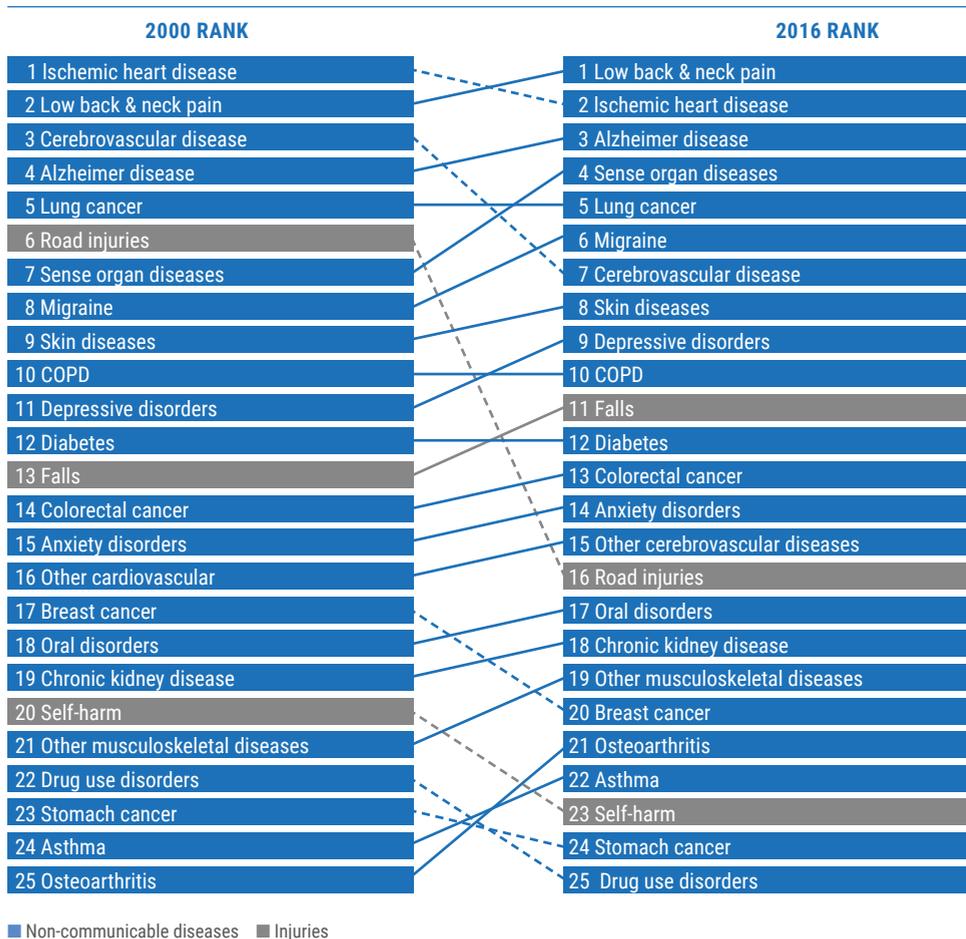
diseases on the individual and on society are the fundamental criteria for the implementation of a correct healthcare policy, as we saw in the previous section.

Economic appraisal and budgetary impact typify health technology and its use and application, while the socio-economic burden of disease is what illustrates the direct budgetary consequences of diseases on health and the social costs associated with the losses in productivity of the patient and/or the caregiver (Angelis et al., 2015).

An accurate knowledge of the socio-economic burden is essential to be able to formulate and prioritise healthcare policies and technologies, and to allocate resources in accordance with budgetary constraints and make the healthcare policy efficient. Thus, Multi-Criteria Decision Analysis (MCDA) is a useful method to help decision-making, prioritising resources (financial and healthcare) with explicit criteria.

By measuring and comparing the socio-economic burden of different diseases on society, healthcare authorities and taxpayers could benefit from optimising the setting of priorities and the allocation of resources.

FIGURE 2.7. DISEASE BURDEN MEASURED IN DALYs, SPAIN 2000-2016



Source: author's preparation from Institute for Health Metrics and Evaluation (IHME) data, 2017.

Data about the burden of disease, efficiency and budgetary impact should be updated in order to understand the economics of diseases and their changing cost structures. This will allow policy-makers to better understand the factors that affect expenditures related to disease and will also allow a better distribution of resources, informed by unmet health needs and healthcare inequalities (Figure 2.7).

In the near future, the pressures on the healthcare system will come from the ageing of the population and from the new technologies. The latter promise better and earlier diagnoses and a greater range of treatment options, but they also have a high cost. Changes can be obtained, but only if the European healthcare system becomes more efficient by channelling resources to where they have the greatest impact on health outcomes. In particular, a new shift towards primary care can help promote a kind of care that is more integrated and patient-centred (OECD/EU, 2016).

2.6. POLICIES TO PROMOTE INNOVATION

Policies which promote innovation have often focused on strengthening the incentives of companies that develop and sell new products by offering them lucrative rights to exclude competitors from the market. Regulators rely on these same companies – and on similar incentives – to gather information about the effects of these new drugs on patients, despite the obvious conflict of interest. The result can be a distorted understanding that leads to the excessive use of new, more expensive technologies.

Not all the possible ways of reducing spending on health depend on a strategy of cutbacks. Currently, most of the policies related to medicine, in our country and others, have been limited to the use of macro-level regulatory instruments (price regulation, reference price systems, global agreements with the pharmaceutical industry, regulation of commercial margins in wholesale and retail distribution, prior authorisation, etc.). Nevertheless, recent technological advances have placed regulators and funders in an excellent position to play a more important role in future innovation to improve health care and reduce costs.

We refer specifically to ‘big data’ about healthcare (and to evidence from the real world). Insurance companies and health service providers keep data about the provision of health services and about outcomes which can provide very valuable information about the true effects of medical treatment without the need for expensive clinical trials. So innovation, on the demand side, by health care providers has tremendous potential for improving the quality of healthcare and for reducing costs. Regulators/funders have access to the health data of millions of patients, which gives us the opportunity to gather new information about the toxicity of drugs, comparative effectiveness (not efficacy), more precise medications (therapeutic targets), and other forms of innovation.

But encouraging innovation on the demand side may require political tools which are very different from those commonly used to motivate companies to develop new products, so innovations on the side of regulators/financers will help not only to improve efficiency but also to obtain important improvements in health results.

Despite the fact that the regulator/payer/financer (both public and private) continually faces gaps in knowledge about the efficacy and safety of new drugs (surrogate variables, comparison

with placebo, time horizon...), and also about the opportunity cost that in a context of limited resources may involve administering a drug to a patient to whom it is not going to provide additional therapeutic value (and will therefore be ineffective), in our country different alternative management formulae have been used: flat rate, price/volume agreement, spending ceiling, silos of pathologies (with programme oriented funding – hepatitis C), and others have been postponed, such as economic evaluation, which is the big absentee from financing and price-setting (basic cost-effectiveness threshold ‘weighted’ by various factors – innovation, burden of the disease, considerations of equity).

So one of the most important regulatory challenges in our country is to convert a binary authorisation model (approved/not approved) into a continuous process of adaptive approval, and to finance on the basis of comparative effectiveness (incremental efficacy). This measure could be of supranational nature, provided that the comparators were the relevant ones for this country and the measurements of results the adequate ones, and as far as possible from intermediate and subrogated measurements for which the causal link with health outcomes has not been adequately established (Puig-Junoy y Oliva, 2017).

Equating the benefit of a drug to its attributable efficacy and assigning a price can be catastrophic for a public system in charge of assessing the real value of an innovation, especially when we know that we are in a situation of asymmetric information in the case of new drugs.

For this reason, a new way of rewarding (financing) new drugs based on the results obtained (especially for those with a high budgetary impact) has been implemented for some time in certain autonomous communities (especially at hospital level). In this way, a desirable pharmaceutical policy should make the industry share responsibility for the sustainability of spending and for health outcomes (Puig-Junoy and Meneu, 2005), in other words, it is necessary to further risk-sharing agreements, not those of a financial nature, but those which are linked to results (no cure, no pay). So innovation should be rewarded in accordance with the value that it brings to patients.

But here the question is multiple: What kind of innovation should be financed? Who should finance it? and How much money are they willing to pay to finance it?

2.6.1. From the regulation of prices to the regulation of healthcare outcomes. Risk-Sharing Agreements (RSAs) as a case study

Risk transfer processes can be framed within the policies promoted by regulators as a reaction to the growth in health spending being greater than the growth in national wealth, based on the fact that regulation through prices and co-payments has been insufficient. Moreover, if the pressure on prices continues, how is true innovation to be stimulated?

A Risk-Sharing Agreement (RSA) is an agreement between a company and a service provider/financer/payer that allows access to a technology under previously specified conditions (Klemp et al., 2011). These agreements generally respond to the need to improve the efficiency of new advances in therapeutics. If payment depends on the health results that the drug provides (that is, if we only finance/reimburse those drugs that improve the health of the population according to previous determined health indicators), the drug as a healthcare tool will be more efficient.

But RSAs also respond to new questions such as: Are the current price and reimbursement systems for drugs able to cope with, the following situations?

- Drugs with multiple indications and a different value in each of them.
- Drugs and healthcare technologies that are administered in combination with others, increasing their value, but also their cost of treatment.
- Drugs with different value for different patients, even with the same indication.

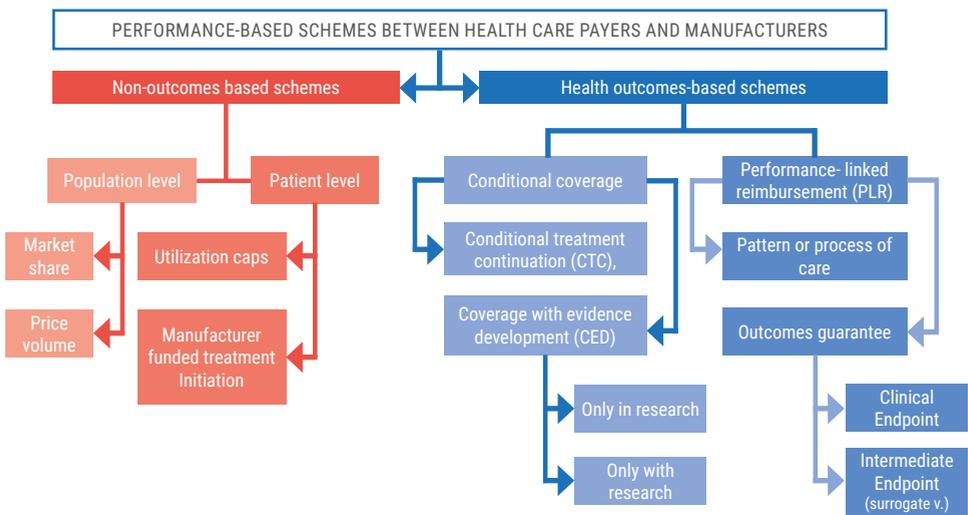
RSA strategies are tools that allow one to compensate for the dilemmas that may exist about the efficacy of an innovative product (real effectiveness in clinical practice). They are particularly useful when there are doubts about the expense involved in a new treatment and how cost-effective it is compared with existing therapies.

In many countries the formulation has begun of various proposals that could be summarised in the following categories:

- Agreements to minimise the budgetary impact (when there is concern about whether the entry of a new technology may have a budgetary impact greater than what is affordable).
- Agreements to minimise uncertainty about use (in the case of drugs, for example, the intention is to avoid off-label uses).
- Agreements to minimise uncertainty about the results (clinical or cost-effectiveness) of the treatment in real clinical practice.

The type of agreement to be subscribed to (when any of these uncertainties occurs) will vary depending on these categories. For example, for the first category, measures that soften the budgetary impact (price reductions, sales volume discounts, expense ceilings, etc.) are recommended. The second would include agreements with limitations of use (agreements on coverage, subject to evidence) to avoid inappropriate use (for example, an agreement aimed at favouring correct use, continuation of treatment in an individual patient subject to their reaching certain clinical 'goals'). And for the third category, either special conditions of access depending on the clinical results or cost-effectiveness (which would constitute a shared risk). Figure 2.8 shows an outline of possible RSAs between pharmaceutical laboratories and funders, based on the types of results desired.

FIGURE 2.8. AGREEMENTS BASED ON RESULTS, BETWEEN FUNDERS AND INDUSTRY



Source: author's preparation from Carlson et al., 2010.

Table 2.6 shows how the risk sharing agreements have advantages and limitations for each of the stakeholders.

TABLE 2.6. ADVANTAGES, BARRIERS AND DIFFICULTIES IN RISK-SHARING AGREEMENTS

| | FUNDERS | PROFESSIONALS | PATIENTS | PHARMACEUTICAL INDUSTRY |
|---------------------------|---|--|--|---|
| Advantages | <ul style="list-style-type: none"> Access to innovation. Limiting uncertainty about the economic impact. Generating incentives for cost-effective use. Aligning the incentives of the industry with those of the healthcare system. Generating experience and knowledge about results. | <ul style="list-style-type: none"> Having effective diagnostic and/or therapeutic alternatives. | <ul style="list-style-type: none"> Access to innovative treatments. | <ul style="list-style-type: none"> Gaining access to the financing of innovative products. Stable framework for financing innovation. Obtaining information and knowledge about its innovative products in a real environment. Minimising costs in promotional areas. |
| Barriers and difficulties | <ul style="list-style-type: none"> Generalisation of payment models based on risk transfer is a slow process. Design of a microagreement: case by case. Valid for some cases, but difficult to extend to all products (technology and/or drug). High technical complexity of the risk-sharing agreement: availability of information systems, data confidentiality and transparency. Negotiating between financier and company generally long and with extensive legal agreements. Difficulty in defining 'performance', both in its measurement and in the follow-up. | | | |

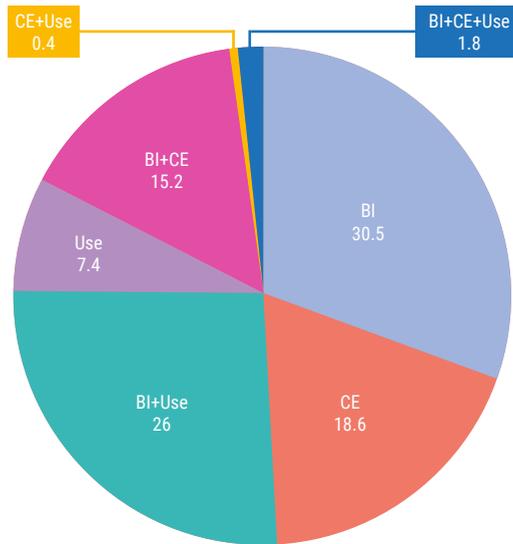
Source: author's preparation.

In European countries, a wide variety of instruments are being used to address the uncertainty arising from the lack of information about budgetary impact, profitability, use in actual clinical practice and access. And despite the number of agreements applied, there is little information

about the impact of these systems and whether they achieve their objectives (although in some cases this has been made transparent: Multiple sclerosis risk-sharing scheme: a costly failure, BMJ 2010).

In addition, the limited amount of information available in the public domain is hampering learning in the countries that apply it, and the ability of patients to participate in the process. The following figure shows the RSAs applied in Europe from 2009 to 2012 (Ferrario and Kanavos, 2013), according to the objectives that they wish to achieve.

FIGURE 2.9. DISTRIBUTION OF RISK-SHARING AGREEMENTS IN EUROPE, 2009-2012



BI= Budgetary Impact
 CE= Cost-Effectiveness
 Use= Monitor use in clinical practice

Source: author's preparation from Ferrario and Kanavos, 2013.

In a study of the different schemes (Managed Entry Agreements - MEAs) for gaining access to new drugs, comparing the United Kingdom, Italy and Spain, the authors (Tolley and Palazzolo, 2014) found that out of a total of 95 agreements, 56% were for oncological drugs, 12% for the musculoskeletal system, 10% for the ophthalmological area, and 7% for other therapeutic areas. The majority of access agreements that were awarded in the United Kingdom were not based on results (76%). In Italy and Spain, all the agreements were based on results, Italy focusing exclusively on RSAs.

RSAs can be very positive because they can serve to allow patients access to innovative drugs by providing an adequate reward for innovative effort and by ensuring the sustainability of the health system. And based on the experience in recent years with RSA, and on real-world data, *"it has been demonstrated that the focus of the drug policy and the pharmaceutical provision*

of health outcomes has allowed the development of good and powerful mechanisms for the registration of treatments that have allowed the monitoring of health outcomes, their comparison and the performance of benchmarking in order to adopt the best practices”, and “the development of a payment and incentive system based not simply on product and activity but on health outcomes achieved” (Gilabert, 2017).

RSAs should not become a quick solution for introducing expensive drugs, but should be integrated into a process of access to new drugs that starts from researching into the needs and passes to the forecast, the appraisal of the healthcare technology, the fixing of prices and reimbursement, and the surveillance of new technologies and medicines (serious adverse effects).

These new approaches to the use of the EU's current regulatory framework for medicines are designed to accelerate the approval process and improve the proper access of patients to new drugs. This is considered particularly applicable in areas of high medical need, in which it is difficult to collect data through traditional pathways, and in which large clinical trials would unnecessarily delay access to patients who are unlikely to benefit from the drug. In this way, we can reinforce the opportunity cost in decision-making, and improve efficiency by knowing the added value that a drug provides in relation to its cost.

2.7. CONCLUSIONS

As we have seen throughout this chapter, the research and development of new drugs is progressing adequately, especially in pathologies of the developed world, but there are still areas of therapeutics that are not fully covered, and others in which we lack authentic innovations that offer true benefits for patients when compared with existing options.

The industry and society in general should recognise that, although drugs have been on many occasions true promoters of big leaps ahead in the improvement of people's life expectancy and quality of life, they are nevertheless just one more tool in terms of determinants in health outcomes, making sense of public health policies, including at social level.

We have also explained how an accurate knowledge of the socio-economic burden of a disease should be essential when formulating and prioritising the policies on medicines and health technologies, as well as when allocating resources in accordance with budgetary constraints and achieving greater efficiency in healthcare policy. In this context, MCDA can also be a useful methodological tool to help decision-making.

From our perspective, the healthcare policies of 'coffee for everyone' should cease to be a maxim, and should accept that true healthcare equity will be achieved when it affects health outcomes, not equity in access or equity in use. But for this change of paradigm to take effect, demand on the part of society is also necessary.

The regulation of prices based on incremental health outcomes would thus favour true innovations, and of course, these must be recognised and properly reimbursed. Applying some of the new forms of management, such as RSAs or different schemes for access to drugs based on real-world data, can help optimise health outcomes, and thus improve the efficiency of pharmacological treatments and health technologies, whether new or not.

Finally, our actions should be aimed at guaranteeing the sustainability of healthcare systems (in the long-term), maximising patients' access (static efficiency) to the most appropriate drugs for them (value) and doing this by adequately rewarding innovation (dynamic efficiency). Added value and health outcomes are therefore two fundamental elements that should be included in any current healthcare policy.

Acknowledgements. The authors thank Francisco Martos Crespo and Ana Montesinos for their comments on a previous version of this chapter.

REFERENCES

- Angelis A, Tordrup D, Kanavos P. Socio-economic burden of rare diseases: a systematic review of cost of illness evidence. *Health Policy*. 2015; 119: 964–79.
- Aronson JK, Ferner RE, Hughes DA. Defining rewardable innovation in drug therapy. *Nature Review*. 2012; 11: 254.
- Baker D, Fugh-Berman A. Do New Drugs Increase Life Expectancy? A Critique of a Manhattan Institute Paper. *J Gen Intern Med*. 2009; 5(24): 678–82.
- Bousquet et al. Scaling up strategies of the chronic respiratory disease programme of the European Innovation Partnership on Active and Healthy Ageing (Action Plan B3: Area 5). *Clin Transl Allergy*. 2016; 6:29.
- Carlson JJ, Sulliva SD, Garrison LP. Linking payment to health outcomes: A taxonomy and examination of performance-based reimbursement schemes between healthcare payers and manufacturers. *Health Policy*. 2010; 96: 179-90.
- Catalá-López et al. Does the development of new medicinal products in the European Union address global and regional health concerns? *Population Health Metrics*. 2010; 8: 34.
- Civan A, Koksal B. The Effect of Newer Drugs on Health Spending: Do They Really Increase the Costs? 2008; MPRA Paper No. 6846.
- European Medicin Agency. Annual report 2016; 2017.
- Ferrario A, Kanavos P. Managed entry agreements for pharmaceuticals: The European experience. EM-iNet; 2013.
- Ferner RE, Hughes DA, Aronson JK. NICE and new: appraising innovation. *BMJ*. 2010; 340: b5493.
- Frascati Manual: Proposed standard practise for surveys on research and experimental development, 6th edition. 2002. Paris: OECD.
- Hay, S et al. Global, regional, and national disability-adjusted life-years (DALYs) for 333 diseases and injuries and healthy life expectancy (HALE) for 195 countries and territories, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. *Lancet*. 2017; 390: 1260–344.
- Gilabert-Perramón A. El "real world data" permite distinguir entre los medicamentos que más aporten al paciente, facilitar su uso y conseguir los mejores resultados en salud. *iSanidad*. Jueves 20 de abril de 2017. ASP. Disponible en: <http://isanidad.com/86100/el-real-world-data-permite-distinguir-entre-los-medicamentos-que-mas-aporten-al-paciente-facilitar-su-uso-y-conseguir-los-mejores-resultados-en-salud/>
- Goerlich Gisbert FJ, Pinilla Pallejà R. Esperanza de vida en España a lo largo del siglo XX: las tablas de mortalidad del Instituto Nacional de Estadística. Working paper 11. Fundación BBVA. 2006.
- Ibern Regás P. Incentivos para la innovación en el mercado farmacéutico. On: Jaume Puig-Junoy. Análisis económico de la financiación pública de medicamentos. 1st ed. Barcelona: Masson S.A; 2002.

Institute for Health Metrics and Evaluation (IHME). 2017. Global Burden of disease study 2016 (GBD). Data resources. Available from: <http://www.healthdata.org/>.

Instituto Nacional de Estadística (INE), 2017. Disponible en: <http://www.ine.es/welcome.shtml>.

International Society of Drug Bulletins (ISDB). Declaración de la Sociedad Internacional de Boletines de Medicamentos sobre avances en terapéutica con medicamentos. 2002. Available from: <http://www.isd-bweb.org/documents/uploads/Declaracion/Spanish-DeclaracionISDB.pdf>

Kaplan W, Wirtz VJ, Mantel-Teeuwisse A, Stolk P, Duthey B, Laing R. Priority medicines for Europe and the World 2013 update. Geneva: World Health Organization; 2013.

Klomp M, Fronsdal KB, et al. What principles should govern the use of managed entry agreements? *Int J Tech in Health Care*. 2011; 27(1): 77-83.

Klenk J et al. Changes in life expectancy 1950-2010: contributions from age- and disease-specific mortality in selected countries. *Popul Health Metr*. 2016; 14:20.

Kulisevsky J et al. Enfermedad de Parkinson avanzada. Características clínicas y tratamiento. Parte II. *Neurology*. 2013; 28(9): 558-83.

Lichtenberg FR. Benefits and costs of newer drugs: an update. *Managerial and Decision Economics*. 2007; 28: 4-5.

Lichtenberg FR. Pharmaceutical innovation and longevity growth in 30 developing and high-income countries, 2000-2009. NBER Working Paper. 2012.

Lichtenberg FR. Pharmaceutical innovation and longevity growth in 30 developing and high-income countries, 2000-2009. *Health Policy Technol*. 2014; 3: 36-58.

Marcus HJ et al. Technological innovation in neurosurgery: a quantitative study. *J Neurosurg*. 2015; 123(1): 174-181.

McCombie L et al. Beating type 2 diabetes into remission. *BMJ*. 2017; 358: j4030.

McGovern L, Miller G, Hugues-Cromwick P. Health policy brief: The relative contribution of multiple determinants to health outcomes. *Health affairs*. August 2014. Available from: <https://www.rwjf.org/en/library/research/2014/08/the-relative-contribution-of-multiple-determinants-to-health-out.html>.

Montesinos Gálvez AC. Estudio sobre la incorporación a la vida laboral en el paciente ostomizado [Doctoral Thesis]. Málaga: Departamento de Farmacología y Terapéutica Clínica, Repositorio Institucional de la Universidad de Málaga; 2017.

Murphy KM, Topel RH. The Value of Health and Longevity. *J Polit Econ*. 2016; 114(5): 871-905.

National Center for Health Statistics. Health, United States, 2012: With Special Feature on Emergency Care. Hyattsville; 2013. Available from: <https://www.cdc.gov/nchs/data/health/us/17.pdf>

OECD/EU, Health at a Glance: Europe 2016 – State of Health in the EU Cycle, Paris: OECD Publishing; 2016.

Pinazo L. Experiencia en práctica clínica real de los nuevos antivirales de acción directa en hepatitis crónica C. [Doctoral Thesis]. Málaga: Departamento de Farmacología y Terapéutica Clínica, Repositorio Institucional de la Universidad de Málaga; 2017.

Puig-Junoy J. Por una política del medicamento integrada y de calidad. *Gest Clin y Sanit*. 2001; 3(1): 3-4.

Puig-Junoy J. Gasto farmacéutico público. On: Jaume Puig-Junoy. Análisis económico de la financiación pública de medicamentos. Barcelona: Masson S.A; 2002.

Puig-Junoy J, Meneu R. Aplicación de los contratos de riesgo compartido a la financiación de nuevos medicamentos. *Gest Clin y Sanit*. 2005; 7(3): 88-94.

Puig-Junoy J, Oliva J. La evaluación económica de medicamentos en España: el traje del emperador. 2017. Disponible en: <http://nadaesgratis.es/sergi-jimenez/>.

Remington et al. The county health rankings: rational and methods. *Population Health Metrics*. 2015; 13:11. DOI 10.1186/s12963-015-0044-2.

Tolley C, Palazzolo D. Managed entry agreements in UK, Italy and Spain. *Value Health*. 2014; 17: A323-A686.

Vignola-Gagné E, Rantanen E, Lehner D, Hüsing B. Translational research policies: disruptions and continuities in biomedical innovation systems in Austria, Finland and Germany. *J Community Genet*. 2013; 4:189–201.

Wang Y, Kotsis SV, Chung KC. Applying the Concepts of Innovation Strategies to Plastic Surgery. *Plast Reconstr Surg*. 2013; 132(2): 483–490.

Wirtz VJ. Priority medicines for Europe and the World: setting a public-health-based medicines development agenda. *Journal of Pharmaceutical Policy and Practice*. 2015; 8(Suppl 1):K4.

Wouterse et al. The effect of trends in health and longevity on health services use by older adults. *BMC Health Services Research*. 2015; 15: 574.



MULTI-CRITERIA DECISION ANALYSIS: WHAT IS IT AND WHAT IS IT FOR?

Jaume Puig-Junoy

CHAPTER

3

3.1 INTRODUCTION

The purpose of this introductory chapter about Multi-Criteria Decision Analysis (MCDA) is to present the basic concepts of its application in decisions about the evaluation and prioritisation of drugs, medical technologies and healthcare programmes. First, the need for a 'multi-criteria' rationale, as opposed to the rationale of cost-effectiveness analysis and of opportunity cost, is analysed. Secondly, it describes what an MCDA consists of, as well as the main methods and stages for carrying out an analysis of this type. Thirdly, a hypothetical illustration of the use of MCDA is presented in order to show the conditions for its appropriate use in each of the three main methods. It concludes by reviewing the conceptual and empirical debate, in the field of health technology evaluation, about the complementarity or substitutability between the cost-effectiveness analysis and the MCDA.

3.2 MULTIPLE CRITERIA IN HEALTHCARE DECISION-MAKING

In recent years, several OECD countries have adopted measures to incorporate economic evaluation (EE) into the set of health technology assessment (HTA) tools that guide the strategies for coverage and reimbursement of drugs and medical technologies (Auraaen et al., 2016; Angelis et al., 2017). Many of them, in addition to comparative or incremental efficacy and the incremental cost-effectiveness ratio (ICER), also require and use budget impact studies which, from the outset, should not be interpreted so much as a conflict between the rules of efficiency and budgetary rules, but, instead, as the way in which insurers take into account the opportunity cost. Cost-effectiveness thresholds cannot be established independently of budget availability in a context of maximising health outcomes (Culyer, 2017). However, the magnitude of the budgetary impact in itself does not provide a good representation of the opportunity cost in relation to lost profits (McCabe et al., 2007). Thus, for example, the case of the new oral agents in the treatment of chronic hepatitis has highlighted the fact that the cost-effectiveness analysis does not adequately take into account the budgetary impact, since it focuses its attention on individual treatments without assessing the impact on the healthcare system as a whole (Neumann and Cohen, 2015). The OECD countries have been adopting different decision-making processes in their respective attempts to include information coming from the economic evaluation of health technologies (Angelis et al., 2017).

From the social perspective, the concepts of therapeutic usefulness and degree of innovation of the new medicines must be related to the social value which they add compared with the available treatment alternatives, and to the added costs that they entail, that is, their ICER. The rationale of incremental cost-effectiveness may be appropriate for the decisions of public insurers about coverage for a particular treatment, the price that one is willing to pay for it, and the clinical situations and groups of patients for which it is recommended. The incremental cost-effectiveness analysis and the establishment of a threshold indicative of the maximum cost that one is willing to pay for 'quality-adjusted life-years' (QALYs) gained – as a reference to the opportunity cost – are the essential elements of this rationale, which does not require setting the price of new drugs at the threshold of willingness to pay.

In practice, HTA procedures to assess and evaluate clinical evidence differ in each country in terms of, for example, the levels of evidence required, the variables of clinical outcome

adopted (mortality, morbidity, survival or quality of life), the criteria for the choice of comparator to evaluate comparative efficacy or the use of measurements of clinical benefit such as QALYs or discrete scales of classification of the added value of innovations. In the same vein, the stakeholders involved (citizens/patients, clinical and economic experts, public payers/public regulators, healthcare professionals/provider institutions/industry) and their role and influence in the evaluation process are variable from one HTA to another. It is not surprising that on many occasions HTA agencies reach different conclusions about the value of the clinical benefit of the same technology.

These decisions of the HTA agencies face some common problems, added to the more specific controversies about the function and the details of the application of economic evaluation techniques, among which two stand out: the limited availability of evidence at the time of carrying out the evaluation, and the difficulties in establishing and justifying a cost-effectiveness threshold above which it is considered that innovations are not cost-effective (Auraaen et al., 2016).

Although cost-effectiveness is widely recognised as a necessary element to guide decision-making with limited resources, none of the HTA agencies that take EE into account use the ICER as the only measurement. This practice recognises that the criterion of efficiency/opportunity cost is a necessary but insufficient condition to guide the allocation of health resources and the need to consider distributive effects on costs and results. From a conceptual point of view, the vast majority of economic evaluations carried out neither incorporate criteria of equity (one year of life or a QALY are valued independently of the disease, the age and the group of patients), nor claim to be, at least from an extra-welfarist approach, the only decision criterion with a metric based solely on the ICER. The very logic of using cost per QALY relies on extra-welfarist theories, and this approach is the one that provides theoretical bases for the inclusion of other criteria in the MCDA (Culyer, 1991; Garau and Devlin, 2017). The main justification, beyond the specific limitations of the QALY – which do not need to be reviewed in detail here – (Angelis and Kanavos, 2016; Nord, 2017), is that the ICER corresponds only to the criterion of efficiency (Devlin and Sussex, 2011; Drummond et al., 2015).

The cost-effectiveness approach implies prioritising not only according to effectiveness but also according to the balance between the comparative costs and health outcomes of one intervention and those of its best alternative, but it does not imply a single threshold (see, for example, the case of the criteria for treatments at the end of life or for rare diseases), nor does it imply that the cost-effectiveness threshold is the only criterion for prioritisation (Culyer, 2017; Neumann and Cohen, 2015). Thus the decision to cover or finance a treatment with an ICER above the threshold could be interpreted in this context as a measure of the opportunity cost of taking into account the 'other' factors in the decision-making in addition to the cost-effectiveness.

In practice, decisions about the allocation of health resources of HTAs take into account, in addition to comparative efficacy and safety, the ICER and the budgetary impact, five groups of 'other' factors: a) the incidence, prevalence and severity of the illness; b) the affected population group; c) the availability of therapeutic alternatives; d) the quality of the available evidence; and e) the degree of technological innovation (Devlin and Sussex, 2011; Regier and Peacock, 2017; Garau and Devlin, 2017). This list of factors could be somewhat more

extensive in certain therapeutic areas, for example, in the case of orphan drugs (Paulden et al., 2014) and is also quite variable between countries, and even between decision-makers.

An international review of the literature of 11 countries about criteria for prioritising technologies groups them into three major groups of principles: a) need, adequacy and clinical benefits; b) efficiency (including cost-effectiveness); and, c) equality, solidarity and other ethical or social values (Golan et al., 2011). Table 3.1 presents the enumeration of allocation principles and the criteria that are associated with them according to this review of the literature.

Decision-making, whatever the level of resource allocation, requires the prioritising and weighting of these criteria in such a way that implicit interchange relationships are established between them. For example, the Dutch healthcare system uses four prioritisation criteria in a rather complex decision process: care must be necessary, effective and efficient, and cannot be left to individual responsibility (Stolk and Poley, 2005). France classifies pharmacological innovations into three levels: essential, important and ease of administration, associating with these categories public financing of 100%, 65% and 35% respectively (Sandier et al., 2004).

TABLE 3.1. MAIN CRITERIA AND 'OTHER' CONSIDERATIONS USED INTERNATIONALLY TO PRIORITISE NEW HEALTH TECHNOLOGIES

| PRINCIPLES OF ALLOCATIVE JUSTICE | CRITERIA |
|---|--|
| Need | General Severity of the condition Availability of alternatives |
| Appropriateness | Efficacy and safety Effectiveness |
| Clinical benefits | General Effect on mortality (life saving) Effect on longevity Effect on health-related quality -of-life |
| Efficiency | Cost -effectiveness /benefit Budgetary impact Cost |
| Equality | General Accessibility to the service Affordability to the individual |
| Solidarity | |
| Other ethical or social values | Autonomy Public health value Impact on future generations |
| OTHER CONSIDERATIONS | |
| Quality of the clinical and economic evidence | |
| Other considerations not elsewhere classified | Strategic issues consistency with previous decisions and precedents |

Source: Golan et al., 2011.

A pragmatic and widely-shared conclusion from the practice observed in the HTA agencies, without going into considerations of its theoretical foundations, is not whether multiple criteria should be taken into account, but, rather, how they should be taken into account with

transparency and consistency and without omission of relevant information. In this sense, for example, when an agency of HTA affirms that in its decisions 'other' factors are considered in addition to cost-effectiveness, the decisional problem is far from being solved, and there are immediate doubts about the manner of applying this multi-criteria approach.

Moving away from the approach of cost-effectiveness or opportunity cost towards the multi-criteria approach, far from solving the problem, makes it essential to give an adequate response to the need to be explicit about the 'other' factors that influence decisions: it requires improving the transparency and accountability of the decision-making process; improving the consistency of the process; taking account of the preferences of the different social agents involved in decision-making and giving precise signals to the industry about the aspects of innovation that are valued (Devlin and Sussex, 2011; Angelis and Kanavos, 2016). By way of example, in many multi-criteria decision-making processes, the way to incorporate the 'other' factors in the decision-making process is not clear in respect of: a) the weight assigned to each of these factors and how they are balanced among themselves and with respect to the ICER; b) whether the effect on the decision is of an additive or multiplicative nature; and c) whether the list of explicit criteria, when there is an explicit list, is or is not exhaustive (Devlin and Sussex, 2011). For example, in the case of England and Wales, when the decisions taken by the National Institute for Health and Clinical Excellence (NICE) in the period 2004-2009 are analysed retrospectively, a significant association is observed between the recommendations of that agency and four variables: a) evidence of clinical superiority of the primary outcome variable in clinical trials; b) the ICER; c) the number of drugs included in the same assessment; and d) the year of the evaluation (Cerri et al., 2013). In Australia, the retrospective review of the decisions taken between 1993 and 2009 (n=425) by the Pharmaceutical Benefits Advisory Committee highlights the fact that an increase of A\$10,000 in cost per QALY reduces the average probability of public funding from 37% to 33%, but if it is a treatment for a life-threatening disease that has no effective alternatives, then these two conditions together are considered as being equivalent to an increase of A\$46,000 in the cost per QALY (Harris et al., 2015). Contrary to what some authors have affirmed (Drummond et al., 2007) for the case of chronic diseases, it does not seem true that the real experiences of HTA are based on the application of a pure criterion of efficiency in a unique way (Laupacis, 2009).

The recent profusion of scales for measuring the innovation value of new drugs ('innovometers'), quite possibly in response to proposals for pricing based on value and the high prices of technological innovations (Neumann and Cohen, 2015; Walton et al., 2016; Zaragoza and Cuéllar, 2017; Angelis and Kanavos, 2017), especially in the field of oncology, is a good example of the risks of inappropriate use of multi-criteria decision-making processes. The non-systematic and *ad hoc* use of value dimensions and the lack of transparency in judgments and preferences about value can often lead to inconsistencies and arbitrariness in the process of evaluating these scales (Angelis and Kanavos, 2016). Simple comparisons of some of these scales of innovation value highlight the fact that there is no consensus about the criteria that must be taken into account, and the fact that they use different strategies to weight the criteria and calculate an overall score. The simple sum of the value scores of each criterion does not necessarily result in an overall score that is consistent with the preferences of the parties (Neumann and Cohen, 2015). Also, in most cases these scales lack any theoretical basis for the measurement of value, which would require, among other things, "*an estimate of the rate at which stakeholders are willing to forgo one attribute of health for another*" (Walton et al., 2016). All these limitations and criticisms directly affect, for example, the quantitative

scale of innovation value of new drugs recently proposed for Spain by a working group led by the Ministry of Health, Social Services and Equality (Zaragozá and Cuéllar, 2017).

The difficulty of decision-making lies in the multiplicity of criteria to evaluate the alternatives, with objectives that may enter into contradiction and with different groups of stakeholders involved in the decision-making, with different preferences. The traditional methods of MCDA constitute a tool at the service of this type of decisions by providing a formalised method of assisting multi-criteria decision-making that comes from operational research models (Belton and Stewart, 2002). In this sense, the need for MCDA in the HTA, despite what some authors claim (Angelis and Kanavos, 2016), cannot be sustained as an alternative to the sole rationale of cost-effectiveness or opportunity cost – which has hardly ever been a unique logical proposal by the economic literature or used in a unique way by the agencies – but can be upheld as an instrument to bring order (transparency, consistency and being comprehensive) in an explicit manner into multi-criteria decisions (Belton and Stewart, 2002; Thokala and Duenas, 2012; Drummond et al., 2015; Regier and Peacock, 2017). MCDA is a set of techniques to help the deliberative processes of multi-criteria decision-making, that is, a support instrument without prescriptive value, and which does not replace decision-making.

Although the application of MCDA to healthcare decisions is relatively recent (Devlin and Sussex, 2011; Marsh et al., 2014; Adunlin et al., 2014; Wahlster et al., 2015; Thokala et al., 2016), these methods have been widely used during recent decades, both in the public sector and in the private sector, as an aid to decisions about transport, immigration, investments, the environment, energy, defence, etc. (Thokala et al., 2016; Communities and Local Government, 2009). It is necessary to recognise that part of the recent interest in MCDA falls within the political debate in order to get away from some HTA processes which, according to some stakeholders, are too concerned with the threshold of the maximum cost per QALY. One has to recognise that there are results or benefits beyond those related to health, and to explore methods of measuring everything that may be relevant for patients. The theoretical and political debate about value-based assessment in the United Kingdom could be interpreted as an attempt to complement the maximisation of QALYs with the explicit consideration of other factors or attributes of value such as the severity of the disease.

In spite of the above, it would be a mistake to regard MCDA merely as a simple alternative to a cost-effectiveness analysis: the correct performance of an MCDA is much more complex than most simple exercises that have so far been disseminated in the literature. Most issues to be resolved and most cost-effectiveness problems affect MCDA with the aggravating circumstance that they extend to other attributes or dimensions: *“Without a proper assessment of the other attributes of benefit forgone, decisions may reduce both health and the other attributes of benefit that originally motivated the use of MCDA”* (Garau and Devlin, 2017).

3.3 WHAT IS MCDA?

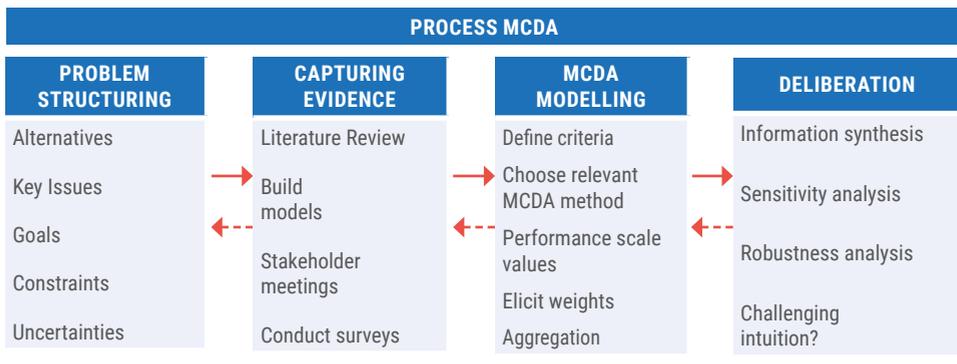
The starting point that justifies the use of MCDA by decision-makers who have to choose between two or more alternatives is that they take into account more than one objective when judging the desirability of a given alternative. Rarely, one alternative is superior to the other or others in respect of all the objectives, that is, it is dominant with respect to the others. The most common situation is that each alternative satisfies the objectives at different levels, and the decisions involve conflicts and relationships of exchange between the degrees of fulfil-

ment of the objectives. In any environment, multi-criteria or not, the choice of one alternative over the others implies an opportunity cost. Thus, the logic of applying MCDA to evaluate alternatives has as its starting point some simple basic assumptions (Regier and Peacock, 2017). First, decisions are made in a context of limited resources, and any decision involves forgoing the benefits of the others (opportunity cost). Secondly, the objectives or criteria that the decision-makers take into account correspond to their discretionary scope and cannot be determined in a normative way by means of ethical or economic theories such as utilitarianism or social justice. Thirdly, an alternative or programme is not a homogeneous benefit, but is described by its multiple characteristics as the combination of several levels for each criterion or dimension. And fourthly, it is possible to establish the relative importance of each criterion and the relationships of exchange between them that allow scores to be obtained and alternatives to be ordered in consequence.

The common purpose of MCDA methods is to take into account explicitly the multiple criteria involved in individual or group decision-making (Belton and Stewart, 2002; Thokala et al., 2016). In a somewhat broader way, Devlin and Sussex, 2011, define MCDA as a set of methods and techniques to help decision-making, applicable when they are based on more than one criterion, which explicitly deal with the impact on the decision of each of the criteria applied, as well as the relative importance of each of them. Two MCDA conditions are particularly relevant for HTA: the replicability and the transparency of the decisions. In fact, the same measurements of quality of life related to state of health that are used to calculate QALYs in a cost-utility analysis are still a form of MCDA – for example, the scale of quality of life EQ-5D-5L takes into account six criteria (survival and five quality-of-life criteria), with five scores for each of them (Drummond et al., 2015).

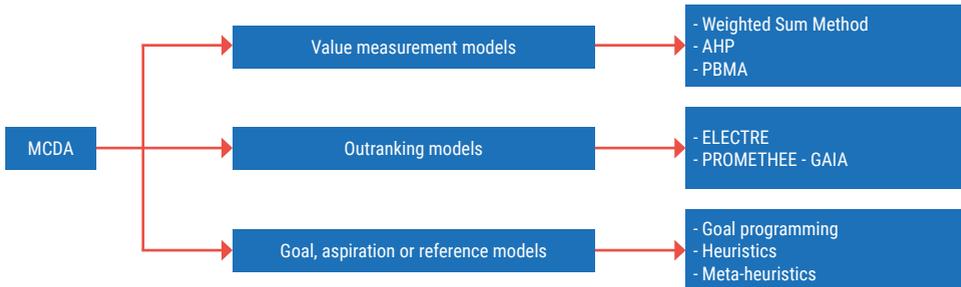
In a multi-criteria decision context, there are three main issues to be resolved (Drummond et al., 2015), which will be decisive in reaching the recommendation or decision. First, the criteria that will be used in the MCDA, that is, the criteria of the benefit or result that will be evaluated together with the improvement in health. Secondly, how to assign weights to the value of the attributes of each criterion. And thirdly, the characteristics or dimensions of benefit that are lost or forgone if additional costs are incurred (opportunity cost).

Any application of MCDA in the healthcare sector or in other sectors includes the following phases (Figure 3.1): identification of alternatives and decision criteria (structuring the problem), construction and use of the model, and development and action plans. The criteria included in the structuring of the problem must fulfil some essential conditions: they must be non-redundant, independent, complete, operational and measurable (Belton and Stewart, 2002; Regier and Peacock, 2017). Each criterion must contribute to the result or benefit independently of the others and avoiding duplication. In the phase of modelling or constructing the model, the information or evidence collected is the subject of quantification and will be used as input in mathematical models to identify the best alternative by incorporating explicit weights and scores of the criteria and attributes. The way in which these models are constructed is what differentiates the different MCDA methods. Their construction involves the production of behavioural models that quantitatively represent the preferences or value judgments of the decision-makers, which, ideally, should reflect the preferences of society. These models have in common that preferences are expressed for each criterion of each alternative and that the aggregation model allows the criteria to be compared with each other in order to combine the estimates of preferences.

FIGURE 3.1. PROCESS OF PERFORMING AN MCDA

Source: Angelis and Kanavos, 2016.

The main MCDA methods can be classified into the following three groups (Devlin and Sussex, 2011; Thokala and Duenas, 2012; Mühlbacher and Kaczynski, 2015; Thokala et al., 2016; Garau and Devlin, 2017): models for measuring value, outranking models and models by objectives or reference levels (Figure 3.2).

FIGURE 3.2. CLASSIFICATION OF MCDA MODELS

AHP: analytic hierarchy process. ELECTRE: Elimination Et Choix Traduisant la Réalité (Elimination and Choice Expressing Reality). PBMA: programme budgeting and marginal analysis. PROMETHEE-GAIA: preference ranking organization method for enrichment evaluations.

Source: Thokala and Duenas, 2012.

The **value measurement models** calculate and compare numerical scores which synthesise the overall value of each alternative as an expression of the degree to which one is preferred to another. The scores of each of the individual criteria are aggregated into a figure that represents the overall value of the alternative. This is the most widely preferred method, almost the only one, in HTA, with the additive aggregation method (the weighted sum method) being the most used. The additive method of aggregation requires that the condition of additive independence be satisfied, that is, that the conflicts between two criteria do not depend on the level of the other criteria (Diaby and Dias, 2017). In a systematic review of the applications of MCDA in the healthcare sector, for the period 1980-2013, it is observed that 60 of the 66 studies selected only used value functions (Diaby and Dias, 2017). Budget programming and

marginal analysis and the analytical hierarchical process are similar techniques that can be included within this group of methods.

Two alternative methods of measuring value functions are identified: the multi-attribute utility theory (MAUT) and the multi-attribute value theory (MAVT). The main difference between the two is that the MAUT uses utility functions which take into account the decision-makers' attitudes to risk, using the concept of lotteries, as opposed to the MAVT, which builds an overall value function for each alternative to obtain the overall score of each alternative based on the decision criteria, using the concept of intensity of preferences. The MAVT's models for measuring value require strict compliance with certain conditions related to criteria and weights: independence and transitivity of criteria preferences and weights that fulfil the requirements of the exchange relationship between criteria. The independence of preferences requires that the decision be based on criteria in which the alternatives appear as different. The relative weights of criteria i and k (w_i, w_k) must be such that the ratio of relative weights (w_i/w_k) represents the change in the value of the score of criterion k for alternative A , $v_k(A)$, which is necessary to compensate for a loss of one unit in the value of the score of criterion i , $v_i(A)$. The techniques for obtaining these weights are analysed in detail in Chapter 4 of this book. Likewise, the methods of scoring and weighting the criteria are described in detail in Chapter 5 of this book and in Marsh et al., 2017. The theoretical bases of the various methods can be found in Regier and Peacock, 2017.

The use of value measurements or functions implies accepting that a low result in one criterion can be compensated with a better result in one of the other criteria. These methods may not be appropriate when these compensatory effects are not considered adequate for the decision process: for example, when the criteria relate to different stakeholders (patients vs. professionals or hospitals) or when the criteria relate to very different dimensions of the value (economic versus social) (Diaby and Dias, 2017).

In **outranking models**, the alternatives are initially compared by pairs in terms of each criterion, in order to confirm the degree of preference (dominance) of one relationship over the other for each specific criterion. If the two alternatives are very similar, they cannot be compared. Next, one adds the degree of preference of the different criteria between the alternatives, in order to establish the overall level of preference of one over the other. This method is based on a direct comparison of the characteristics of the alternatives, and it is appropriate for the HTA despite having been little used so far.

Modelling by objectives or reference levels is based on determining which alternative is closest to predetermined levels of results for each criterion. The use of a cost/utility threshold in pricing based on value would be similar to the way of proceeding in this group of methods based on mathematical programming techniques. Programming by objectives implies minimising deviations from the objectives, taking into account the relative importance of each objective or criterion. The two main techniques used in this group of methods are programming by objectives and lexicographic programming by objectives, which differ in the way of prioritising and reaching the optimal solution (Thokala and Duenas, 2012).

Table 3.2 presents the comparison of these three groups of MCDA, in relation to the weights used, the measurement of the performance of the criteria, the complexity of the MCDA model, the presentation of results and the treatment of uncertainty.

TABLE 3.2. COMPARISON OF MCDA METHODS

| | VALUE MEASUREMENT MODELS | OUTRANKING MODELS | GOAL PROGRAMMING |
|---|---|--|--|
| Weights | Swing weights are used to capture both the effect of measurement scales and the importance of the criteria. Weights need to satisfy the preferential independence of criteria and the trade-off requirements. | Weights are uninfluenced by the scale of the value functions. They convey the relative importance of criteria in the assertion that one alternative is better than the other. Weights do not have to satisfy any conditions. | Weights are attached to the deviations and represent the relative importance of criteria by specifying an overall measure of deviations from the goals. Weights do not have to satisfy any conditions. |
| Measuring the performance of the criteria | Performance scores $v_i(a)$, monotonic functions of the attribute values $z_i(a)$, need to be developed for all criterion i . Significant effort is needed to develop these performance scores. | Outranking approach can use either performance value scores $v_i(a)$ or the attribute values $z_i(a)$, saving on the effort needed to develop performance scores. | Goal programming method operates directly on the attribute values, $z_i(a)$. No need to develop performance scores. |
| Complexity of the MCDA model | Weighted sum approach is easy to understand and use by the decision makers. The parameters can be changed in real time to observe their effect. | Intuitive and easy to follow. With right software, assumptions can be changed and results can be observed almost instantaneously. | Easy to understand but requires significant computational time to provide results. Real-time updating is not possible. |
| Presentation of the results | Easy to follow and enables further deliberation, well suited for good visual presentation of the results. | Moderately easy to follow. Can be presented visually but difficult with multiple alternatives. | Results easy to follow, but they cannot be represented visually. |
| Incorporating uncertainty | Probabilistic sensitivity analysis can be used to propagate parameter uncertainty quite easily. | Moderately difficult to include uncertainty, needs specialist software. | Quite difficult to include uncertainty, complex stochastic programming techniques are needed. |

Source: Thokala and Duenas, 2012.

From a practical point of view, the phases for the performance of an MCDA can be summarised as the following five (Angelis and Kanavos, 2016), although there are authors and guides that present more detailed schemes or stages (Devlin and Sussex, 2011; Thokala et al., 2016): problem structuring, model building, model assessment, model appraisal and action plans.

In the first stage, structuring the problem, in which researchers and decision-makers intervene to establish the context of the decision: the problem about which a decision has to be made and the objectives to be pursued, as well as the decision-makers and parties involved. For example, the decision problem could be to assess the benefits and costs of a new technology from the social perspective and compared with usual clinical practice, in order to identify the intervention that contributes more value to the healthcare system. The decision-makers, in this context, could be the payers or the insurers; and the parties involved would be health professionals, patients and their caregivers, the supplier industry and experts in methods.

In the second stage, model building, researchers and decision-makers also intervene. This phase consists basically of the selection of criteria and attributes that reflect the decision-makers' objectives and concerns, the selection of alternatives and the obtention of evidence for the results of the alternatives for the selected criteria. For example, in the evaluation of a new drug compared with a previous, more effective one, the selected criteria could be the therapeutic benefit, the safety profile, the burden of the disease, the level of innovation, the socio-economic impact and the quality of the evidence.

The third stage of the MCDA consists of the assessment of the performance of options against the identified criteria: the score for each criterion that provides information for the intra-criterion comparison, and weighting of the criteria according to their relative importance that contributes information for combining criteria. This phase generally requires the construction of value functions by means of various techniques (see Chapter 5 of this book), transforming the values of the results into scores on the value scale. The appraisal requires the obtention of an indicator of added value from the combination of scores and weights, the technical details of which are different depending on the MCDA method adopted (Diaby and Dias, 2017). The result of this phase, subject to the performance of a sensitivity analysis, consists of an arrangement of the alternatives based on the value score obtained with the MCDA.

3.4. WHAT CAN MCDA BE USED FOR? A SIMPLE ILLUSTRATION

MCDA can be useful as an aid to decision-making in the healthcare sector, which, in addition to the HTA, can include authorisation decisions, about prioritisation of coverage or financing of benefits, prioritisation of access to treatment for patients, classification of diseases, allocation of resources for R&D, etc. (Thokala et al., 2016; Marsh et al., 2016; Mühlbacher and Kaczynski, 2016). Castro et al., 2017, present an interesting description of three case studies of the application of MCDA in the HTA for various countries and regions (Colombia, Lombardy and Belgium), which are explained in detail in sections 4.3 and 7.3.7. Also, a review of the experiences in the application of MCDA in the prioritisation of benefits and treatments in low- and middle-income countries can be found in Tromp et al., 2017.

Included in this section is a simple illustration of a hypothetical case of using the various MCDA methods to evaluate two pharmacological treatments A and B, being A the current treatment and B the new treatment. This illustration is based on a summary presentation of the case study published by Thokala and Duenas in 2012. Table 3.3 presents the main characteristics of the two drugs under evaluation based on five criteria: cost/effectiveness (in terms of net benefit), equity, innovation, adherence to treatment and quality of evidence. The cost-effectiveness ratio is presented as a net benefit calculated from a willingness to pay per QALY or 20,000 monetary units (mu). The net monetary benefit is the difference between the monetary value of the QALYs and the cost of treatment (Hounton and Newlands, 2012). Let us assume that the net benefit of drug A depends on the price (x) according to the following function: $f(p) = 25,000 - 1,000x$. If the logic of cost-effectiveness is used alone, treatment B would be recommended. Nevertheless, let's see what might happen if the three MCDA methods mentioned above were used. In this hypothetical example, the opportunity cost is not included as a criterion, but only the net benefit of both alternatives.

TABLE 3.3. CHARACTERISTICS OF THE DRUGS IN THE APPRAISAL PROCESS

| CHARACTERISTICS | DRUG A $z(a)_i^*$ | DRUG B $z(b)_i^*$ |
|------------------------|----------------------|----------------------|
| C/E (in terms of NB) | £15,850 | £25,600 |
| Equidad (%) Equity (%) | 0.14 | 0.08 |
| Innovation | Innovative | Less Innovative |
| Patient compliance (%) | 0.93 | 0.85 |
| Quality of evidence | Good | Good |

C/E: cost-effectiveness.

* $z(a)_i$ and $z(b)_i$ are values of attribute i for drug A and drug B respectively.

Source: Thokala and Duenas, 2012.

As a step prior to the application of the three MCDA methods, table 3.4 presents the performance scores for each of the five criteria according to an objective measure or scale. For each criterion, desirable performance levels have been identified. When there is a linear relationship between the value of the attribute and the value of the result of the criterion, as in the case of adherence, the same value of the attribute can be taken as a measure of result. In most cases, for which this linear relationship cannot be assumed, it is necessary to construct a scale to represent the performance of each alternative in which the highest values are preferred. For criterion i , the performance score $v_i(A)$ is a non-decreasing function of the attribute value $z_i(A)$. More generally, for any criterion i the score of any alternative is defined as $v_i = f(z)$, the function f being the same for the alternatives (pharmacological treatments) compared. These scores are generally presented standardised on a scale ranging from the least desirable to the most desirable value (1, 10 or 100). Here we shall assume that we have defined these functions and that the scores are those presented in table 3.4. Once these scores have been obtained, we are ready to start applying each of the three MCDA methods.

TABLE 3.4. PERFORMANCE SCORES OF DRUGS

| CRITERION (i) | DRUG A $v(a)_i^*$ | DRUG B $v(b)_i^*$ |
|-----------------------|----------------------|----------------------|
| CIE | 0.72 | 0.84 |
| Equity(%) | 0.14 | 0.08 |
| Innovation | 0.91 | 0.62 |
| Patient compliance(%) | 0.93 | 0.85 |
| Quality of evidence | 0.82 | 0.79 |

C/E: cost-effectiveness.

* $v(a)_i$ and $v(b)_i$ are values of the result of attribute i for drug A and drug B respectively.

Source: Thokala and Duenas, 2012.

Value measurement models. In order to achieve a single overall score for each alternative, the performance levels of both drugs are constructed for each criterion (preference modeling) as we did in table 3.4 (partial value function), in this case, on a scale between 0 and 1. A partial value function reflects how the value of an attribute varies for the decision-maker along the scale of measurement. It can be a function for an attribute such as quality of life, or decreasing, as for cost. Next, to each criterion a weighting or relative weight w_i is assigned that depends on the scale of the value function of each criterion and indicates its relative importance. Next, the partial value functions are aggregated, taking into account the weightings. If an additive aggregation function (weighted sum approach) is adopted, then: $V(a) = \sum_{i=1}^n w_i \cdot v_i(a)$. Assuming that we have identified the following relative weights (w_i) using the most appropriate technique: cost-effectiveness (8), equity (1), innovation (3), adherence to treatment (2) and quality of evidence (3), then one obtains $V(a) = 12.95$ and $V(b) = 12.73$. With this method, alternative A is preferred to alternative B, contrary to the recommendation with the sole use of cost-effectiveness, since the best results of A in the other 4 criteria more than compensate for the disadvantage in the cost-effectiveness criterion.

Outranking approach. The first step in estimating the agreement or disagreement is to construct the matrix of outranking relations from the individual scores of the alternatives in each criterion (Table 3.4), as shown in table 3.5. With this method, alternative A is preferred to B only if $v_i(A) - v_i(B)$ is greater than a certain 'indifference threshold' (t_i). In our case, for example, if the threshold for the quality of the evidence is 0.05, then the difference observed is less than this threshold, and the alternatives for this criterion cannot be compared. This method also requires the estimation of relative weights for each criterion, which may be different from those of the value measurement model since they only represent the relative importance of the different criteria that allows us to affirm that one alternative is better than the other (they do not need to be exchange relations between criteria). In this example, the concordance index $C(A,B)$ is calculated by the ELECTRE I method, which consists of the ratio between the sum of the weights of the criteria in which A is at least as good an alternative as B (Q weights) and the sum of weights of all

criteria (m weights): $C(A,B) = \frac{\sum_{i \in Q(A,B)} w_i}{\sum_{i=1}^m w_i}$ Alternatively, we could also have calculated a

discordance index $D(A,B)$ which can be defined as: $D(A,B) = \begin{cases} 1 & \text{if } v_i(B) - v_i(A) > t_i \\ 0 & \text{otherwise} \end{cases}$

For our hypothetical example, the concordance index of A against B is $C(A,B) = 8/18 = 0.44$ and the concordance index of B against A is $C(B,A) = 10/18 = 0.56$. Supposing now that we define a veto threshold for the cost-effectiveness criterion $t_1 = 0.1$ table 3.4 shows that B is better than A in cost-effectiveness by 0.12, which is higher than the veto threshold t_1 . If, for example, the threshold of agreement C' were less than 0.56: then, with this method it could be concluded that alternative B is better than A provided that the thresholds of the other criteria are respected and in accordance with the threshold of agreement adopted.

TABLE 3.5. OUTRANKING RELATIONS AND WEIGHTS

| CRITERION (i) | WEIGHTS w_i^* | DRUG A | DRUG B |
|------------------------|-----------------|--------|--------|
| C/E (in terms of NB) | 10 | | ✓ |
| Equity(%) | 2 | ✓ | |
| Innovation | 1 | ✓ | |
| Patient compliance (%) | 3 | ✓ | |
| Quality of evidence | 2 | ✓ | |

C/E: Cost-Effectiveness. NB: Net Benefit.

* w_i^* is the weighting of criterion i.

Source: Thokala and Duenas, 2012.

Goal programming. With this method, first, objective values are defined for each criterion of the alternatives, understood as the desired levels of results for each criterion. In table 3.6 these values of the g_i objective are defined. Thus, in our example we define objectives for the criteria cost-effectiveness, equity and adherence. We shall assume that whereas cost-effectiveness can be modified through changes in price, equity and adherence cannot be modified.

TABLE 3.6. ATTRIBUTES OF DRUGS, THE GOALS AND WEIGHTS AGAINST DIFFERENT CRITERIA

| CRITERION (i) | DRUG A $z(a)_i^*$ | DRUG B $z(b)_i^*$ | GOALS g_i | WEIGHTS w_i^{**} | WEIGHTS w_i^{***} |
|---------------------------|----------------------|----------------------|----------------|-----------------------|------------------------|
| C/E (in terms of NB) | £15.850 | £25.600 | £20.000 | 10 | 0 |
| Equity(%) | 0.14 | 0.08 | 0.20 | 0 | 5 |
| Innovation | Innovative | Less Innovative | - | 0 | 0 |
| Patient compliance (%) | 0.93 | 0.85 | 0.95 | 0 | 5 |
| Quality of evidence | Good | Good | - | 0 | 0 |

C/E: Cost-Effectiveness. NB: Net Benefit.

* $z(a)_i$ and $z(b)_i$ are values of attribute i for drug A and drug B respectively.

** Weights of the first level of priority.

*** Weights of the second level of priority.

° Ω^i and Ω^j are the weights related to the deviations of criterion i from the objective g_i .

Source: Thokala and Duenas, 2012.

We shall use the lexicographical method of goal programming, with cost-effectiveness being the criterion with the highest priority and the others with the next priority. We know that the net benefit of A (C/E) varies with the price according to the following function: $f(x) = 25,000 - 1,000x$ (where x is the unit price of A). If we use the net benefit function of A, this alternative would only reach the level of the net profit objective if the price were equal to 5 mu, that is, 45% less than the initial price (9.5 mu). Only if both treatments attain the objective of net benefit can we proceed to the next step. Now that, with the price reduction, the objective of top priority has been attained, we can move on to the next level of priority, which includes the other criteria. Calculat-

ing the weighted sum of the deviations from the objective value, for the other two criteria with a non-zero relative weight, for each alternative we obtain: $D(A) = (5 \times 0.06) + (5 \times 0.02) = 0.4$ and $D(B) = (5 \times 0.12) + (5 \times 0.1) = 1.1$. So with this method we can conclude that drug A is superior to drug B, being closer to the objectives of equity and adherence, provided that the price of A is reduced by 45% to reach the cost-effectiveness objective.

3.5 GOOD PRACTICES IN MCDA: ADVANTAGES AND RISKS

MCDA applied to healthcare decisions has some advantages which are widely mentioned in the literature, but it also has its disadvantages or limitations, which can be amplified if this method moves away from the conditions of good practice on which the validity and advantages attributed to it are based. One useful function of MCDA is its consideration of complex information when making decisions such as those in which multiple criteria intervene and which may be in conflict; complexity is 'cognitively demanding' and can lead to incongruent decisions. Some of the most important advantages come from the greater transparency of the process through which a decision or recommendation is reached, as well as its congruence or coherence and the fact that it can be easily replicated (Devlin and Sussex, 2011; Garau and Devlin, 2017). Additionally, from the point of view of the political decision-makers, it is important to bear in mind that an explicit focus on criteria and weights or weightings, such as that of MCDA, may also have its disadvantages when compared with less formal and explicit processes of decision-making, since decision-makers can perceive that it contributes to reducing the discretion and degrees of freedom of their decisions (the cost of being "too explicit"; Garau and Devlin, 2017).

In the actual application of MCDA in the wide range of healthcare decisions in which it may be applicable, it is essential to observe the desirable properties of the criteria, attributes, weights and aggregation method, as well as compliance with good practice guidelines such as those of the ISPOR's Task Force (Thokala et al., 2016) or those identified by other authors (Marsh et al., 2016; Thokala and Duenas, 2016) (see Chapter 6 of this book). There are still few MCDA studies in the health field which prove that the criteria have been defined in such a way that they fulfil the requirements of this analysis, such as the avoidance of double counting. Table 3.7 summarises a detailed set of 16 principles of good practice in MCDA in the healthcare sector covering all stages of conducting such an analysis.

Even when the application of MCDA in healthcare decisions respects the conditions of good practice, there are some important costs and risks which may affect the use of this method. In this section, starting from the literature reviewed, three groups of practical risks are identified, leaving the subject of opportunity cost for more detailed comment in the next section.

First, MCDA can be used generically for all decisions with the same purpose (for example, decisions about reimbursement and price), or it can be applied in a specific way, *ad hoc* for each specific case. In this second case, the criteria, weights, value functions and aggregation functions may be different for each drug or technology evaluated, whereas in the first case, they are the same for all evaluations. With a generic MCDA approach, the preferences represented must be those of society, with the added difficulty of appropriately capturing these preferences. A case-by-case approach in which criteria, weights and functions of *ad hoc* value are chosen for each decision can be a serious limitation to the consistency of, and coherence between, the decisions adopted (Garau and Devlin, 2017), with negative implications for the allocative efficiency of the decisions, in addition to reproducing the discretion and inconsist-

ency which MCDA precisely tries to mitigate. The limited practical experience indicates that, the greater the uncertainty about the criteria and the weights, the greater the likelihood that the HTA committees and agencies will be tempted to convert the MCDA into a case-by-case analysis in which each of them is concerned in a different way (Thokala and Duenas, 2012).

TABLE 3.7. PRINCIPLES OF GOOD PRACTICE IN MCDA IN THE HEALTHCARE SECTOR

| STAGE | PRINCIPLE OF GOOD PRACTICE |
|--|---|
| Establishment of the context of the decision | 1. Define the limits of the problem. |
| | 2. Identify the objectives and key actors. |
| | 3. Explore the context and issues with the group. |
| Identification of alternatives | 4. Establish a clearly defined set of requirements for the options. |
| Identification of objectives and criteria | 5. Establish a defined set of operational criteria. |
| | 6. Make sure that the criteria are not redundant. |
| | 7. Make sure that there is no double counting in and between criteria. |
| | 8. Make sure that the criteria are mutually independent of preferences. |
| Scoring of the alternatives | 9. Iterate between options and objectives to create a set of requirements for each one. |
| | 10. Make sure that the evaluators understand the type and meaning of the scale of preferences. |
| | 11. Conduct consistency tests during and after scoring the options. |
| Weighting of criteria | 12. Make sure that the comparison ranges are taken into account when the trade-offs are valued. |
| | 13. Keep it simple. |
| Calculation of weighted scores | 14. Use an algorithm to add evidence and criteria. |
| Analysis of results | 15. Use software to present the results in the form of graphs and tables. |
| Sensitivity and scenario analysis | 16. Explore the robustness of the conclusions by means of sensitivity and scenario analysis. |

Source: author's preparation from Phillips, 2017.

Secondly, there is a risk that the choice of preferences could be influenced by pressure groups, as predicted by the theory of public choice. If the preferences to be taken into account in the MCDA are those of the members of the decision committees, as representatives of social preferences in a generic approach, or those of the parties involved in a case-by-case approach, there is a risk that parties with a more self-serving interest in the decision will be selected instead of those who are really involved. The transaction cost of the participation in a deliberative process of the groups most directly affected, and who have the most to gain or lose as a result of the decision (for example, groups of patients with a certain disease) is much lower than the cost for the rest of society or for the insured parties who bear the opportunity cost of that decision, but with a smaller and more dispersed individual impact (Devlin and Sussex, 2011).

And thirdly, the limitations deriving from uncertainty about the evidence relating to the characteristics of the criteria chosen for each of the alternatives are, at least, the same in a traditional HTA and in an MCDA. In an MCDA, the areas of uncertainty include the structuring

of the problem (choice of the most appropriate MCDA model, criteria, level of detail, etc.), the evidence for each alternative and the variation in preferences (uncertainty about scores, weights of criteria, thresholds, etc.) (Thokala and Duenas, 2012). In an MCDA, the degree of uncertainty is still potentially greater than when the incremental cost-effectiveness ratio is used as the sole decision criterion because of the multiple attributes of the asset or benefit (Garau and Devlin, 2017). For example, is a measurement of uncertainty necessary for each criterion? The influence of uncertainty on healthcare decisions has still been studied very little, and depends, for example, on the parties' attitude to the risk and on the adequate understanding of the uncertainty. In practice, the MCDA must ensure that the parties whose preferences are to be used have an adequate understanding of the uncertainty which exists in the information available for each decision, in addition to using appropriate methods to capture this uncertainty (scenario analysis, multi-attribute utility theory, diffused logic, stochastic analysis of multi-criteria acceptability).

By way of summary, it is of interest for any application of MCDA in HTA, in addition to reviewing the criteria of good practice, to make explicit, to both the decision-makers and the users of the results, the list of key issues to be resolved in order to incorporate MCDA in HTA with guarantees that this will not make things worse (Table 3.8).

3.6. THE 'MULTI-CRITERIA' RATIONALE VERSUS THE RATIONALE OF OPPORTUNITY COST ARE THEY ALTERNATIVES OR COMPLEMENTARY?

The growing literature about MCDA in healthcare decisions points out three reasons why this group of methods may represent an alternative to economic appraisal in HTA procedures (Angelis and Kanavos, 2016): 1) the inclusion of a comprehensive list of value dimensions in an explicit manner, beyond what traditional economic evaluation methods capture; 2) the assignment of quantitative weights across the different evaluation criteria, so that their relative importance is incorporated explicitly, thereby improving the transparency of the preference-elicitation process; 3) the stakeholder participation and the possibility of including them all in the value assessment process, which helps to increase the legitimacy of the process.

The previous potential advantages may be contrasted with the difficulty of the MCDA, if not the failure to consider the opportunity cost. The inclusion of the opportunity cost in an MCDA requires, in theory, all the benefits lost and potentially produced as a result of implementing an intervention to be compared, and this requires having a threshold or maximum limit to understand the opportunity cost. The efficient allocation of limited resources between alternative interventions cannot, under any pretext about the dimensions of value, set aside or neglect consideration of the opportunity cost (taking into account the budgetary constraint), that is, the benefits that are not going to be obtained because of the displacement of resources towards the selected or recommended intervention. Beyond the correct application of the MCDA method (Table 3.7) and the appropriate resolution about the key aspects of an MCDA (Table 3.8), the main problem, and the source of the most important criticisms presented by the use of MCDA in HTA, derives from the difficulty to adequately consider the logic of opportunity cost (the value of the lost benefit when considering the ICER as a decision criterion).

Table 3.8 shows the two conventional options for taking into account the opportunity cost (Garau and Devlin, 2017). The first option is to consider it as one more criterion, together with the other factors. In this case it is necessary, at least, to avoid the risk of overlapping

with other criteria (for example, between cost and cost-effectiveness). As an example, the EVIDEM framework (Goetghebeur et al., 2008) incorporates the compared cost ('economic consequences of the intervention') as one of the five criteria considered as representative of value, together with need, compared results, type of benefit and knowledge about the intervention. For example, Angelis and Kanavos, 2017, propose an Advanced Value Framework (AVF) which includes a criterion of socio-economic impact together with the burden of the disease, the therapeutic impact, the safety profile and the level of innovation. The value of the socio-economic impact criterion is based on three intermediate criteria: public health (risk reduction and prevention), direct incremental costs (medical and non-medical) and incremental indirect costs (absenteeism, presentism, premature abandonment, premature mortality and caregivers). The inclusion of cost, be it the total cost or the incremental cost, as a criterion or dimension of the value of the intervention has been the subject of criticism, since if what is pursued is to obtain a value index, then the criteria should represent attributes of benefit. The AVF tries to evade this criticism by considering not the cost of the intervention, but the incremental cost, as a measurement of the impact on costs instead of the total cost or the cost of acquisition. The treatment of opportunity cost in the majority of currently available MCDA schemes and software only partially deals with the subject of opportunity cost (Garau and Devlin, 2017). When cost is treated as an individualised value criterion, a careful definition of this criterion is necessary to avoid overlapping with other criteria (for example, when both cost and cost-effectiveness are included), and additional information is needed to identify whether it represents good value for money (whether it is efficient). In most applications there is a need to know where the adoption boundary is located (the maximum cost per incremental point) to the extent that the decisions affect a limited budget.

The second option for taking into account the opportunity cost in the framework of an MCDA would be to construct a composite measure of the benefit (net), compared with the cost (net), an incremental cost-value ratio (ICVR) (Angelis and Kanavos, 2016). This option, the incremental approach of MCDA, requires the establishment of an acceptable level or threshold of incremental cost per benefit point which reflects the opportunity cost and, for consistency, that the cost itself has not been included as one of the criteria in the construction of the value index. The ICVR aims to be used as a guide for the allocation of resources in a similar way to the ICER. It is obvious to point out that, among other conditions, value scores must guarantee conditions of comparability so that interventions can be prioritised according to the incremental cost per point of value. This option would be no more than an ICER with an extended measurement of value.

The MCDA can be used in complementary way to the ICER to adjust it, taking into account additional dimensions of value, the 'other' factors that are not measured with the QALYs (complementary approach). In this case, the criticism of the inclusion of cost as a criterion within the MCDA and the explicit absence of the lost benefit as an opportunity cost gains weight. A particular case of the application of this rationale could be, for example, a decision process in two stages. In the first one, the ICER is considered in comparison with the maximum cost threshold per QALY. When the ICER is higher than this threshold, in a second stage the appraisal is complemented by obtaining a value index through the MCDA. Possibly the outranking approach is better suited to a two-stage process of this type than the value-measurement approach.

With a different perspective, although very rare, MCDA can be used as a single method (the pure approach) to estimate the value of an intervention without using the ICER (or, only includ-

TABLE 3.8. KEY ISSUES FOR THE USE OF MCDA IN HTA

| | OPTIONS | ISSUES TO CONSIDER |
|---|--|---|
| Which criteria and weights? | <ol style="list-style-type: none"> 1. Set in advance; similarly in all decisions. 2. Selected case-by-case and varying throughout technologies or disease areas. | <ol style="list-style-type: none"> 1. It allows the use of the same metric to measure losses and added benefits; it considers consistently all the criteria. 2. Flexible method. Even so, a systematic consideration of all the criteria and the predictability of decision-making can be difficult. |
| Which criteria? | <ol style="list-style-type: none"> 1. The current criteria of the HTA agencies. 2. Those of the members of the HTA committee, representing those of the healthcare system. 3. Reflection of the opinions of the general public. | <ol style="list-style-type: none"> 1. It is assumed that there is an obligation to consider these criteria. 2. The involvement of holders of the budget for healthcare can promote an alignment of objectives through various decision-makers in the healthcare system. 3. They reflect the opinions of the users of the healthcare system/taxpayers. |
| Which are the preferences for weighting the criteria? | <ol style="list-style-type: none"> 1. Any of those involved, defined by the decision-maker. 2. Members of an HTA committee. 3. Members of the general public. | <ol style="list-style-type: none"> 1. In congruence with the current extra-welfarist principles of the HTA. Also, the variations of those involved throughout the diseases require flexible weights. 2. A paradigmatic method that can avoid carrying out large studies based on preferences. 3. Consistent with the method used to assess the quality of life in QALYs. |
| How are opportunity costs incorporated? | <ol style="list-style-type: none"> 1. Different criteria for costs. 2. An aggregate measure of benefits (net) to be set against costs (net). | <ol style="list-style-type: none"> 1. Risk of overlapping with other criteria (for example, cost and cost-effectiveness). 2. Requires the fixing of an 'acceptable cost per increase in the benefit/score scale'. |
| How can uncertainty be dealt with? | <ol style="list-style-type: none"> 1. A separate, different criterion for uncertainty. 2. Sensitivity analysis techniques. | <ol style="list-style-type: none"> 1. Measuring and assessing this criterion presents difficulties. Different criteria can be associated with different types and degrees of uncertainty. 2. One ensures that the fragility of the premises about key evidential issues is taken into account. Similarly, the question of how the results of the sensitivity analysis should affect decision-making is left unanswered. |

QALY: Quality-Adjusted Life-Years.

Source: Garau and Devlin, 2017.

ing variables related to the ICER as criteria of the MCDA). The pure approach could better reflect the opportunity cost of the resources used, provided a threshold is established, although it is very likely that the limitations derived from the specific context and the preferences of the stakeholders included, represent still more serious limitations for the points of value than for the QALYs and the ICER (see Tables 3.7 and 3.8).

Acknowledgements. The author thanks Carlos Campillo and Juan Oliva for their comments on a previous version of this chapter.

REFERENCES

Adunlin G, Diaby V, Xiao H. Application of multicriteria decision analysis in health care: a systematic review and bibliometric analysis. *Health Expectations*. 2014; 18(6): 1894-1905. doi:10.1111/hex.12287.

Angelis A, Kanavos P. Value-Based Assessment of New Medical Technologies: Towards a Robust Methodological Framework for the Application of Multiple Criteria Decision Analysis in the Context of Health Technology Assessment. *Pharmacoeconomics*. 2016; 34(5): 435-446. doi:10.1007/s40273-015-0370-z.

Angelis A, Kanavos P. Multiple Criteria Decision Analysis (MCDA) for evaluating new medicines in Health Technology Assessment and beyond: The Advance Value Framework. *Soc Sci Med*. 2017; 188: 137-156. doi:10.1016/j.socscimed.2017.06.024.

Angelis A, Lange A, Kanavos P. Using health technology assessment to assess the value of new medicines: results of a systematic review and expert consultation across eight European countries. *Eur J Health Econ*. 2017; 19(1): 123-152. doi:10.1007/s10198-017-0871-0.

Auraaen A et al. How OECD health systems define the range of goods and services to be financed collectively. Paris: OECD Health Working Papers 90. OECD Publishing; 2016.

Belton V, Stewart TJ. Multiple criteria decision analysis: an integrated approach. Boston: Kluwer Academic; 2002.

Castro H, Tringali M, Cleemput I, Devriese S, Leoni O, Lettieri E. Advancing MCDA and HTA into coverage decision-making. On: Marsh K, Goetghebeur M, Thokala P & Baltussen R, editors. *Multi-Criteria Decision Analysis to Support Healthcare Decisions*. Cham: Springer International Publishing; 2017.

Cerri KH, Knapp M, Fernandez J. Decision making by NICE: examining the influences of evidence, process and context. *Health Econ Policy Law*. 2013; 9(02): 119-141. doi:10.1017/s1744133113000030.

Communities and Local Government. Multi-criteria analysis: a manual. London: Department for Communities and Local Government; 2009.

Culyer A. The normative economics of health care finance and provision. On: McGuire A, Fenn P, Mayhew K, editors. *Providing health care: the economics of alternative systems of finance and delivery*. Oxford: Oxford University Press; 1991.

Culyer A. Ethics, priorities and cancer. *J Cancer Policy*. 2017; 11: 6-11. doi:10.1016/j.jcpcp.2016.09.007.

Devlin NJ, Sussex J. Incorporating multiple criteria in HTA. London: Office of Health Economics; 2011.

Diaby V, Dias L. Beyond value function methods in MCDA for health care. On: Marsh K, Goetghebeur M, Thokala P & Baltussen R, editors. *Multi-Criteria Decision Analysis to Support Healthcare Decisions*. Cham: Springer International Publishing; 2017.

Drummond MF, Sculpher MJ, Claxton K, Stoddart GL, Torrance GW. *Methods for the economic evaluation of health care programmes*. Oxford: Oxford University Press; 2015.

Drummond MF, Wilson DA, Kanavos P, Ubel P, Rovira J. Assessing the economic challenges posed by orphan drugs. *Int J Technol Assess Health Care*. 2007; 23(01), 36-42. doi:10.1017/s0266462307051550.

Drummond MF, Evans B, Leloirier J, Karakiewicz P, Martin D, Tugwell P, MacLeod S. Evidence and values, requirements for public reimbursement of drugs for rare diseases – A case study in oncology. *Can J Clin Pharmacol*. 2009; 16(2): e282-284.

Garau M, Devlin NJ. Using MCDA as a decision aid in health technology appraisal for coverage decisions: opportunities, challenges and unresolved questions. On: Marsh, K., Goetghebeur, M., Thokala, P., & Baltussen, R, editors. *Multi-Criteria Decision Analysis to Support Healthcare Decisions*. Cham: Springer International Publishing; 2017.

Goetghebeur M, Wagner M, Khoury H, Levitt R, Erickson L, Rindress D. Evidence and Value: Impact on Decision Making - The EVIDEM framework and potential applications. *BMC Health Serv Res.* 2008; 11(3). doi:10.1016/s1098-3015(10)70584-2.

Golan O, Hansen P, Kaplan G, Tal O. Health technology prioritization: Which criteria for prioritizing new technologies and what are their relative weights? *Health Policy.* 2011; 102(2-3): 126-135. doi:10.1016/j.healthpol.2010.10.012.

Harris A, Li JJ, Yong K. What Can We Expect from Value-Based Funding of Medicines? A Retrospective Study. *PharmacoEconomics.* 2015; 34(4): 393-402. doi:10.1007/s40273-015-0354-z.

Hounton S, Newlands D. Applying the net-benefit framework for assessing cost-effectiveness of interventions towards universal health coverage. *Cost Eff Resour Alloc.* 2012; 10(1): 8. doi:10.1186/1478-7547-10-8.

Laupacis A. "Evidence and values: requirements for public reimbursement of drugs for rare diseases – A case study in oncology" Commentary. *Can J Clin Pharmacol.* 2009; 16(2): e282-e284.

McCabe C, Tsuchiya A, Claxton K, Raftery J. Assessing the economic challenges posed by orphan drugs: A comment on Drummond et al. *Int J Technol Assess Health Care.* 2007; 23(03): 397-401. doi:10.1017/s0266462307071024.

Marsh K, Lanitis T, Neasham D, Orfanos P, Caro J. Assessing the Value of Healthcare Interventions Using Multi-Criteria Decision Analysis: A Review of the Literature. *PharmacoEconomics.* 2014; 32(4): 345-365. doi:10.1007/s40273-014-0135-0.

Marsh K, Ijzerman M, Thokala P, Baltussen R, Boysen M, Kaló Z, et al. Multiple Criteria Decision Analysis for Health Care Decision Making—Emerging Good Practices: Report 2 of the ISPOR MCDA Emerging Good Practices Task Force. *Value Health.* 2016; 19(2): 125-137. doi:10.1016/j.jval.2015.12.016.

Marsh K, Thokala P, Mühlbacher A, Lanitis T. Incorporating preferences and priorities into MCDA: selecting an appropriate scoring and weighting technique. On: Marsh K, Goetghebeur M, Thokala P, Baltussen R, editors. *Multi-Criteria Decision Analysis to Support Healthcare Decisions.* Cham: Springer International Publishing; 2017.

Mühlbacher AC, Kaczynski A. Making Good Decisions in Healthcare with Multi-Criteria Decision Analysis: The Use, Current Research and Future Development of MCDA. *Appl Health Econ Health Policy.* 2015; 14(1): 29-40. doi:10.1007/s40258-015-0203-4.

Neumann PJ, Cohen JT. Measuring the Value of Prescription Drugs. *N Engl J Med.* 2015; 373(27): 2595-2597. doi:10.1056/nejmp1512009.

Nord E. Beyond QALYs: Multi-criteria based estimation of maximum willingness to pay for health technologies. *Eur J Health Econ.* 2017. doi:10.1007/s10198-017-0882-x.

Paulden M, Stafinski T, Menon D, McCabe C. Value-Based Reimbursement Decisions for Orphan Drugs: A Scoping Review and Decision Framework. *PharmacoEconomics.* 2014; 33(3): 255-269. doi:10.1007/s40273-014-0235-x.

Phillips LD. Best Practice for MCDA in Healthcare. On: Marsh K, Goetghebeur M, Thokala P, Baltussen R, editors. *Multi-Criteria Decision Analysis to Support Healthcare Decisions.* Springer, Cham; 2017.

Regier DA, Peacock S. Theoretical foundations of MCDA. On: Marsh K, Goetghebeur M, Thokala P & Baltussen R, editors. *Multi-Criteria Decision Analysis to Support Healthcare Decisions.* Cham: Springer International Publishing; 2017.

Sandier S, Paris V, Polton D, Thomson S, Mossialos E. Health care systems in transition. Copenhagen, Denmark: European Observatory on Health Systems and Policies; 2004.

Stolk EA, Poley MJ. Criteria for determining a basic health services package. *Eur J Health Econ.* 2005; 6(1): 2-7. doi:10.1007/s10198-004-0271-0.

Thokala P, Devlin N, Marsh K, Baltussen R, Boysen M, Kalo Z, et al. Multiple Criteria Decision Analysis for Health Care Decision Making—An Introduction: Report 1 of the ISPOR MCDA Emerging Good Practices Task Force. *Value Health*. 2016; 19(1): 1-13. doi:10.1016/j.jval.2015.12.003.

Thokala P, Duenas A. Multiple Criteria Decision Analysis for Health Technology Assessment. *Value Health*. 2012; 15(8): 1172-1181. doi:10.1016/j.jval.2012.06.015.

Tromp N, Mikkelsen E, Viergever R, Baltussen R. Embedding MCDA in priority-setting processes: experiences from low- and middle-income countries. In: Marsh K, Goetghebeur M, Thokala P & Baltussen R, editors. *Multi-Criteria Decision Analysis to Support Healthcare Decisions*. Cham: Springer International Publishing; 2017.

Wahlster P, Goetghebeur M, Kriza C, Niederländer C, Kolominsky-Rabas P. Balancing costs and benefits at different stages of medical innovation: a systematic review of Multi-criteria decision analysis (MCDA). *BMC Health Serv Res*. 2015; 15(1). doi:10.1186/s12913-015-0930-0.

Walton SM, Basu A, Mullahy J, Hong S, Schumock GT. Measuring the Value of Pharmaceuticals in the US Health System. *PharmacoEconomics*. 2016; 35(1): 1-4. doi:10.1007/s40273-016-0463-3.

Zaragozá F, Cuéllar S. Valoración de la innovación terapéutica: importancia y medida. En: del Llano JE, editor. *Innovación y regulación en biomedicina: obligados a entenderse*. Madrid: Fundación Gaspar de Casal; 2017.



SELECTION OF CRITERIA IN MULTI-CRITERIA DECISION ANALYSIS

Jaime Espín Balbino

CHAPTER

4

4.1. INTRODUCTION

One of the usual criticisms of decisions based on studies of economic appraisal (regardless of the existence or not of a previously determined threshold) is the lack of consideration of criteria other than those of mere efficiency to make those decisions. Multi-Criteria Decision Analyses (MCDA) have, as one of their objectives, that of trying to cover this gap by selecting a broad set of criteria (including economic ones, but also equity, public healthcare, etc.) and weighting each according to its importance; that is, MCDA will address complex problems through multiple dimensions, using qualitative and quantitative approaches, and making explicit the multiple factors that are taken into account (which can sometimes be contradictory) (Castro et al., 2016).

MCDA considers three important steps (Devlin and Sussex, 2011): the selection of the criteria, the weighting or assignment of weights to them and the use of MCDA in decision-making. In this chapter we shall deal with the first of these aspects.

4.2. SELECTION OF CRITERIA IN MCDA

In general terms, the criteria to be selected for the performance of an MCDA must have at least three characteristics: they must be 1) broad, to collect all the areas related to the health problem about which a decision is going to be made; 2) relevant, in such a way that they take into account those criteria that are crucial to the decision; and 3) explicit, as a key element of transparency and giving account in the decision-making process. As it can be seen in table 4.1, international experience shows us that in making decisions about the allocation of resources, important criteria such as the budget impact, equity or the burden of disease have sometimes not been taken into account; or, if they have been taken into account, this has not been done explicitly.

TABLE 4.1. CRITERIA USED TO ASSESS DRUGS FOR PRICE AND REIMBURSEMENT IN THE EUROPEAN CONTEXT

| CRITERIA | AT | BE | CH | DE | ES | FI | FR | NL | NO | SE | UK |
|--|----|----|----|----|----|----|----|----|----|----|----|
| Therapeutic benefit | X | X | X | X | X | X | X | X | X | X | X |
| Patient benefit | X | X | X | X | | X | X | X | X | X | X |
| Cost-effectiveness | X | X | | | X | X | | X | X | X | X |
| Budget impact | | X | | | X | X | X | X | X | | X |
| Innovation | X | X | | | X | | X | X | | | X |
| Availability of therapeutic alternatives | X | | | | X | | | X | | X | X |
| Equity considerations | | | | | | | | | X | X | X |
| Public health impact | | | | | | | X | | | | |
| R&D | | | | | | X | | | | | |

AT: Austria. BE: Belgium. CH: Switzerland. DE: Germany. ES: Spain. FI: Finland.
FR: France. NL: The Netherlands. NO: Norway. SE: Sweden. UK: United Kingdom.

Source: author's preparation from Kanavos et al., 2011, updated with Spain for this book.

Continuing with the previous example, we can draw two major conclusions which allow us to advance in the relevance of conducting an MCDA: the first is that we do not know explicitly what made a country use some criteria and not others; secondly, once the criteria have been determined, we do not know which ones are more important or whether they all have the same weight when it comes to setting the price and/or reimbursement of a new drug.

In some cases, the criteria can be pre-defined and grouped in a broad and general way (into what can be called domains or dimensions), as can be seen later in the Belgian experience (where they are called clusters), where the population is consulted about the specific criteria that they would take into account when making a decision. This situation would be perfectly applicable to the Spanish context within the framework of the Law on Guarantees and Rational Use of Medicines. Article 92 of *Real Decreto Legislativo* 1/2015 establishes the criteriaⁱ for the public financing of medicines and healthcare products (BOE, 2015). However, we do not know what methods are used for their selection, and still less the weighting criteria used to fix prices and reimbursement for medicines in Spain, as the resolutions are not made public.

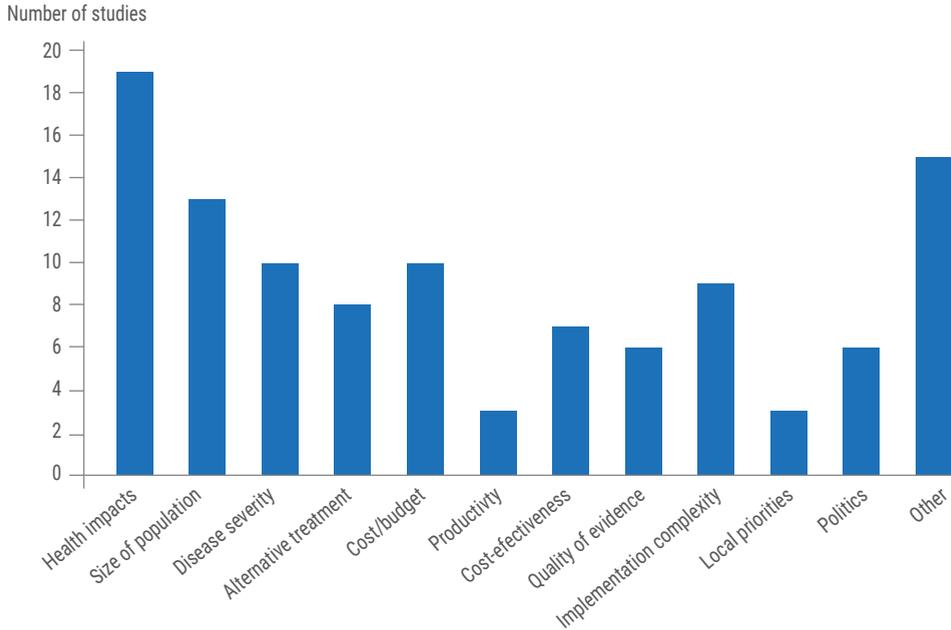
In those countries where there are no pre-defined criteria for decision-making, their selection must be one of the first steps in the application of MCDA, either using quantitative or qualitative methods (focus groups, public consultations, qualitative research, declared preferences, etc.) (Devlin and Sussex, 2011). A good start for the selection of criteria would be the use of the priority-setting frameworks which have been designed in many international contexts.

A recent review of the literature (Marsh et al., 2014) shows us the predominant criteria included in the studies conducted with MCDA (Figure 4.1). A first analysis of the data allows us to draw some conclusions: first, the economic aspects (either budgetary impact or cost-effectiveness studies) play a minor role as criteria and do not even include 50% of the MCDA (the number of studies in this review is 23); second, the clinical aspects (impact on health, severity of the disease, etc.) are the most common criteria; third, criteria that are usually taken into account in a non-explicit way, such as political aspects or local priorities, are explicitly included in some cases.

In 2016, the International Society for Pharmacoeconomics and Outcomes Research (ISPOR) published a set of good practice recommendations for the implementation of MCDA (Thokala et al., 2016; Marsh et al., 2016). Within the framework of these recommendations, the identification and selection of criteria constitute the second step (Table 4.2), (see Chapter 6 for more details).

ⁱ The criteria are:

- a) Severity, duration and consequences of the different pathologies for which they are indicated.
- b) Specific needs of certain groups.
- c) Social value of the drug and its incremental clinical benefit, taking into account its cost-effectiveness ratio.
- d) Rationalization of public spending for pharmaceutical benefit and budget impact in the healthcare system.
- e) Existence of drugs or other or other therapeutic alternatives for the same conditions at a lower price or lower cost of treatment.
- f) Innovation level of the drug.

FIGURE 4.1. CRITERIA INCLUDED IN MCDA FOR INVESTMENT DECISIONS

Source: author's preparation from Marsh et al., 2014.

TABLE 4.2. STEPS IN THE PROCESS OF MEASURING VALUE IN AN MCDA

| STEP | RECOMMENDATION |
|---------------------------------------|---|
| Defining the decision problem | Identify objectives, type of decision, alternatives, stakeholders, and output required |
| Selecting and structuring criteria | Identify criteria relevant for evaluating alternatives |
| Measuring performance | Gather data about the alternatives' performance on the criteria and summarize this in a "performance matrix" |
| Scoring alternatives | Elicit stakeholders' preferences for changes within criteria |
| Weighting criteria | Elicit stakeholders' preferences between criteria |
| Calculating aggregate scores | Use the alternatives' scores on the criteria and the weights for the criteria to get "total value" by which the alternatives are ranked |
| Dealing with uncertainty | Perform uncertainty analysis to understand the level of robustness of the MCDA results |
| Reporting and examination of findings | Interpret the MCDA outputs, including uncertainty analysis, to support decision making |

Source: Marsh et al., 2016.

As the ISPOR guide points out, the selection of criteria is the next step after the problem has been identified. Sometimes it may happen that the criteria are exclusively clinical, like some examples that are indicated in the guide itself. The criteria can be selected by various means, such as the review of previous decisions, or the holding of focus groups or seminars, as already noted above.

The first phase of criteria selection usually results in a list with numerous criteria that have to be reduced later, trying to ensure that the selected set fulfils the following requirements (Marsh et al., 2016):

- **Completeness:** the criteria should capture all factors relevant for the decision.
- **Non redundancy:** we must avoid those criteria that are unnecessary or unimportant.
- **Non overlap:** we should avoid overlaps between criteria, because this will prevent some elements from being double counted, thus overvaluing some dimensions. An example that is indicated as a possible overlap is the inclusion, on the one hand, of cost and effectiveness data as separate criteria and, on the other hand, of cost-effectiveness as an additional criterion.
- **Preference independence,** in such a way that we avoid the presence of interactions between criteria.

The same ISPOR guide points out that, once the criteria have been identified, they must be defined so that they contain the following characteristics: to be precise/unambiguous, comprehensive, direct, operational and understandable.

For the selection of criteria, the use of value trees is recommended, in such a way that they are organised in a hierarchical manner.

There is no norm that indicates the number of criteria that have to be selected, but ISPOR notes that in a review of the literature (Marsh et al., 2014) an average use of 8.2 criteria was found, the number of criteria varying from 3 to 19. However, the recommendation is to have a limited number of criteria, especially when using a discrete choice experiment, where it is recommended that between 4 and 5 be used. One of the best known and most used MCDA frameworks, EVIDEM, takes into account 13 quantitative criteria (although in some specific cases some of the criteria have been deleted).

As Marsh et al., 2014, point out, in their review of the literature about MCDA, “the documents include little discussion about the process of defining the criteria.” Additionally, it states that only one study explains that the criteria were defined in compliance with the requirements of MCDA, such as the avoidance of double counting. Finally, Marsh points out that, although having the opinion of experts is important for the definition and selection of criteria, it is undoubtedly necessary to create guidelines for this purpose. In the article published by Angelis and Kanavos in 2016, it is also pointed out that there are not enough methodological guides to designing, conducting and implementing MCDA as part of the appraisal of health technologies, including how to select criteria appropriately.

In the same article, the authors indicate what the different approaches could be for the selection of objectives and criteria (Table 4.3). In summary, there are three main approaches.

The first is 'value-focused thinking', in which the selection of criteria is always made a priori, even before knowing the different options; the second is 'alternative-focused thinking', in which the criteria arise from the characteristics of the different alternatives; and finally, the third method is 'value-alternative hybrid thinking', perhaps the most widely used in the field of health technology assessment, where decisions have a generic set of objectives and criteria which are adapted to each concrete decision. The example which the authors show for this third approach is that of adverse effects: they are traditionally included as criteria to be evaluated in the different alternatives, but they can be excluded if they are not necessary in a specific case.

TABLE 4.3. APPROACHES FOR SELECTING OBJECTIVES AND CRITERIA

| Approach | 'Value-focused thinking' | 'Alternative-focused thinking' | 'Value-alternative hybrid thinking' |
|----------------------|---|--|---|
| Description | Objectives and criteria are selected before the identification or assessment of the alternative options | The options are previously compared so that objectives and criteria can emerge based on their attributes | A set of criteria and general objectives are first created, and are then adapted to the particular decision problem |
| Value tree formation | Top-down approach | Bottom-up approach | Top-down followed by bottom-up |

Source: author's preparation from Angelis and Kanavos, 2016.

4.3. EXPERIENCES IN THE SELECTION OF CRITERIA

EVIDEM framework

One of the best known and most used MCDA frameworks is EVIDEM (Evidence and Value: Impact on Decision-Making). The initial framework was designed in 2006 by EVIDEM Collaboration, based on the common objective of the healthcare systems (the imperative of alleviating/preventing the suffering of patients, prioritising those who are worse, ensuring sustainability) and on the ethical imperatives underlying the said objective.

The EVIDEM has a central model of quantifiable criteria, whose scales can be defined a priori, and complements it with a contextual tool of qualitative criteria, which serve to adapt the framework more closely to the context of each decision. In turn, most criteria include sub-criteria that can be added to the final model.

The criteria selected by EVIDEM are generic and universally operable (Goetghebeur et al., 2012). To select the criteria, an extensive analysis of the literature and current procedures for decision-making has to be carried out (Goetghebeur et al., 2008). They are defined in accordance with the methodological principles of MCDA: non-redundancy, independence, operability and completeness.

In its latest update (version 4.0), of 2017, the EVIDEM framework groups the 13 quantitative criteria into five domains (need for intervention, results of the intervention, type of benefit of the intervention, economic consequences and knowledge of the intervention), while the 7 contextual criteria are divided into two groups or domains (normative or feasibility contextual criteria) (Evidem Collaboration, 2017) (Table 4.4).

TABLE 4.4. CRITERIA INCLUDED IN THE EVIDEM FRAMEWORK

| CORE MODEL | CONTEXTUAL TOOL |
|--|---|
| 1 Disease severity | 1 Mandate and scope of the healthcare system |
| 2 Size of affected population | 2 Population priorities and access |
| 3 Unmet needs | 3 Common goal and specific interests |
| 4 Comparative effectiveness/efficacy | 4 Environmental impact |
| 5 Comparative safety/tolerability | 5 Opportunity costs & affordability |
| 6 Comparative patient-reported outcomes (PRO) | 6 System capacity & appropriate use of the intervention |
| 7 Type of preventive benefit | 7 Political, historical and cultural context |
| 8 Type of therapeutic benefit | |
| 9 Comparative cost of the intervention | |
| 10 Comparative other medical costs | |
| 11 Comparative non-medical costs | |
| 12 Quality of evidence | |
| 13 Expert consensus / clinical practice guidelines | |

Source: EVIDEM Collaboration, 2017.

Note: EVIDEM recognizes that the cost-effectiveness ratio is a composite measure, already considered in other criteria, which therefore does not comply with the requirement of non-redundancy. It could be included in the framework, since many current decision-making processes are based on this measure, but should ultimately be replaced by the criterion with which it overlaps.

Colombia

In 2012, the *Comisión de Regulación en Salud (CRES)* of Colombia (a government agency attached to the Ministry of Health, and which was responsible for the *Plan Obligatorio de Salud* [Mandatory Health Plan] – POS) used the EVIDEM methodology to update the services of the POS, and after performing two nominal group sessions, 15 criteria were selected, of which 13 were from EVIDEM and two from the country context. With the support of the health technology assessment agency of Colombia (IETS), literature reviews and Health Technology Assessment reports were produced, followed by panel meetings with decision-makers to update the POS.

Of the 15 selected criteria, the first three were the severity of the disease, the size of the population and the efficacy of the treatment (Castro et al., 2016). Participants in the sessions considered that the most appropriate approach should contain two steps for decision-making: on the one hand, performing the analysis with the MCDA format, and, in a complementary way, performing a budget impact analysis.

Belgium

In 2010, a document was published in Belgium (le Polain et al., 2010), which proposed a reference framework to improve accountability in the context of reimbursement of public healthcare benefits, focusing on three criteria: feasibility, transparency and consistency. Within this framework, the elements that the decision had to consider (social, medical and therapeutic needs; willingness to pay) were raised, and with them, the possible criteria (Table 4.5).

TABLE 4.5. KEY QUESTIONS AND POSSIBLE CRITERIA FOR A DRUG REIMBURSEMENT APPRAISAL PROCESS IN BELGIUM, IN ACCORDANCE WITH THE MCDA FRAMEWORK

| DECISION | QUESTION | POSSIBLE CRITERIA |
|---|---|---|
| Medical, therapeutic and societal need | Does the product target a medical, therapeutic and societal need? | Medical need: Life-threatening/non-life threatening condition; severe/moderate symptoms; poor initial state of health; reference health level Therapeutic need: Effective alternative treatments available/not available Societal need: High/low prevalence; health inequalities; public spending related to the disease |
| Preparedness to pay out of public resources for a treatment | Are we, as a society, in principle, prepared to pay for a treatment that will improve this indication out of public resources? | Own responsibility; condition related to lifestyle |
| Preparedness to pay out of public resources for the treatment under consideration | Are we, as a society, prepared to pay for this particular treatment, given that we in general would be prepared to pay for a treatment for this indication? | Safety and efficacy of the treatment compared with alternative treatments; Curative, symptomatic, preventive; Therapeutic value; Significance of health gains |
| Preparedness to pay more | Are we, as a society, prepared to pay more for this treatment than for the best available alternative treatment? | Added therapeutic value; Potentially induced savings elsewhere in the health care sector; Quality and uncertainty of the evidence; Acceptability of co-payments; Rarity of the disease |
| Willingness to pay (price and reimbursement basis) | How much are we willing to pay out of public resources for this particular treatment? | Added therapeutic value; Budget impact/ability to pay; Cost-effectiveness ratio; Medical, therapeutic and societal need; Quality and uncertainty of the evidence; Limits to cost sharing. |

Source: author's preparation from le Polain et al., 2010.

In 2012, the public were consulted about what they thought should be the most appropriate criteria for setting the price and reimbursement of drugs in the previous frame of reference, in order to give answers to the questions raised (the results were published in Cleemput et al., 2014). The criteria finally selected as the most relevant, as well as their weights, can be seen in table 4.6.

TABLE 4.6. SELECTED CRITERIA AND THEIR WEIGHTS

| THERAPEUTIC NEED | WEIGHT |
|--|--------|
| Impact of the condition on quality of life, given current treatments available | 0.43 |
| Impact of the condition on life expectancy, given current treatments available | 0.14 |
| Inconvenience of current treatment | 0.43 |
| SOCIETAL NEED | WEIGHT |
| Condition-related public expenditures per patient | 0.65 |
| Frequency of the condition | 0.35 |

| ADDED VALUE | WEIGHT |
|--------------------------------|--------|
| Change in quality of life | 0.37 |
| Change in prevalence | 0.36 |
| Change in life expectancy | 0.14 |
| Impact on public expenditures | 0.07 |
| Impact on treatment discomfort | 0.06 |

Source: author's preparation from Cleemput et al., 2016.

The document is accompanied by a definition of what is meant by each of the criteria. The document also highlights the fact that, although the criterion of convenience was not included as one of the criteria by the participants, the Belgian law that sets the criteria for the reimbursement of medicines places it as one of those to be taken into account.

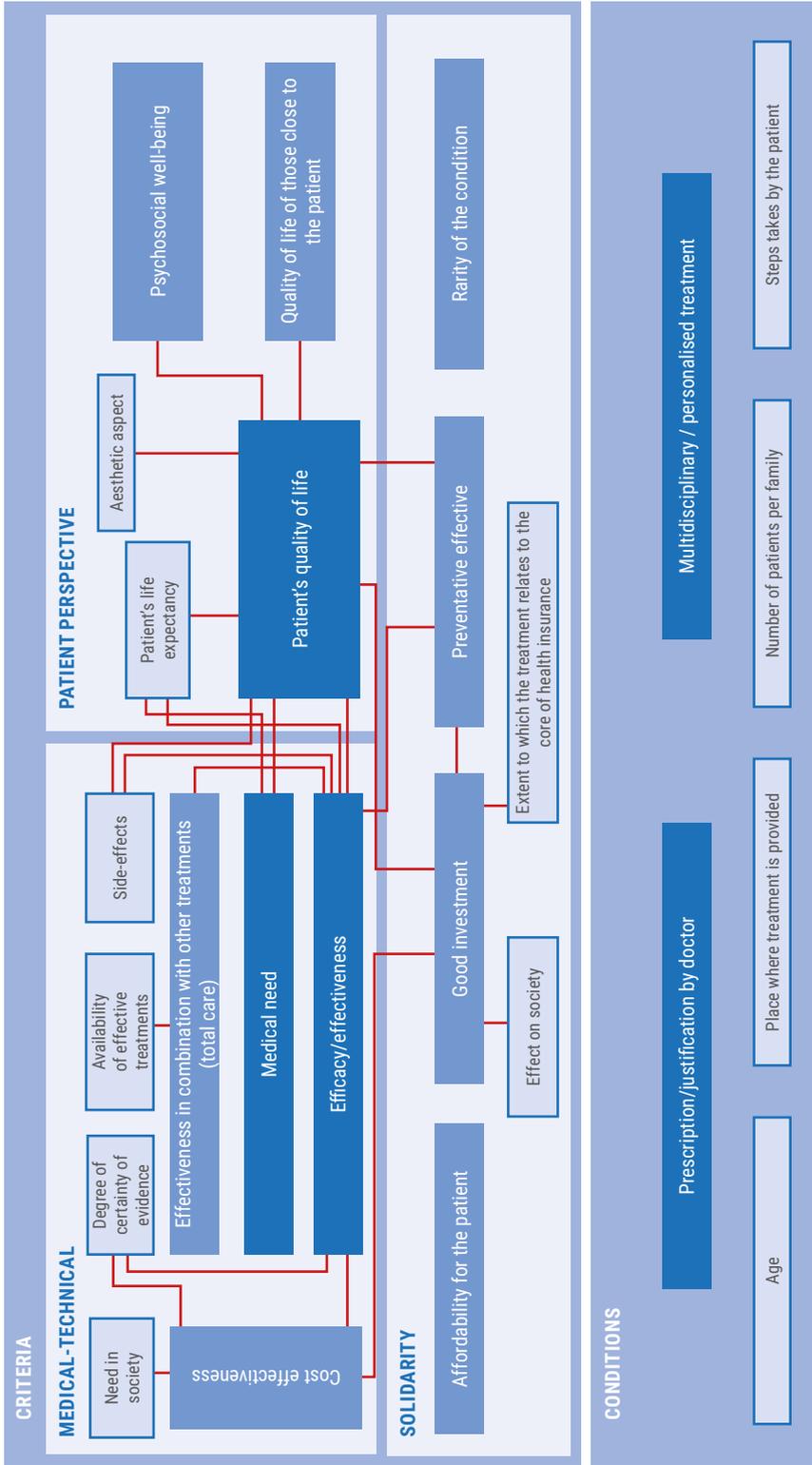
Finally, the document mentions that, in 2015, the King Baudouin Foundation consulted the public about what they considered were the key criteria for the reimbursement of public healthcare benefits. The result was 19 criteria and 6 conditions for reimbursement, as can be seen in figure 4.2. The criteria are grouped into three large clusters: technical/medical, the patient's perspective, and solidarity. In the medical section, not only medical needs are taken into account, but also effectiveness, side effects, efficiency, etc. In the patient section, criteria such as quality of life expectancy, psycho-social well-being, etc., are mentioned. Finally, in the section of solidarity some of the criteria taken into account are the rarity of the condition, the affordability for the patient, etc.

4.4. CONCLUSIONS

The choice of criteria for MCDA is a key aspect which ensures that the final result of the analysis can adequately capture all the necessary dimensions. Whether through a review of the literature or through pre-established MCDA frameworks such as EVIDEM, there is a high degree of agreement about the 10-15 keys that are repeated in different international experiences (see, for example, Tanios et al., 2013). So the great differences in results will not be determined as much by the selection of the criteria, as by the weighting that is given to these criteria.

Nevertheless, despite the coincidence mentioned in relation to the key criteria to be used in MCDA, the participation of citizens and healthcare decision-makers is crucial to ensuring that the final choice is endowed with legitimacy in the context of each country.

FIGURE 4.2. CRITERIA AND CONDITIONS FOR REIMBURSEMENT IDENTIFIED BY THE CITIZENS PANEL



Source: King Baudouin Foundation, 2015.

REFERENCES

Angelis A, Kanavos P. Value-Based Assessment of New Medical Technologies: Towards a Robust Methodological Framework for the Application of Multiple Criteria Decision Analysis in the Context of Health Technology Assessment. *Pharmacoeconomics*. 2016; 34(5): 435-46. doi: 10.1007/s40273-015-0370-z.

Castro H, Tringali M, Cleemput I, Devriese S, Leoni O, Lettieri E. Advancing MCDA and HTA into Coverage Decision-Making. In: Marsh K, Goetghebeur M, Thokala P, Baltussen R, editors. *Multi-Criteria Decision Analysis to Support Healthcare Decisions*. Cham: Springer International Publishing, 2017: 119-146.

Castro HE, Goetghebeur M, Moreno-Mattar O. Testing Multi-Criteria Decision Analysis for More Transparent Resource-Allocation Decision Making in Colombia. *Int J Technol Assess Health Care*. 2016; 32(4): 307-314. doi: 10.1017/S0266462316000350.

Cleemput I, Devriese S, Christiaens W, Kohn L. Multi-criteria decision analysis for the appraisal of medical needs: a pilot study. *Health Services Research (HSR)*. Brussels: Belgian Health Care Knowledge Centre (KCE); 2016. KCE Reports 272. D/2016/10.273/68.

Devlin N, Sussex J. *Incorporating multiple criteria in HTA. Methods and processes*. London: Office of Health Economics; 2011.

EVIDEM Collaboration. *Concept & definitions (v4.0)*. 10th Edition. 2017.

Goetghebeur MM, Wagner M, Khoury H, Levitt RJ, Erickson LJ, Rindress D. Evidence and Value: Impact on Decision Making – the EVIDEM framework and potential applications. *BMC Health Serv Res*. 2008; 8(270).

Goetghebeur MM, Wagner M, Khoury H, Levitt RJ, Erickson LJ, Rindress D. Bridging health technology assessment (HTA) and efficient health care decision making with multicriteria decision analysis (MCDA): applying the EVIDEM framework to medicines appraisal. *Med Decis Making*. 2012; 32(2): 376-88. doi: 10.1177/0272989X11416870.

Kanavos P, Nicod E, Espin J, van den Aardweg S. Short- and Long-Term Effects of Value-Based Pricing vs. External Price Referencing, (European Medicines Information Network –EMINET- 2011). DG Enterprise and Industry. European Commission; 2011.

King Baudouin Foundation. *Reimbursement in Health Care: an Agenda for Change*. Final Report 2012-2015. November 2015. Available from: <https://www.kbs-frb.be/en/Virtual-Library/2015/20151105AVCTERUG-BETALING>.

le Polain M, Franken M, Koopmanschap M, Cleemput I. Drug reimbursement systems: international comparison and policy recommendations. *Health Services Research (HSR)*. Brussels: Belgian Health Care Knowledge Centre (KCE). 2010. KCE Reports 147C. D/2010/10.273/90. Available from: https://kce.fgov.be/sites/default/files/page_documents/KCE_147C_Drug_reimbursement_systems_4.pdf.

Marsh K, Lanitis T, Neasham D, Orfanos P, Caro J. Assessing the Value of Healthcare Interventions Using Multi-Criteria Decision Analysis: A Review of the Literature. *Pharmacoeconomics*. 2014; 32:345–365.

Marsh K, et al. Multiple Criteria Decision Analysis for Health Care Decision Making—An Introduction: Report 2 of the ISPOR MCDA Emerging Good Practices Task Force. *Value Health*. 2016; 19: 125-137.

RD Legislativo 1/2015 de 24 de julio, por el que se aprueba el texto refundido de la Ley de Garantías y Uso racional de los medicamentos y productos sanitarios. BOE número 177.

Tanios N, Wagner M, Tony M, Baltussen R, van Til J, Rindress D, Kind P, Goetghebeur MM; International Task Force on Decision Criteria. Which criteria are considered in healthcare decisions? Insights from an international survey of policy and clinical decision makers. *Int J Technol Assess Health Care*. 2013; 29(4): 456-65. doi: 10.1017/S0266462313000573.

Thokala P, Devlin N, Marsh K, Baltussen R, Boysen M, Kalo Z, Longrenn T, Mussen F, Peacock S, Watkins J, Ijzerman M. Multiple Criteria Decision Analysis for Health Care Decision Making—An Introduction: Report 1 of the ISPOR MCDA Emerging Good Practices Task Force. *Value Health* 19 (2016) 1-13.



TECHNIQUES OF WEIGHTING, SCORING, MODELLING AND MANAGEMENT OF UNCERTAINTY

Carlos Martín Saborido

CHAPTER

5

5.1. INTRODUCTION

As has been mentioned in previous chapters, MCDA is basically a tool to help find a solution to a problem posed with several possible solutions, where a set of criteria will allow us to decide which solution is the most appropriate for our problem. We need to note the way in which these criteria are weighted, that is, how much each criterion 'weighs' in the global decision and how information about the stakeholders involved in each criterion will be obtained.

As the reader may suppose, all these procedures, as well as those mentioned in previous chapters, introduce a certain degree of uncertainty into this type of model. An example of this is the different ways in which a group of patients and a group of decision-makers will weight the criteria.

In the following sections we are going to delve into these concepts of weighting and scoring the criteria, as well as the methods of aggregation and the identification and management of uncertainty.

5.2. WHAT DOES THE WEIGHTING OF CRITERIA REFLECT?

In general in life, when we have to make a decision in which many actors are involved, we try to get everyone to participate with their opinion, and we usually group these opinions. In decisions in the field of healthcare about the incorporation of new technologies (drugs, devices, programmes...), many actors have something to say. For example, for the National Institute for Health and Care Excellence (NICE), these actors are divided into consultees and commentators, and are available at an early stage in the process of appraisal through an open consultation, and subsequently in a face-to-face meeting called 'the scoping workshop'. In this process, the NICE collects the information from those involved, but only uses it to define the scope of the appraisal. In MCDA, this information about the people involved is collected in a structured manner and grouped into criteria which will be weighted to find out which group of information (criterion) 'weighs' the most. This weighting must therefore reflect the preferences and priorities of those involved between each criterion, that is, how important criterion X is compared with criterion Z, and, obviously, this will be different for each group of people, and even among those of the same group.

For example, let us imagine a health problem for which we have 4 drugs (A, B, C and D, as shown in table 5.1), about which we have information regarding different criteria (shown in the columns of the table). These drugs are evaluated by comparing them with another reference drug. The weighting (in the last column of the table) would show what weight or relative importance each of the criteria has for each agent. It would be logical to think that each type of stakeholder can weight differently (provider, payer, patient, clinician...) depending on their own perception, personal and professional experience, training, value judgements, interests, etc. In this example, after multiplying the score by the weighting of each criterion, it is seen that drug B obtains the highest total score. In the next section we discuss the scoring methods.

TABLE 5.1. EXAMPLE OF MATRIX PROBLEM

| | PRICE | SURVIVAL ADDED | HEALTH PERCEIVED BY THE PATIENT DURING THE TREATMENT CYCLE | MEDICAL COSTS | TOTAL |
|-----------|---|---|---|---|-------|
| | From 5 to -5 in discrete units 5= substantial savings; -5= substantial additional costs compared with the reference treatment | From 5 to -5 in discrete units 5= much better than the comparator; -5= much worse than the comparator | (From 0 to 5, 0 being the lowest degree of perceived health and 5 the highest degree of perceived health) | From 5 to -5 in discrete units 5= much better than the comparator; -5= much worse than the comparator | |
| Drug A | 4 | 4 | 1 | 4 | 3.1 |
| Drug B | 4 | 5 | 3 | 4 | 4.1 |
| Drug C | 4 | 4 | 3 | 4 | 3.7 |
| Drug D | 3 | 4 | 3 | 4 | 3.55 |
| Weighting | 15% | 40% | 30% | 15% | 100% |

Source: author's preparation.

5.3. WHAT IS BEHIND THE SCORING OF THE CRITERIA?

If, as mentioned in the previous section, the weighting of the criteria reflects the importance that each stakeholder gives to the selected criteria, each criterion must also be scored in terms of how it affects the decision. That is, if a criterion were 'the health perceived by the patient', depending on each actor, that criterion could weigh 20% of the total, but the question would arise as to how we could measure the quality perceived by the patient, that is, how we could value it. To answer that question we would have to think about the best way to assess this criterion (for example, it could be good/bad, from 1 to 5). Keep in mind that the different scales and methods of scoring are going to affect the final results. Continuing with the example of table 5.1, the scoring or evaluation of the criteria would be done in the grey shaded cells, and this evaluation can be done in different ways:

- Directly through a quantitative value. For example, in the criterion 'health perceived by the patient during the treatment cycle', a scale between 0 and 5 with the discrete or continuous values of table 5.1.
- By comparing alternatives. For example, we consider the drug of reference A, and we ask the participant how much better/worse each of the other alternatives is (drugs A to D of table 5.1) using the categories: much better the same, worse, much worse; or

using ordinal scales: for example between -5 to 5 in discrete values, that is, zero meaning (the same), negative values meaning 'worse' and positive values meaning 'better'.

- By comparing scenarios which include different combinations of criteria and ratings. For example, "Which of the following scenarios would you prefer?"
 - High price, survival gain 20% higher than the reference, same level of health perceived by the patient during the treatment cycle, or high medical costs.
 - Equal price, survival gain 5% higher than the reference, lower level of health perceived by the patient during the treatment cycle, or lower medical costs.

Collectively, the weighting methods and the criteria score are called 'elicitation modes'. In the next section we shall present different methods of obtaining information and a brief explanation of their similarities and differences.

5.4. HOW ARE THE CRITERIA SCORED AND WEIGHTED?

In general, we can consider four large groups of methods for weighting and scoring the criteria within an MCDA:

- **Direct methods:** Those methods that require the people participating in the analysis to provide objective data that directly address the weights and scores for each individual criterion. This is what we have described above with the example of the criterion 'Health perceived by the patient during the treatment cycle', a scale between 0 and 5 with discrete or continuous values.
- **Hierarchical methods:** Methods in which, instead of using individual data for each criterion, the relative value between criteria is used to calculate weights and scores. To illustrate these methods we can use the example described in the previous section (we consider drug A as the reference and we ask the participant how much better/worse each of the other alternatives is using the categories: much better, the same, worse, much worse; or using ordinal scales: -3, -2, -1, 0, 1, 2, 3, with zero meaning "the same", negative values meaning "worse" and positive values meaning "better").
- **Choice experiments:** In this case, each stakeholder chooses between hypothetical options which have several criteria simultaneously. The weights and scores are derived from comparing several sets of experiments. In this case, the example would be to offer the evaluator scenarios with different combinations of criteria and valuations.
- **Matching methods:** These methods offer different sets of weightings and criteria scores for each alternative, and the evaluator is asked to indicate in which cases the alternatives presented are commutative.

As highlighted by Marsh et al., 2017, there are several techniques within each of the methods described above, although not all are used in the healthcare sector. Below, we illustrate the different methods by describing four weighting and scoring techniques found in the literature.

5.4.1. Techniques of direct scoring

Among the direct scoring techniques, the EVIDEM (Evidence and Value: Impact on Decision Making) framework is being used exponentially in the healthcare sector. As discussed in

Chapter 4, the EVIDEM framework currently includes 13 criteria that can be quantified using direct scores, in addition to 7 qualitative criteria. Both groups of criteria have been developed by analysing the opinions of experts, together with an analysis of the literature (Table 5.2).

TABLE 5.2. EXAMPLE OF DIRECT ESTIMATES: EVIDEM FRAMEWORK

| QUANTITATIVE CRITERIA | SCORING SCALE | CONTEXTUAL CRITERIA | SCORING SCALE |
|---|---|---|--------------------------------------|
| Disease severity | 5= very severe; 0= not severe | Mandate and scope of the health system | Positive, neutral or negative impact |
| Size of affected population | 5= common disease; 0= very rare disease | Priorities of the population and access | Positive, neutral or negative impact |
| Unmet needs | 5= many unmet needs; 0= no unmet needs | Common goal and specific interests | Positive, neutral or negative impact |
| Comparative effectiveness | 5= much better than comparator; -5= much worse than comparator | Environmental impact | Positive, neutral or negative impact |
| Comparative safety/tolerability | 5= much better than comparator; -5= much worse than comparator | Opportunity costs & affordability | Positive, neutral or negative impact |
| Comparative patient-reported outcomes | 5= much better than comparator; -5= much worse than comparator | System capacity & appropriate use of the intervention | Positive, neutral or negative impact |
| Type of preventive benefit | 5= eradication; 0= no preventive benefit | Political, historical and cultural context | Positive, neutral or negative impact |
| Type of therapeutic benefit | 5= cure / save lives; 0= no therapeutic benefit | | |
| Comparative cost of the intervention | 5= substantial savings; -5= substantial additional expenditures | | |
| Comparative other medical costs | 5= substantial savings; -5= substantial additional expenditures | | |
| Comparative non-medical costs | 5= substantial savings; -5= substantial additional expenditures | | |
| Quality of evidence | 5= highly relevant and valid; 0= not relevant and/or invalid | | |
| Expert consensus / Clinical practice guidelines | 5= strong recommendation for intervention above all other alternatives; 0= not recommended | | |

Source: EVIDEM Collaboration, 2017.

The reason why this group of techniques are the most used is the simplicity with which they score the criteria and their subsequent calculation. Nevertheless, this simplicity implies that direct scoring methods may be susceptible to a loss of information, because they use cate-

gorical (ordinal) rather than continuous (cardinal) scales. Another weakness of these methods is that the weighting of the criteria is done separately from their scoring.

Among these methods, the following are also important:

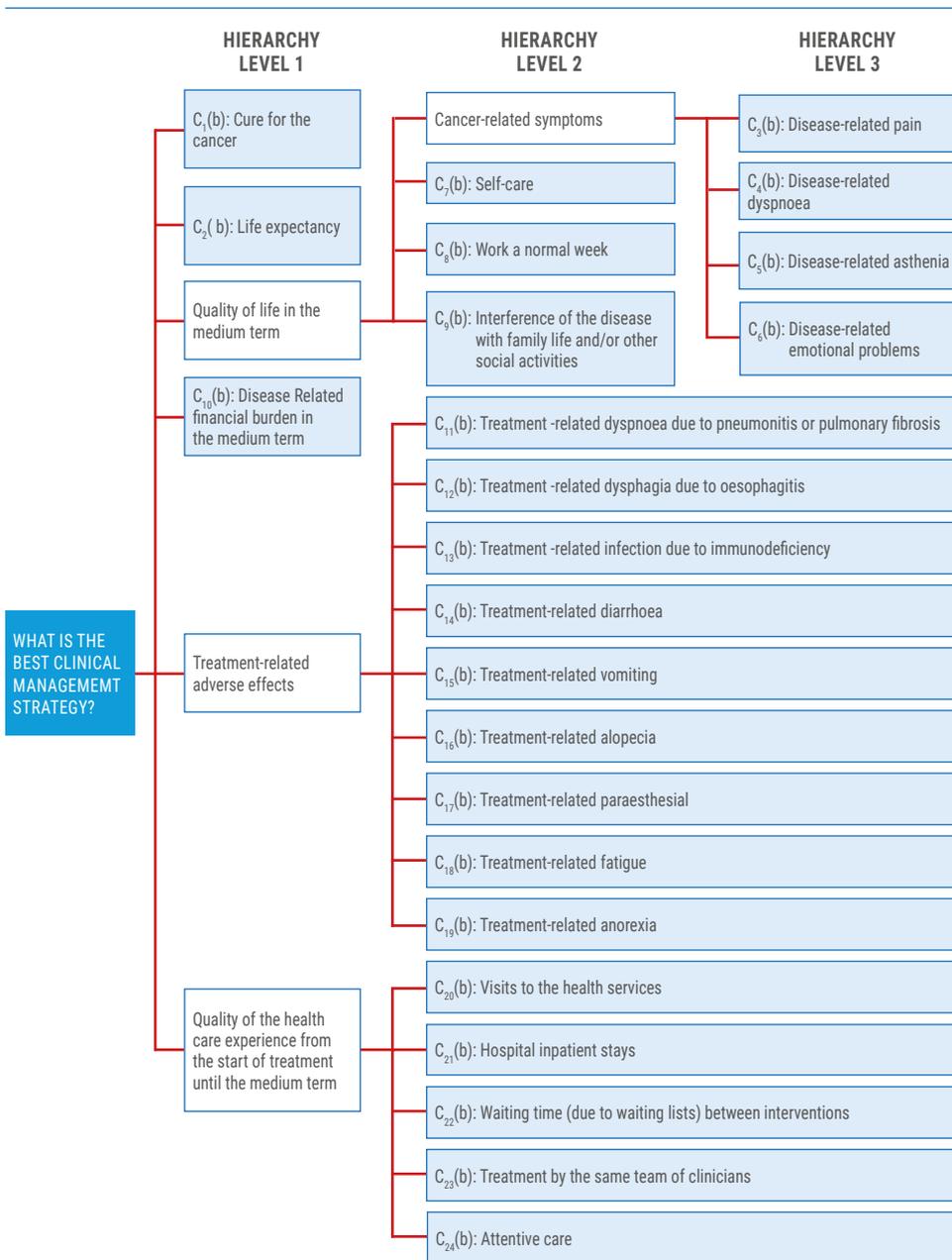
- **ELECTRE (Roy, 1991):** A comparative method which establishes a relationship of superiority between alternatives, i.e., which operates on the assumption that alternative 1 is equal to or better than 2, 3, etc. The analysis is performed by generating matrixes of agreement and discrepancy. The elements of the matrix of matches include an index created to see how much better each alternative is than another for each criterion. The elements of the discrepancy matrix include an index that calculates the maximum positive difference between alternatives. Finally, a table of dominances is generated, in which the dominances of the alternatives are seen and a ranking is established.
- **PROMETHEE (Brans and Vincke, 1985):** The set of PROMETHEE methods (I and II) are based on the relationships of dominance of different alternatives compared in pairs. The analyst compares each pair of alternatives using the defined criteria. In order to determine the partial ordering (PROMETHEE I) or complete ordering (PROMETHEE II) between the different alternatives, the decision matrix is calculated, and then a preference index matrix similar to that of the ELECTRE method. One analyses the 'positive flows' or superiority relationships of some alternatives compared with others, and the 'negative flows' or inferiority relationships. From the analysis of these flows the ranking of alternatives is established.
- **VIKOR (Muñoz and Romana, 2016):** This enables us to work with quantitative and qualitative criteria by applying various techniques to the previous methods, but which have the advantage that they allow us to use different units or scoring methods, which is a limitation of ELECTRE and PROMETHEE. The analysis is carried out in four steps. In the first, the best and worse values are calculated for each criterion, or positive/negative values if the scoring method is of the '-5 to 5' type. In the second step, the parameters S and R, which compare each alternative in each criterion, and the parameter Q, which links the two previous ones, are calculated. In the third step, the alternatives are ordered by the parameters of the previous step. Finally the alternative is selected based on the minimum Q value fulfilling the conditions of 'acceptable advantage' and 'acceptable stability'.

5.4.2. Pair-wise comparison with ordinal scales (hierarchical technique)

Pair-wise comparison is a technique which has traditionally been used to compare the different hierarchies within the hierarchical process of analysis. To understand this technique, we propose the following figure (Figure 5.1).

In this figure we find the criteria for a study (Pozo-Martín, 2015) of various kinds of software for an MCDA implemented in portfolio decision analysis (PDA), and which included this hypothetical hierarchical tree. In this paper, two of the most used types of software are evaluated: Annalisa in Elicia (Annalisa In Elicia, 2015) and M-MACBETH (M-MACBETH, 2015). This hierarchy allows us to simplify complex criteria and to make comparisons within the same hierarchical level.

FIGURE 5.1. HIERARCHY OF CRITERIA FOR A HYPOTHETICAL PATIENT



Source: Pozo-Martin, 2015.

Following the example of Figure 5.1, in order to determine the weights of each criterion, pairs of comparisons would be proposed that would include criteria or sub-criteria but of the same hierarchical level. This pair of criteria or sub-criteria would be accompanied by a 9-point scale, as indicated in Figure 5.2. In this figure we see how we have taken as an example two sub-criteria of hierarchical level 2:

- Diarrhoea related to the treatment
- Vomiting related to the treatment

We ask the participant to indicate which one of both sub-criterion seems more relevant to him/her, and to what extent, by using the scale shown in Figure 5.2.

FIGURE 5.2. EXAMPLE OF COMPARISON BY PAIRS TO DETERMINE THE WEIGHT OF THE CRITERIA IN A HIERARCHICAL ANALYSIS



Source: author's preparation.

Once we have all the answers for all possible pair-wise comparisons, an analysis is performed to check the consistency of the answers. Saaty proposed what is known as the consistency rate (Saaty, 1980; Saaty, 1990), in which a value of this rate lower than 0.1 will mean that the matrix may be considered as sufficiently consistent. If this rate is exceeded, one should consider performing the analysis again.

For the scoring of the criteria an ordinal scale would be used for their quantification, as described in the direct estimation (section 5.4.1).

5.4.3. Discrete choice method

Unlike direct estimation, this method presents the participants with hypothetical alternatives about which they have to decide. These alternatives include the criteria that have been defined for the specific MCDA. With this technique the intention is that the participant should make an assessment of each criterion by combining the different attributes of the hypothetical alternatives presented.

For example, let's go back to the example in Figure 5.1 and take several sub-criteria from levels 2 and 3. We could define drug A and drug B by combining different attributes for each sub-criterion, as shown in table 5.3.

TABLE 5.3. EXAMPLE OF THE DISCRETE CHOICE METHOD

| | PAIN RELATED TO THE AILMENT (0-100) | ASTHENIA RELATED TO THE AILMENT | DYSPNOEA RELATED TO TREATMENT | RISK OF DEATH |
|--------|-------------------------------------|---------------------------------|-------------------------------|---------------|
| Drug A | 30 | Moderate | Severe | 5% |
| Drug B | 10 | Moderate | No | 8% |

Source: author's preparation.

In the case of our experiment, the participant must choose one of the two scenarios. Once there is a large number of results from these experiments, a regression model is generated

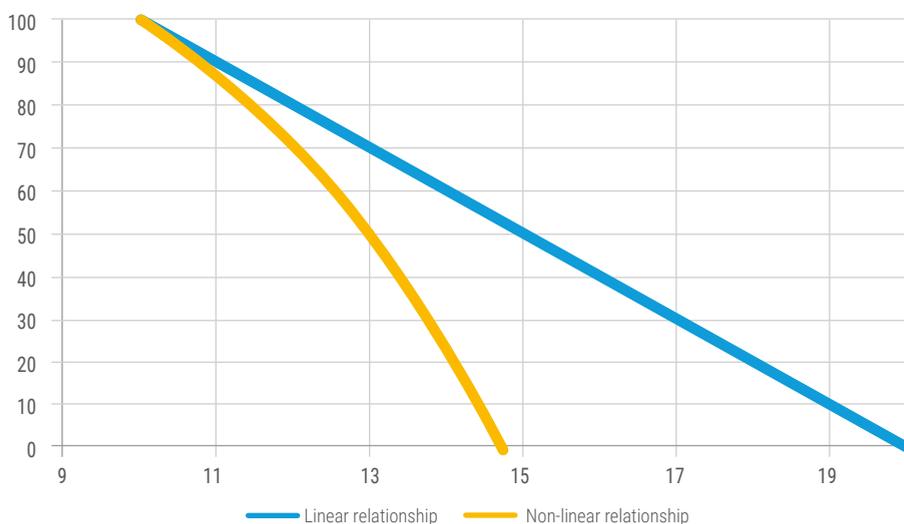
which will allow us to obtain the weighting and the change in preferences of the participants among the criteria considered.

5.4.4. Keeney-Raiffa method

This method is based on the idea of the existence of a utility function associated with each of the attributes that are considered in the decision problem (Keeney and Raiffa, 1976). The objective is to obtain a measure of the total utility of each of the alternatives by means of the composition of the 'n' utility functions corresponding to the 'n' attributes. That is, for each attribute a utility function is calculated (usually partial functions), and then a single utility is calculated for each alternative evaluated in the MCDA.

For example, if we have a criterion which is '10-20% probability of suffering an adverse event', the partial utility function of this criterion is constructed from the relationship between the criterion range (10-20%) and a discrete numerical scale, usually 0-100. This can be seen graphically in Figure 5.3, in which we have represented two value functions, one linear and one non-linear, since both circumstances can occur. If the relationship is linear, the final calculation will be simpler and if it is non-linear, it will be necessary to use more complex methods. To construct the function, the parties are asked about the distances within the range. For our example, a question could be "What value would you consider the average between 10% and 20% probability of suffering an adverse event?" or "Do you consider the difference between 10% and 11% probability equivalent to the difference between 19% and 20% probability?". This process is repeated as many times as considered necessary until it has the form of the defined function.

FIGURE 5.3. EXAMPLE OF PARTIAL UTILITY FUNCTION



Source: author's preparation.

The Keeney-Raiffa method also includes the permutation of weights coupled with partial utility functions. This permutation of criteria is done as follows:

- The most important criterion for the decision is taken and assigned a value of 100.
- The participants in the MCDA are shown a comparison of this criterion with another one and they are asked to give us the relative value of the criterion to be compared with the first one.

Continuing with the previous example, if we have as the most relevant criterion '10-20% probability of suffering an adverse event', that means that we have given it the score of 100. We can take another criterion, such as 'probability of between 20% and 30% of developing mild dyspnoea', and present the following comparison: if an increase of between 10% and 20% in the probability of suffering an adverse event has an assessment of 100 (scale from 0 to 100). How would you rate an increase of between 20% and 30% in the probability of developing mild dyspnoea?

This process would be repeated with each of the criteria until a ranking of criteria is obtained.

5.5. HOW DO WE COMBINE THE WEIGHTING AND THE SCORING?

Within the methods used in the MCDA described in the previous sections, we have described the weighting and scoring methods that are most frequently used in the healthcare area. In the same way, there are different methods of combining weighting and scoring. These methods are collectively called aggregation methods, since their purpose is to add the preferences to obtain a global value, which represents the attributes of each alternative and helps us to make the decision. Among the aggregation methods, within the multi-attribute value measurement models (MAVT, MAUT and AHP), we find:

- Additive method
- Regression method
- Multiplicative method
- Matrix algebra

The additive method is simply to multiply each score in each criterion by its weighting and thus calculate the value of each alternative. It is also one of the most used, since it has the advantage of being easier to understand and communicate to participating stakeholders and to the decision-makers.

The regression method is the one used when the weights and scores have been extracted using a discrete choice experiment. Thus, after presenting all pair-wise comparisons, one has a score which shows the preference of each alternative for each attribute. In this way, a regression model is constructed, which estimates the probability for each alternative of being the preferred one.

The multiplicative method is similar to the additive method, but sometimes the scores are raised to the weighting of the corresponding criterion, or the individual utilities of each participant are taken into account, as in the case of multi-attribute utility techniques.

Finally, we can perform the synthesis or aggregation of values using matrix algebra.

In the four methods that we have explained in this chapter (5.4.1 to 5.4.4), those of direct estimation and Keeney-Raiffa can use the additive or multiplicative technique, in the case of the pair-wise comparison (AHP) matrix algebra would be used, whereas, in the case of discrete choice, we would use the regression method (Table 5.4).

TABLE 5.4. SCORING AND AGGREGATION METHODS

| SCORING METHODS | AGGREGATION METHODS |
|-----------------------------|----------------------------|
| Direct scoring method | Additive or multiplicative |
| Keeney-Raiffa method | Additive or multiplicative |
| Pair-wise comparison method | Matrix algebra |
| Discrete choice method | Regression |

Source: author's preparation.

At this point we might wonder when to use the additive technique and when the multiplicative technique. Well, the general rule is that when the criteria are independent in their development, that is, when there is no interaction between criteria, the additive method will be used. If we are in situations in which this interaction or dependency exists, then we will use the multiplicative method.

The choice of aggregation method is relevant, since it can influence the result of the model. So it is appropriate to take this fact into account in two ways:

- If the choice of method is fixed (e.g. the additive method), then heterogeneity and uncertainty should be explored from the beginning by the uncertainty management methods explained in section 5.7.
- By directly making a final sensitivity analysis by exchanging the additive and multiplicative method and studying the differences.

5.6. DISCUSSION ABOUT SCORING AND WEIGHTING

When one considers the different methods of MCDA that have been presented in the previous chapters, they are usually compared in terms of the complexity of the techniques used (Guitouni and Martel, 1998), the ease of use of the tools, and, on fewer occasions, the weighting techniques (van Til et al., 2014). In the area of healthcare, the decision-maker's point of view is very relevant, since, in the end, it is not just a decision of public policy such as the construction of a bridge, but a decision to finance a health technology which can affect the health of the citizens, and whose value is greater than that of any civil works. For this reason, it is necessary to pay attention not only to the computer support or ease of use of an analysis, but to the validity of its results. This same approach has long been experienced with economic appraisals based on mathematical models in which, owing to their assumptions, the validity of the results can sometimes be questionable.

In this sense, the validity of the results in MCDA will depend on the chosen criteria being correctly weighted and valued, capturing the real preferences of the participants (actors in the

system), and even more, on these preferences having the necessary quality to be applied in the analytical model that is going to be used.

The level of validity required of an MCDA will depend on the objective pursued, since when an MCDA is used to assess a new health technology or to decide the reimbursement price of a drug, the decision must be very well supported. Additionally, if the objective is simply to carry out a ranking, this validity can be more relaxed.

Another matter to be taken into account, and which we see many times in the discussion of studies which use MCDA, is the level of knowledge of the participants (Marsh et al., 2017). This is a factor that can lead us to decide on some techniques or on others, since they do not involve the same effort for the participant, for example, to decide between hypothetical scenarios which we described in section 5.4.3, as to make the comparisons with the permutation of weights of the criteria of the hierarchical techniques of section 5.4.4. In addition, as we explained above, some of the techniques involve answering questions for the weighting and also answering others for the assessment. Finally, this section about the level of knowledge includes the fact that some of the techniques have to be applied in fairly prolonged face-to-face sessions, whereas others can be applied in a survey. This has pros and cons, logically: although a face-to-face session allows interaction, it can also become more burdensome and require participants to reserve a specific period for it. By contrast, the use of surveys (discrete choice) allows the participant to do them whenever he/she wants, but restricts interaction. Other relevant considerations in relation to face-to-face sessions versus surveys are heterogeneity and uncertainty, that is, in face-to-face sessions in which a group of participants interact, they tend to seek and achieve consensus among themselves, whereas in interviews such a consensus cannot be reached.

Another relevant factor in the choice of techniques is the way in which the results are presented and interpreted. This is not a big problem as long as numerical results of type 0-1, such as those offered by the methods which use approximation of the weighted sum, are accompanied by the meaning of the number and the change within the range. In the case of direct choice, the result, in the form of regression, expresses the probability that one participant prefers one alternative to another. Another important issue in relation to the results is the transparency, not only of the final numbers but also of the process by which they have been reached, so direct estimation processes unsupported by complex statistical analyses will tend to be more easily traceable.

Finally, it should be considered that, like any model and analysis, MCDA includes some uncertainty about the robustness of the model, the weighting, the assessment and the data analysis. This is why we are going to devote the next section to tackling this uncertainty.

5.7. ASSESSMENT AND MANAGEMENT OF UNCERTAINTY

As in all modelling exercises, the result depends on the assumptions made and the decisions that were taken while the model was being constructed, and not only on the inputs and the structure itself. That is why when we build a model we begin to make a list of the areas where we are going to find uncertainty, doubt, unexplained variability... Traditionally, in economic appraisal, the following types of uncertainty are discussed:

- Structural uncertainty: this will be connected with the design of the model that we have built, and with the way we did it in order to achieve the stated objective.

- Stochastic uncertainty: this relates to that which is due to chance and which we can find simply by taking a sample of a population and measuring a variable. We shall find that, even when the sample is composed of very similar people, there may be a variability due to chance. This uncertainty is reflected by standard deviation or variance.
- Uncertainty about the parameters: this is associated with a parameter or specific variable within a group of individuals when we compare it with another group of individuals extracted from the same population. In statistics, this uncertainty can be measured through standard error.
- Heterogeneity: this is the uncertainty which is associated with the characteristics of each person. That is, when we compare people, they are not exactly the same and therefore they will differ in something. The quantification of these differences is what gives us the idea of heterogeneity. We must bear in mind that sensitivity analyses can capture observable heterogeneity if they are designed *ad hoc* for each parameter, but unobservable heterogeneity is captured by more global sensitivity analyses, such as probabilistic sensitivity analyses.
- Uncertainty about the clinical evidence: this is associated with the information included in each criterion. That is, the sources of information and the methods and strategies used to find and report the evidence about the alternatives to be evaluated. Some tools, such as EVIDEM, include specific criteria to evaluate the quality of the evidence.

Overall, all this uncertainty is evaluated by a series of sensitivity analyses that will quantify the impact of the uncertainty on the result, thus giving us a measure of the validity of the results, as we mentioned in the previous section.

5.7.1. Structural uncertainty

When we refer to uncertainty in the structure of the model, we are basically referring to how the decision is taken about the model to be used and how the criteria that are going to be part of the model have been decided. Although it seems that by working on these two points we could control the structural uncertainty, we must bear in mind that prior to the choice of the model and the criteria there was a selection of the stakeholders/participants and a framing of the problem to be addressed. This means that if we have not included the relevant participants (patients, decision-makers, industry ...) to address the problem, we shall begin with a problem that is difficult to solve with a sensitivity analysis.

To try to avoid uncertainty at the time of choosing the criteria, it is advisable to use a combined top-down and bottom-up approach, since with the combination of both we can be more sure of having included all the relevant criteria, and of having a consistent hierarchical tree of criteria. That is, if we start by listing the big criteria of the problem which we want to address and then proceed to shred them down to sub-criteria of smaller size, and then go back the other way, we can see whether our hierarchical tree is the right one. All this process must be debated, discussed and agreed upon among the actors. That is why it was remarked earlier that the correct choice of actors/participants is crucial. We must not lose sight of the fact that, despite the combined top-down and bottom-up exercise, the final list of criteria should be as simple as possible, but still exhaustive. Unfortunately, there are still no guides that allow us to know the optimal number of criteria to be included or the type of criteria to be included, so the best way to test structural uncertainty is to launch the model using different alternative groups of criteria or decision trees (van Til et al., 2014; Ijzerman et al., 2008).

5.7.2. Uncertainty in the scoring

This uncertainty is related to stochastic uncertainty, which is uncertainty about the parameters and the heterogeneity that was described at the beginning of this section 5.7. When we analyse methods of assessment, we can quickly think about the types of uncertainty which we mentioned at the beginning of this section, together with the methods described in the chapter. When assessing a criterion, a participant can make different assessments at two different times, especially if scales with a multitude of values are used. Imagine a scale of 100 points for a criterion, and the participant being asked to evaluate the criterion today and again within 15 days (re-test). It is very possible that, in scales with so many values, there is a variation within the same individual. In the same way there may be variation between apparently identical participants, the more so with those who are not identical, or between participants who represent different actors. That is why this stage of the MCDA is a crucial source of uncertainty. To incorporate this uncertainty into the model and to be able to assess it later, one has to use the necessary statistics of centralization (mean, mode, median), as well as the statistics of dispersion (standard deviation, confidence intervals...) which quantify the uncertainty. To assess the heterogeneity that may come from the different participant profiles, subgroup analyses can be performed to see whether a more in-depth analysis is needed in this regard. This can happen when there are, for example, clinical profiles from different areas such as primary/specialised healthcare, public/private sector...

Finally, we should not forget that when criteria are presented based on data from both the literature and experts (when there are no published or accessible data), these data must incorporate the uncertainty inherent in their aggregation. For example, if we talk about a criterion of average survival, we must include the confidence interval so that when handling this data in partial functions, such as those described in section 5.3.4, we can adjust the criterion values to the chosen scales. This point relates to the uncertainty about clinical evidence mentioned at the beginning.

5.7.3 Uncertainty in the weighting

At this point, we have already concluded that all the uncertainty in the weighting comes from the participants in the MCDA. This is because the weight of the criteria, whatever the method, is influenced by the variability between participants, intra-subject variability, etc.

Classically, the arithmetic mean and its standard deviation are used as means of obtaining the overall weight of each criterion. Nevertheless, there are other approaches such as the geometric mean used in hierarchical techniques.

In terms of heterogeneity, as in the uncertainty surrounding the scoring of the criteria, we must pay attention to the possible subgroups and evaluate them in depth, looking for sources of heterogeneity that may affect the result. Regarding intra-subject variability, if, for the weighting, the participant is only asked for one per criterion, this variability would not exist, otherwise, and if a re-test has been carried out, we would have to reflect it, just as before.

Finally, once we have seen how to include or reduce uncertainty in scoring and weighting, we must mention the aggregation methods, that is, those which are used to combine scoring and weighting. The simplest method, which is also the easiest to understand, is the additive method, in which each criterion is multiplied by its valuation and then these quantities are

added for each alternative. This involves making a decision about whether to first take the average of the weightings of all the participants and multiply it by the average of all the scores, or whether to multiply the weighting and the score of each participant and then average the product of each participant. The two methods they would give us different results in terms of variability, so, once again, this uncertainty must be reflected by means of standard errors or adequate confidence intervals.

5.7.4. Sensitivity analysis in MCDA

Once we have identified the different sources of uncertainty in the model, we must decide how to make explicit the impact of this on the results. A first idea could be to perform different MCDA for the different uncertainties described in each phase, but this would be too cumbersome, in addition to being an infinite task. It is recommended to perform a sensitivity analysis on the model as a whole, incorporating all the sources of uncertainty. In the conventional economic appraisal, this analysis is carried out in two ways: deterministically or probabilistically.

A deterministic analysis consists of varying the values of the parameters each time and observing the influence of this change on the result and on the conclusions reached. We can vary one parameter or several parameters at a time. For example, we can vary the weight of a criterion by a value of its confidence interval or its range, and observe the impact on the result. It is a simple method, but it would involve a lot of analysis and we would not be able to observe the interaction between several parameters at the same time.

By contrast, a probabilistic analysis allows us to combine several sources of uncertainty, for example, the weights of the criteria, the scoring or uncertainty about the data used to produce the criteria. A statistical distribution is attributed to each parameter that is considered relevant or influential in the analysis, taking into account the definition and characteristics of each parameter, and random estimates of each of the values are made. So, instead of using, for example, the average weight of a criterion, we construct a probability function with the average value and the standard deviation. In this way, the distribution reflects the uncertainty so that we would be able to see, graphically, one probability density function, that had, for example, a normal distribution, very narrow at the base, indicating little uncertainty, and another with a broad base, which would indicate the opposite.

Once we have defined the probability functions for each parameter in the model, we take a random value of each parameter, calculate the result of the model and repeat this process many times (e.g. 1,000; 10,000). Later, we calculate the average result of the 'n' times that we have repeated it, and we observe whether the result is stable for a high percentage of the times. For example, if in 95% of the times the same chosen alternative emerged out of two possible alternatives, this would suggest that uncertainty was not affecting the model. By contrast, one alternative emerged as the best option 30% of the time, another 30% of the time another one emerged, and a third the remaining 40%, then we would be faced with a doubtful result which was highly influenced by uncertainty.

5.8 CONCLUSIONS

As a summary of what is mentioned in this chapter, regarding the methods of scoring and weighting, we can retain two main ideas. The variety of MCDA techniques is so great that a

consensus or guide would be necessary to choose the most appropriate technique, since a researcher in health or in economic evaluation who is not an expert in MCDA may not know which technique is the most appropriate to achieve their objective. Furthermore, as has been happening with modelling in economic appraisal with Markov models (many evaluators use this model because they do not know any other and do not investigate other modalities), and as the law of the gold hammer states, "if your only tool is a hammer, then every problem looks like a nail" (Maslow, 1966), we run the risk of only using the EVIDEM framework regardless of the objective of our analysis.

With regard to uncertainty, it is introduced into an MCDA in different phases (problem structuring, selection of participants...), and through different elements (parameters, clinical evidence...), and it is often not possible or even desirable (in the case of heterogeneity) to reduce uncertainty, but in any case this uncertainty must be made explicit, and there must be a study of how it affects the results of the analysis.

We have reviewed how the quantitative results of the analysis depend on the way in which they weigh and assess the criteria, as well as the method of aggregation, while the interpretation of the result of the model also depends on the way in which the result is presented to the decision-maker.

Finally, the sensitivity analysis is the way to evaluate the uncertainty of the analysis. These analyses (deterministic or probabilistic) will help to quantify the uncertainty of the analysis. We must bear in mind that the possibility of performing sensitivity analyses can be infinite, and we can 'twist' the numbers in innumerable ways, but we must not lose sight of the fact that the MCDA is a tool for decision-making and that the uncertainty must ultimately be evaluated by the decision-makers.

REFERENCES

Annalisa in Elicia. Available from: <http://www.maldaba.co.uk/products/elicia/>.

Brans JP, Vincke PH. A preference ranking organization method – (the Promethee method for multiple criteria decision-making). *Manag Sci.* 1985; 31(6): 647–656. doi:10.1287/mnsc.31.6.647

EVIDEM Collaboration. *Concept & definitions (v4.0)*. 10th Edition. 2017.

Guitouni A, Martel J-M. Tentative guidelines to help choosing an appropriate MCDA method. *Eur J Oper Res.* 1998; 109: 501–521.

Keeney RL, Raiffa H. *Decisions with multiple objectives: preferences and value tradeoffs*. New York: Wiley; 1976.

Ijzerman MJ, van Til JA, Snoek GJ. Comparison of two multi-criteria decision techniques for eliciting treatment preferences in people with neurological disorders. *Patient: Patient-Cent Outcomes Res.* 2008; 1(4): 265–272.

Marsh K, Goetghebeur M, Thokala P, Baltussen R. *Multi-criteria decision analysis to support healthcare decisions*. Springer; 2017.

Maslow AH. *The Psychology of Science a Reconnaissance*. 1966.

M-MACBETH. Available from: <http://www.m-macbeth.com/en/m-home.html>. Accessed: 5 May 2015.

Muñoz B, Romana M. Aplicación de métodos de decisión multicriterio discretos al análisis de alternativas en estudios informativos de infraestructuras de transporte. *Pensamiento matemático*. 2016; (2):27-45.

NICE Technology Appraisal Guidance. Available from: <https://www.nice.org.uk/about/what-we-do/our-programmes/nice-guidance/nice-technology-appraisal-guidance>. Accessed: 10 oct 2017.

Pozo-Martín F. *Multi-Criteria Decision Analysis (MCDA) as the basis for the development, implementation and evaluation of interactive patient decision aids [Doctoral Thesis]*. London: London School of Hygiene and Tropical Medicine; 2015.

Roy B. The outranking approach and the foundations of electre methods. *Theor Decis.* 1991; 31(1): 49–73. doi:10.1007/Bf00134132.

Saaty TL. *The analytic hierarchy process: planning, priority setting, resource allocation*. New York: McGraw-Hill; 1980.

Saaty TL. How to make a decision: the analytic hierarchy process. *Eur J Oper Res.* 1990; 48(1): 9–26.

van Til J, Groothuis-Oudshoorn C, Lieferink M, Dolan J, Goetghebeur M. Does technique matter; a pilot study exploring weighting techniques for a multi-criteria decision support framework. *Cost Eff Resour Alloc.* 2014; 18: 12-22 doi:10.1186/1478-7547-12-22.



REVIEW OF METHODOLOGICAL AND GOOD PRACTICE GUIDELINES IN MCDA

Marta Trapero-Bertran

CHAPTER

6

6.1. INTRODUCTION

This chapter provides an emerging good-practice guidance in the use of MCDA to support healthcare decisions. The aim is to transmit the basic technical notions underlying the methods used in this type of analysis. The chapter is intended mainly for healthcare professionals who have heard about the methodology of MCDA but are not specialists in it.

An MCDA requires a technical design which reflects both the social decisions (who participates, when and how, etc.) and the techniques used (what methods, what software, etc.). MCDA is a structured and explicit approach to better inform decision-making in our context, decisions about healthcare. Decision-making that occurs in the absence of criteria and objective evaluation processes can result in variability in the factors considered, discrepancies in the way in which the importance of the criteria is weighted, and inconsistent choices. Thus, using structured and explicit approaches that require the evaluation of many criteria can significantly improve the quality of decision-making. Methodological guidelines and good practices try to homogenise the different steps in performing an MCDA, while accepting the subjectivity of the value judgments used throughout the process (priorities, importance of values, etc.). In an MCDA, the decision-makers are those who choose between alternatives, using weights and rankings. So establishing a common definition, framework and good practice guidelines for conducting MCDA can contribute significantly to more informed, more efficient and more equitable decision-making in healthcare. The main objective of this chapter is to summarise and explain coincidences and differences among the various good-practice guidelines in MCDA in healthcare, by reviewing the literature about the guides to this type of analysis when applied to healthcare decisions.

The chapter is structured as follows: after this first introductory section, the second section focuses on explaining and describing the methods used to carry out this review of the literature about methodological and good-practice guidelines in MCDA. In the third section, the results of this review are described in five sub-sections. The first sub-section aims to describe the process of conducting an MCDA. The second sub-section describes the most commonly used criteria in MCDA. The third sub-section aims to list the different types of MCDA models distinguished in this review. The fourth sub-section cites some examples found in the review of healthcare decisions to which MCDA could easily be applied. The fifth and last sub-section describes the checklist of ISPOR good practice guidelines for MCDA. Finally, the fourth and last section of this chapter consists of a discussion, offering a reflection on the aspects which generate the most controversy about the methodology of this analysis, and mentioning the potential next steps that could guide research into this methodology in the coming years.

6.2. METHODS FOR THE REVIEW OF METHODOLOGICAL AND GOOD PRACTICE GUIDELINES IN MCDA

A systematic search was made of reviews of MCDA in the healthcare field to identify studies and methodological guidelines which have applied MCDA in a healthcare context in the last ten years (published since 2007). The objective of this review was to identify and summarise the MCDA techniques used in the field of healthcare, so it was not necessary to include reviews published long ago because, although these techniques have been used for many years, interest in reviewing this methodology and in producing methodological guides about it has only re-appeared recently. So the systematic reviews published about this methodology

are recent and none were published more than 10 years ago. The search was restricted to articles published in English and Spanish. The criterion for including studies was that they had to be systematic reviews of the application of MCDA in healthcare, or methodological guides to MCDA produced by a public or private healthcare institution or organization and published by one of the evaluation agencies of the countries with the most experience in appraising interventions or healthcare programmes, such as the Canadian Agency for Drugs and Technologies for Health (CADTH) or the National Institute for Health and Care Excellence (NICE) in the United Kingdom. Both are relevant examples, but they are not the only ones.

The following databases were used to perform the search: Medical Literature Analysis and Retrieval System Online (MEDLINE), Public Medline (Pubmed) and Web of Knowledge. The search terms used were: 'multi-criteria decision making' OR 'multi-criteria decision making' OR 'multi-criteria decision analysis' OR 'multi-criteria decision analysis' OR 'MCDA' OR 'multiple criteria decision' OR 'multi-criteria analysis' OR 'multi-criteria analysis'. These search terms were used in conjunction with healthcare-related words such as 'health' OR 'health care' OR 'healthcare' OR 'medic*', and terms related to methods or reviews of the literature such as 'systematic review' OR 'review' OR 'bibliometric' OR 'method*' OR 'technique*'.

In addition to these databases, the following catalogues and conference programmes were searched manually: Society for Medical Decision Making (SMDM), Health Technology Assessment International (HTAi), International Society for Pharmacoeconomics and Outcomes Research (ISPOR), International Health Economics Association (iHEA) and European Health Economics Association (EuHEA). The author of this chapter reviewed the bibliographic references of the relevant studies found in the review to identify studies of interest for this review. The author examined the studies identified in the review to determine whether they fulfilled the specified eligibility criteria.

6.3. RESULTS

This review found several good-practice guidelines published by various public and private bodies that have examined the application of MCDA. For example, the Agency for Healthcare Research and Quality (AHRQ) (U.S. Department of Health & Human Services, 2012), the CADTH (Husereau et al., 2010), the Department of Health in the United Kingdom (Department for Communities and Local Government, 2009; Marsh et al., 2013), the NICE in England and Wales (Thokala and Duenas, 2012), the Office of Health Economics (OHE) in the United Kingdom (Devlin and Sussex, 2011), the German Institute for Health Quality and Efficiency (IQWiG) (Danner et al., 2011), the International Health Policy Program (IHPP) (Youngkong et al., 2012) and the Health Intervention and Technology Assessment Program (HITAP) in Thailand (Youngkong et al., 2012). These organizations/agencies have used and proposed MCDA as an approach to incorporating the preferences of stakeholders in research into comparative effectiveness (U.S. Department of Health & Human Services, 2012), evaluating new health technologies (Husereau et al., 2010), prioritising investment in public healthcare interventions (Marsh et al., 2013), evaluating orphan drugs (Sussex et al., 2013), supporting risk/benefit appraisal and weighting considerations in the assessment of quality and efficiency in health care (Danner et al., 2011), and developing a portfolio of healthcare services with universal coverage (Youngkong et al., 2012).

The author has also identified two bibliometric analyses of applications of MCDA in healthcare (Diaby et al., 2013; Adunlin et al., 2015) and a series of documents explaining the different

methods used in an MCDA: reviews of the literature about the different methodologies used (Marsh et al., 2013; Liberatore and Nydick, 2008), documents specifying all the steps of this type of analysis (Diaby and Goeree, 2014), reviews of the criteria used in healthcare decisions (Guindo et al., 2012), methods of using MCDA for health technologies (Angelis and Kanavos, 2017; Thokala and Duenas, 2012), methods of MCDA to incorporate the opinions of patients (Dolan, 2010) and good-practice guidelines in MCDA (Marsh et al., 2016; Thokala et al., 2016). In the next sub-sections, the different methodologies will be described, with special emphasis on their advantages and disadvantages for each of the cases and recommendations which have been established for this type of analysis when applied in the field of healthcare.

6.3.1. An overview of the steps necessary to conduct an MCDA

MCDA is a way of evaluating complex problems, characterised by any combination of monetary and non-monetary objectives, and which divides each problem into more manageable pieces to allow data and judgements to be applied to those pieces, and then re-assembles the pieces to present a coherent general picture to the healthcare actors. This process must help us to structure, think and reflect about decisions, but it does not replace decision-making. MCDA can be used retrospectively to evaluate the alternatives to which resources have already been assigned, or to evaluate the alternatives which, so far, have only been proposed. So in the following explanations of MCDA there is no need to distinguish between these two uses, although in practice the approaches will be different.

The steps to be followed when performing an MCDA vary depending on the source and the nature of the information used to inform decision-making, but they include four common steps: 1) identification of the alternatives to be evaluated; 2) identification and establishment of the appraisal criteria for the different alternatives; 3) measurement and prioritisation of interventions in accordance with the established criteria; and 4) weighting criteria which measure the relative importance of each criterion in comparison with the others in order to produce a general appraisal of each intervention. This sub-section provides an overview of the process of performing an MCDA.

Four studies have been found which clearly specify the different steps in performing an MCDA (Department for Communities and Local Government, 2009; Thokala and Duenas, 2012; EVIDEM Collaboration, 2017; Thokala et al., 2016).

The following list of eight steps is intended to be a compilation of the different steps described in the articles found in this review, although they all share the same general principles and concepts (they are also shown in table 3.7 of this book):

- 1. Definition of the decision problem and its context:** identifying the objectives, the type of decision, the alternatives, the members that will constitute the committee which makes the decision, and the required results. Also identifying potential biases which could generate uncertainty in the analysis, and considering the specific context of this appraisal.
- 2. Selection and structuring of the criteria:** identifying and structuring the elements of value or relevant criteria for evaluating the different alternatives. These criteria can be identified in different ways. Sub-paragraph 6.3.2 specifies the different options for identifying them, and those which are the most commonly used.

3. **Performance measurement:** collecting data about the performance of the alternatives according to the different criteria, and summarising it in a performance matrix.
4. **Weighting of criteria:** obtaining the committee members' preferences for the different criteria. The consistency of the scores of each participant for each criterion should be measured or evaluated.
5. **Scoring of the criteria:** knowing the score that each member gives to each criterion according to the value which they estimate that each of them contributes to the different alternatives.
6. **Calculation of the total value of the alternatives:** adding the scores of the different alternatives for the different criteria, weighted by the weights of the criteria, to obtain the aggregate total value. To do this aggregation it will be necessary to choose a model. In sections 5.5 and 6.3.3 more details are given about the different types of existing models or methods.
7. **Evaluation of uncertainty:** performing an analysis of uncertainty in understanding or interpreting the level of robustness of the MCDA results. This analysis should answer the question: If the interviewees had given other weightings or scores, how would this have affected the final ranking of the different alternatives? List the advantages and disadvantages of the different options selected and compare them in pairs; and evaluate alternative scenarios that could be better and worse than those originally considered.
8. **Interpretation of the results and synthesis of the information:** interpretation of the results of the MCDA, including an analysis of uncertainty to support the final decision-making about the different alternatives.

6.3.2. Criteria most commonly used in MCDA

Once the decision problem has been identified, one of the most important steps is to decide which criteria will be taken into account when evaluating the different alternatives. For example, and as specified in Thokala et al., 2016, authorisation decisions must be informed exclusively by clinical results, whereas decisions about prioritisation may involve a larger number of criteria. Criteria can be identified in different ways, from reviewing previously made decisions to organising focus groups and/or workshops or sessions for that purpose. In an additive model, what is important is that the criteria meet a series of requirements, such as completeness or non-redundancy, which do not overlap, and independence in the preferences which they try to express (see sections 3.5 and 4.3). Once the criteria have been identified, they can be represented as a tree of values (von Winterfeldt and Fasolo, 2009), which break down the total value into criteria and sub-criteria in a visual way. This reviewer found four reviews of the literature which summarise the criteria that have been used the most in MCDA in healthcare (Guindo et al., 2012; Marsh et al., 2013; EVIDEM Collaboration, 2015; Angelis and Kanavos, 2017). The most common criteria considered by these four reviews are: benefits and health outcomes, economic impact of the intervention (also related to the cost, budget or considerations of efficiency), the quality of the evidence and the complexity or severity of the disease. There is a study which performs an MCDA of a portfolio of universal coverage ser-

vices in Thailand and a review of the criteria most used in different agencies and institutions which use the MCDA to make healthcare decisions (Youngkong et al., 2012). Effectiveness and efficacy, cost-effectiveness and ethical and equity implications, and their social implications, are the criteria which this review identified as the most used by institutions such as the CADTH, Agency for Health Quality and Assessment of Catalonia, German Agency for Health Technology Assessment (DAHTA), NICE of the United Kingdom, Swedish Agency for Health Technology Assessment and Assessment of Social Services (SBU), and Veterans Affairs Technology Assessment Program.

Guindo et al., 2012, conducted a review of the literature which is specifically about the criteria for the allocation of resources and healthcare decisions. The authors summarise a list of criteria which they classify into categories. The following list of nine categories is ordered by number of citations of each of the classifications and criteria included: (i) overall context (for example, mission of the healthcare system or interests and pressures of the stakeholders); (ii) benefits and health outcomes of an intervention; (iii) priorities, justice and ethics (for example, priorities of the population, access, vulnerability, etc.); (iv) economic impact of the intervention; (v) complexity of the intervention (organisational requirements, integration, flexibility of the intervention, etc.); (vi) quality of the evidence and uncertainty; (vii) impact of the disease in question (severity and burden of disease, etc.); (viii) therapeutic context of the intervention (unmet need, alternative treatments, etc.); and ix) type of service provided (prevention, effect of the medical service on the patient, etc.).

The review by Marsh et al., 2013, focuses on studies of investment, and not on studies of intervention or authorisation. The greatest variety of criteria was adopted in studies focused on investment. Less than 50% of the latter included cost as a criterion, although three others included the concept of cost through the criterion of cost-effectiveness. One of the reflections made by these authors, and which it is important to highlight, is whether or not cost-effectiveness should be included as one more criterion. For example, there is a risk that the inclusion of cost-effectiveness and cost and health benefit would result in double counting in the sense that the same characteristic of an intervention would be counted more than once. The most commonly used criteria to support the investment decisions described in this review are, in order of frequency of use: impact on health, population size, severity of the disease, cost/budget, complexity of implementation, alternative treatment, cost-effectiveness, quality of the evidence, policy, productivity and local priorities.

As discussed in other chapters, the EVIDEM framework (EVIDEM Collaboration, 2017) defines a standard set of decision criteria or elements of value to be considered in evaluation, which can be: quantitative (quantifiable from a universal point of view, and which define the intrinsic value of the intervention in a given context), or qualitative (not quantifiable from a universal point of view, related to the context, priorities and capabilities of the system, and which define the extrinsic value of the intervention evaluated). The decision criteria included in the EVIDEM are intended to fulfil the methodological requirements that dictate good practices applied to the selection of criteria in all MCDA, which are the following: completeness (including all the components or elements of value that constitute the intrinsic value of the intervention), non-redundancy (the components must not be duplicated), mutual independency (the score of each component is independent of the score of the other components), and operability (each component is defined unambiguously, and the data on which to base the eval-

uation should be available). The 13 quantitative criteria currently considered by the EVIDEM (see section 4.3) form the core of the multi-criteria decision analysis. These criteria create a generic interpretative framework with the common healthcare objective and its maximum motivation. It is important to take this into account when adapting or modifying these generic criteria. The cost/effectiveness ratio is a measure composed of other criteria and does not fulfil the non-redundancy design requirement of MCDA. It could be included as a generic criterion, since many decision-making processes are currently based on this composite measure, but it will ultimately have to be replaced by the criterion with which it overlaps. Additionally, the EVIDEM considers 7 qualitative criteria, selected to complement the value of the intervention with the ethical aspects inherent in decision-making (section 4.3). Nevertheless, special emphasis is placed on the cost-effectiveness criterion, insisting that many actors currently use this measure composed of health costs and outcomes, but that it is important that, if it is included, the other included criteria be reviewed to avoid overlapping or duplication of criteria.

The main groups of value dimensions identified by Angelis and Kanavos, to be considered as evaluation criteria in a study of various European countries, identified through the first and second stages of the model construction process, included: 1. Burden disease (severity, availability of alternatives, prevalence); 2. Therapeutic impact (direct endpoints, surrogate endpoints); 3. Safety profile (adverse effects, tolerability, contraindications); 4. Innovation level (clinical innovation, nature of the treatment, ease of use and convenience); 5. Socio-economic impact (public health, budgetary impact, social productivity); 6. Efficiency considerations; and 7. Others (Angelis et al., 2017a). Finally, the authors excluded the last two dimensions, efficiency and other types of concerns, because according to them they could contradict the theoretical necessary properties that all criteria must possess. With these criteria or dimensions they built a tree value for assessing of new medicines.

6.3.3. Types of MCDA models

Value-based methods lead decision-makers, through a series of judgements which produce quantitative scores which show how well the different alternatives fulfil the decision criteria, to establish relative priorities taking into account some of the decision criteria. The methods in this category include linear weighting (scoring), methods based on multi-attribute utility (goal programming), methods based on dominance (outranking approach) and the analytic hierarchy process (AHP). All of these methods begin with the creation of a model for hierarchical decision, sometimes called the 'value tree', which explicitly defines the purpose of the decision, the alternatives which are considered, and the criteria which will be used to judge how the alternatives achieve the stated objective.

In one of the reviews of the literature found (Marsh et al., 2013), it is specified that the most used method for prescribing decisions is the AHP. Thokala and Duenas, compare the different MCDA approaches according to different dimensions to provide an indication of the potential benefits and limitations of each of the models (see section 3.3). First, there are different requirements in the weightings, depending on the type of MCDA model, with the linear weighting models requiring an additional effort in comparison with the outranking approach and the methods based on multi-attribute utility due to the time needed to interpret the oscillations of the weightings. Similarly, linear weighting models require a significant effort to produce the scores, whereas the outranking approach and methods based on multi-attri-

ute utility can be implemented directly in the values of the attributes. The linear weighting models, however, are easy to understand and allow sensitivity analysis in real time. Both the outranking approach and the methods based on multi-attribute utility are easy to follow, but considerable calculation time is required for the programming of objectives. Similarly, linear weighting models lend themselves to easy visual presentation, whereas the results of the outranking approach and of methods based on multi-attribute utility are difficult to follow. Finally, uncertainty is easier to incorporate in linear weighting models than in the outranking approach and in methods based on multi-attribute utility (Thokala and Duenas, 2012).

Dolan, also makes some comparisons between these types of models. Linear weighting models imply that the process of hierarchical ranking and weighting of criteria and alternatives should improve the strength of the decision-making process by assessing the preferences of decision-makers about the pros and cons of alternatives and priorities relating to the decision criteria (Dolan, 2010). This step should also improve communication between the parties about important aspects of the choice which is made. A great advantage of the approaches based on linear weighting models is that the only input required from the actors is rank-ordering, which is easy to understand and achieve. Nor should it consume much time when added to a clinical consultation. Nevertheless, a major unresolved problem is whether they adequately measure the preferences of the actors. These models indicate which elements are better than others, but do not provide information about the magnitude of the differences. Other methods quantify the preferences of the decision-maker more precisely.

According to Dolan, methods based on multi-attribute utility (MAUT) seek to improve the intensity of decision support beyond linear weighting models by creating standardised scales called utility functions, which measure how well the choices fulfil the criteria and incorporate information about the variability of the options within each criterion to determine the priorities of the criteria. Nevertheless, the MAUT method is very difficult to implement in practice.

The multi-criteria method most used for medical and non-medical applications is the analytic hierarchy process (Liberatore and Nydick, 2008). The AHP was created expressly to be an easy-to-use decision support method, capable of addressing a wide range of difficult decision problems, including those which involve both 'hard data' and less tangible considerations (Dolan, 2010). Like the other methods in this category, it breaks down a decision problem, creates a quantitative comparison of the decision alternatives, and provides a format for performing a 'what-if' sensitivity analysis. Its level of support for the decision is comparable to that provided by the MAUT. Differences with the MAUT include the use of ratio level scales to qualify the alternatives, the availability of several alternative methods to create these scales, the use of a simple peer comparison procedure throughout the analysis and the routine calculation of a control over the internal consistency of these comparisons. The advantages of using the AHP to provide support in clinical decision-making include its flexibility, ease of use and strength of measurement. Its flexibility is due to the different formats available. This makes it possible to adapt the same basic decision supporting method to different users and different circumstances. Its ease of use stems from the pairwise comparisons. People easily learn how to make these comparisons, and the same method can be applied equally well to quantitative data and subjective considerations. Because the AHP is also a theory of measurement, the mathematical operations involved in the analysis are theoretically justified and free from assumptions (Gass, 2001). The routine use of the consistency ratio during the

analysis helps users to avoid making technical mistakes and to monitor the quality of the analysis. The main disadvantage of the use of the AHP for clinical decision support is that the process of pairwise comparison is very time-consuming, which can make it difficult to implement in clinical settings.

6.3.4. Examples of healthcare decisions to which MCDA could be applied

As already mentioned in this chapter and in previous ones, MCDA is a type of analysis which can be applied in several disciplines. This sub-section aims to provide examples of the current and potential use of this type of analysis in healthcare decisions. Chapters 7 and 8 review the practical applications of MCDA in prevalent and rare diseases respectively. Both show specific applications of this type of analysis, which could also be applied to other types of healthcare decisions. The first report on MCDA good practices, published by ISPOR (Thokala et al., 2016) details a table with different examples of healthcare decisions which could be evaluated using MCDA (Table 6.1). The following list of examples is not intended to be exhaustive, and MCDA can be useful in many other decision contexts. Nevertheless, it demonstrates the breadth of decision problems which the MCDA can analyse or evaluate, and it is a list of examples validated and prepared by different health economists who are experts in MCDA.

The criteria used in MCDA depend on the type of health problem to be resolved. The results of the MCDA can be used to support different types of decisions, for example, to evaluate the risks and benefits of different alternatives, to appraise a health technology, to analyse a portfolio decision, to prioritise decisions, to take joint decisions, and to prioritise patients' access.

Table 6.1 describes the type of health decision to which an MCDA could respond, provides examples of who makes the decisions in each of these types, and gives examples of relevant criteria for making the decision, examples of the stakeholders that could provide their explicit preferences, and examples of the types of decisions.

In addition to those mentioned in table 6.1, MCDA has been used for other purposes, for example, for preparing classifications of diseases (Aletaha et al., 2010; Johnson et al., 2014) or for purchasing by hospitals (Dolan, 1989; Dolan, 2008; van Til et al., 2008).

Throughout this review, other examples have been found of the application of MCDA to make healthcare decisions. For example, some works have been published in which MCDA was used as a tool to evaluate/analyse the value of orphan drugs in relation to their prices, postulating it as an alternative to the method currently used to make reimbursement recommendations for drugs (Schey et al., 2017; Sussex et al., 2013). MCDA has also been used to evaluate and prioritise the different statins according to the utilities and benefits provided (Ramlil et al., 2013). Other examples have been published, using the AHP as a tool to distinguish between the different criteria in order to obtain a solution that best satisfies the combination of possible alternatives in the case of the possible purchase of a new CT scanner in a public hospital (Pecchia et al., 2013). In Canada, physiotherapy services were assessed by means of MCDA to make decisions at local level (Dionne et al., 2013). In Thailand, MCDA was used to prioritise HIV/AIDS control interventions in order to improve efficiency (Youngkong et al., 2012). Two other studies have applied MCDA to inform decisions about healthcare coverage in Canada (Tony et al., 2011) and Thailand (Youngkong et al., 2012). All these examples show us that the use of the MCDA

TABLE 6. 1. EXAMPLES OF HEALTHCARE DECISIONS TO WHICH MCDA MIGHT BE APPLIED

| TYPE OF HEALTH CARE DECISION | EXAMPLES OF WHO MAKES THESE DECISIONS | EXAMPLES OF CRITERIA RELEVANT TO THE DECISION | EXAMPLES OF STAKEHOLDERS PROVIDING PREFERENCES | EXAMPLES OF THE TYPE OF DECISIONS |
|---|--|--|---|--|
| Benefit-risk assessment (BRA). | Regulators (Phillips et al., 2011). See below for detail of EMA's assessment of MCDA as a method for BRA. (European Medicines Agency, 2015a; European Medicines Agency, 2015b; European Medicines Agency, 2015c). | Criteria are the different aspects of benefits and risk that are relevant to each new medicine under consideration. | Regulatory committees and/or patients. | Categorical. |
| Health technology assessment (HTA). | HTA bodies, such as G-BA in Germany, NICE in England and Wales, and PBCA in Australia. See below for details of IQWiG's pilot of MCDA for HTA (Danner et al., 2011), Thailand's use of MCDA for universal coverage (Youngkong et al., 2012), and the HTA framework in Lombardy Region (Radaelli et al., 2014). | The criteria used differ between HTA systems, but might include effectiveness, patient need/burden of disease/severity. Note: the role of cost, cost-effectiveness and budget impact as criteria in MCDA is contentious (3 rd Plenary session ISPOR 16 th Annual European Congress, 2013). | HTA committees or general public. | Categorical, ranking or understanding 'value'. |
| Portfolio decision analysis (ODA). | Decisions made by life sciences companies; choosing where the best to direct R&D efforts. See below for a pharmaceutical company's experience (Phillips y Banae, 2007). | The likelihood of success and projected profitability (or consistency with other company objectives) of alternative investment decisions. | Board of directors, or a committee appointed by the board. | Ranking or understanding 'value'. |
| Commissioning decisions priority setting frameworks (PSFs). | Resource allocation decisions made by a local budget holders in the English NHS. Decisions made by private insurers about the bundle of services to reimburse. See below for details on English local budget holders' experience with MCDA (Atroldi et al., 2014). Other examples include the use of PBMA (Mitton et al., 2003a; Mitton et al., 2003b; Peacock et al., 2007), DCE (Marsh et al., 2013; Youngkong et al., 2012), and EVIDEM (Goetghebaur et al., 2008; Goetghebaur et al., 2010) to set priorities. | The criteria used vary considerably, but might include effectiveness, meeting unmet need/equity objectives, meeting government targets, etc. | Committee in charge of making the funding decisions. | Ranking. |
| Shared decision making (SDM). | Decisions made by patients in discussion with their doctors, about the choice of treatments. See below for an example of MCDA being used to assess cancer screening alternatives (Dolan et al., 2013). | For example, effect on life expectancy, quality of life, side effects from treatment, and the process of care. | Patients and clinicians. | Ranking. |
| Prioritizing patients' access to health care. | Prioritization of patients for health care services. See below for the use of "points systems" to prioritize patients awaiting elective surgery in New Zealand (Hansen et al., 2012). These can also be used for transparent equitable, and accountable allocation of scarce resources, such as solid organs among patients waitlisted for transplantation (Kamath y Kim, 2007). | Various measures of patient "need" and ability to benefit from treatment. | Clinical leaders, patient groups, and other health professionals. Organ procurement organization at international, national, or regional level. | Ranking. |

DCE: Discrete Choice Experiment; EMA: European Medicines Agency; EVIDEM: Evidence and Value Impact on Decision Making; G-BA: Der Gemeinsame Bundesausschuss IQWiG: Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen; NHS: National Health Service; NICE: National Institute for Health and Care Excellence; PBAC: Pharmaceutical Benefits Advisory Committee; PBMA: Program Budgeting and Marginal Analysis; R&D: Research and Development.

Source: Thokala et al., 2016.

methodology in the field of healthcare has been explored more intensively over the past 10 years. As a result, it is expected that in the next few years more emphasis will be placed on the homogenisation of methods and good practices in the application of this methodology.

6.3.5. Checklist of good practice guidelines in MCDA

A checklist is a tool used in various management areas of organisations for purposes of evaluation, control, analysis and verification. A checklist can be used to deduce the value of an indicator, or to compare several options, or to establish a 'photo' of the current situation.

Normally, the checklist is presented in the form of questions which are answered in a binary way, responding to whether a characteristic is fulfilled or not, although one can also give more than two response options, but always in a closed manner. The closed nature of the answers provides objectivity, and makes this tool one of the most objective ways of assessing the methodology of an analysis.

In this review of the literature, only one work was found which establishes a checklist for the MCDA (Marsh et al., 2016). In 2014, ISPOR created a Working Group on Emerging Good Practices in MCDA to establish a common definition for MCDA, and to develop good practice guidelines for conducting an MCDA in the field of healthcare. In a first report, MCDA was defined, examples of its use in the field of healthcare were provided, the key steps were described, and a general description of the main MCDA methods was provided. In 2016, a second report was published, which provided guidance about good practices in the implementation of MCDA, including a checklist for the MCDA. This list consists of eight steps, beginning with the definition of the decision problem and concluding with the presentation of the results of this analysis. The checklist is not intended to make recommendations as to what specific methods should be used, but it provides a list of key considerations for designing and evaluating whether the methodology followed in an MCDA has been correct. Each step in the checklist includes a recommendation of validation. Table 6.2 shows the ISPOR Checklist of Good Practices for MCDA.

Although the checklist could be interpreted as implying an orderly or sequential process to implement MCDA, this is rarely the case. Designing an MCDA is a recurrent problem. The authors describe all the circumstances described in the checklist, but they may diverge from the order of the steps described in it. The most desirable and logical order of application of the stages would surely be to weight the criteria first and then score the intervention. Weighting the criteria first allows us to become aware of the importance underlying each of them and to explicitly identify which evaluation criteria are important for each of us. Once the criteria have been identified, the intervention is scored with the explicit personal ranking of criteria. In this way, the score of the intervention should reflect and include the importance of each criterion for each individual. If we first score the intervention without having made the effort to order the criteria according to their relevance and our personal judgement, it is difficult to understand what rationale we are following to evaluate this intervention. The checklist does not consider the possible existence of budgetary constraints, but the full report by Marsh et al., 2016, details the implications for conducting an MCDA when such constraints do exist. Explicit criteria are sometimes used in an MCDA without ever having been quantitatively analysed. This is simply a way to summarise the relevant evidence and help to structure and make the best decisions among the various alternatives. This type of non-quantitative analysis is what we would call a partial MCDA.

6.4. DISCUSSION AND NEXT STEPS IN TERMS OF METHODOLOGY

This section describes and explains the similarities and differences between the various existing reviews of the literature about the methodology of the MCDA and the very few existing good-practice guidelines in MCDA when applied in the field of healthcare.

The good practice document published by ISPOR in 2016 (Marsh et al., 2016) identified several areas for future research in MCDA in the healthcare area, including: (a) the level of accuracy required of an MCDA; (b) the cognitive challenges faced by the different types of stakeholders and the support that can overcome those challenges; (c) the preferences of healthcare stakeholders for the theoretical foundations of the MCDA methods; (d) the value functions which best describe the preferences of the stakeholders; and (e) the best methods to incorporate uncertainty and budget constraints into an MCDA. This report also highlights the importance of assessing the impact of the use of MCDA in healthcare decisions by comparing it with other methods that are used or could be used for the same purpose, although, so far, very few studies have made these comparisons (IJzerman et al., 2008; IJzerman et al., 2012; Van Wijk et al., 2006; van Til et al., 2014).

From this review it is clear that in the few guides or reviews which have been published on the practices or good practices in MCDA, the steps for its conduct are quite homogeneous and there do not seem to be many discrepancies. Nevertheless, it is surprising how little emphasis these guides and recommendations place on the analysis of uncertainty in MCDA. In order to obtain robust results from these analyses, it is very important that the uncertainty be handled and analysed in depth, so there is still a long way to go in establishing a more regulated and explicit methodology in relation to uncertainty. Another of the possible lines of future research could be to identify, from among the various criteria which have been used in MCDA, those which could be used transversally in all decisions about healthcare, and then to classify the criteria according to the type of healthcare decision which has to be taken (authorisation, financing, etc.). It would be interesting if in the near future recommendations were made about the different methods that could be used in MCDA according to the different healthcare decisions that may be included in its use. There is only one existing checklist for MCDA, published by the ISPOR Task Force Report (Marsh et al., 2016). It is certainly not necessary to make another checklist, but one could try to improve that checklist by improving its objectivity and generating options for more closed and specific responses to the various questions raised. For example, in the 'selection and structuring of criteria' section of the checklist, it is stated that the methods used to identify and define the criteria must be reported and justified, and the criteria and value tree must be validated and reported. But these statements could be posed in question format to which one could answer: "Yes", "No", "Partly" or "Not applicable". For example: "Does this work include the criterion of benefits and results in healthcare, the economic impact of intervention also related to the cost or budget, or considerations of efficiency, quality of the evidence, and the complexity or severity of the disease?" "If it does not include any of these criteria, does it explain why?" If we look at a checklist in economic evaluation, although it is true that this is a methodology which has been used and developed more, we can see that it is a tool which is much more specific, easy and direct for measuring its quality.

In short, in order to measure the quality of an MCDA it is necessary to have specific tools to be able to better inform decisions, especially because this type of analysis includes a greater

TABLE 6.2. ISPOR CHECKLIST OF GOOD PRACTICES FOR MCDA

| MCDA STEP | RECOMMENDATIONS | BRIEF SUMMARY OF IMPLEMENTATION |
|--|--|--|
| 1. Defining the decision problem. | <ul style="list-style-type: none"> a. Develop a clear description of the decision problem. b. Validate and report the decision problem. | <p>Determine whether the objective is to classify or evaluate alternatives, whether the decision is unique or whether a re-usable model is required; consider alternatives, actors, and restrictions (such as budgetary ones).</p> |
| 2. Selecting and structuring criteria. | <ul style="list-style-type: none"> a. Report and justify the methods used to identify criteria. b. Report and justify the criteria definitions. c. Validate and report the criteria and the value tree. | <p>The criteria can be identified in documents that describe previous decisions; evaluations to support related decisions; studies of the priorities of the stakeholders; and treatment guidelines. Effective criteria imply completeness, non-redundancy, non-overlapping, and independence of preferences. The individual criteria must be unambiguous, exhaustive, direct, operational and understandable.</p> |
| 3. Measuring performance. | <ul style="list-style-type: none"> a. Report and justify the sources used to measure performance. b. Validate and report the performance matrix. | <p>The method for measuring performance must conform to the general principles of medicine, based on evidence and on local methodology guidelines. These guidelines will often recommend the analysis of clinical trial data or network meta-analyses to generate evidence of functioning. When these data are not available, the opinion of experts should be used.</p> |
| 4. Scoring alternatives. | <ul style="list-style-type: none"> a. Report and justify the methods used for scoring. b. Validate and report scores. | <p>The aim of the scoring is to quantitatively assess interventions, health technologies, or healthcare programmes according to the various criteria which have been considered most relevant. The selection of the scoring method should take account of the cognitive load of the actors, the level of precision required, the theoretical foundations and the heterogeneity of the actors.</p> |
| 5. Weighting criteria. | <ul style="list-style-type: none"> a. Report and justify the methods used for weighting. b. Validate and report weights. | <p>The objective of the weighting is to capture the preferences of the actors about the importance that they give to each of the quantitative criteria. The selection of the scoring method will depend on a series of characteristics of the decision problem. The results of this exercise should make explicit the weighting method used. The most common methods of weighting are that of the scale 1 to 5 (hierarchical) and that of the distribution of 100 points (direct).</p> |
| 6. Calculating aggregate scores. | <ul style="list-style-type: none"> a. Report and justify the aggregation function used. b. Validate and report results of the aggregation. | <p>The objective of the aggregation is to combine scores and weightings in a way which is consistent with the preferences of the actors. The aggregation formula most commonly applied in health MCDA is the additive model.</p> |
| 7. Dealing with uncertainty. | <ul style="list-style-type: none"> a. Report sources of uncertainty. b. Report and justify the uncertainty analysis. | <p>The types of uncertainty that may affect the results of an MCDA must be reported, including imprecise or incomplete parameters of the model, variability of the parameters, the quality of the evidence and structural uncertainty. There are two approaches to measuring the impact of uncertainty (that is to say, which include uncertainty as a criterion in MCDA and in the sensitivity analysis).</p> |
| 8. Reporting and examining findings. | <ul style="list-style-type: none"> a. Report the MCDA method and findings. b. Examine the MCDA findings. | <p>The parameters and results of an MCDA can be shown in various tabular and graphical formats. In the end MCDA aims to be a tool to help the actors to make a decision: their decision, not the decision of the tool. This can be facilitated by presenting the MCDA model to actors and allowing them to explore the results and their sensitivity and uncertainty in relation to the parameters.</p> |

Source: author's preparation from Marsh et al., 2016.

degree of interpretability and subjectivity. Future lines of research should expand the guidelines of good practice in parallel with an increase in the use of MCDA to evaluate different healthcare decisions.

6.5. CONCLUSIONS

This review has found several good-practice guidelines published by different public and private bodies which have examined the application of MCDA. It has also identified two bibliometric analyses of applications of MCDA in healthcare, and a series of documents explaining the various methods used in an MCDA. Nevertheless, there is still a long way to go in standardising this methodology and in homogenising the method. The steps to be followed to perform an MCDA include four common steps: 1) identification of the alternatives to be evaluated; 2) identification and establishment of the evaluation criteria of the different alternatives; 3) measurement and prioritisation of the interventions in accordance with established criteria; and 4) weighting of criteria which measure the relative importance of each criterion in comparison with the others in order to produce a general evaluation of each intervention. The most desirable and logical order of implementation of the stages would certainly be to weight the criteria first and then score the intervention. The multi-criteria method most used in medical and non-medical applications is the analytic hierarchy process. The criteria used in MCDA depend on the type of healthcare decision problem. The results of the MCDA can be used to support different types of decisions, such as evaluating the risk and benefit of different alternatives, evaluating a health technology, analysing a portfolio decision (portfolio decision analysis), making prioritisation decisions, making shared decisions, and prioritising patients' access.

REFERENCES

Adunlin G, Diaby V, Xiao H. Application of multicriteria decision analysis in health care: a systematic review and bibliometric analysis. *Health Expect*. 2015; 18(6): 1894-1905.

Airoldi M, Morton A, Smith JA, Bevan G. STAR—people-powered prioritization: a 21st-century solution to allocation headaches. *Med Decis Making*. 2014; 34: 965–75.

Aletaha D, Neogi T, Silman AJ, et al. Rheumatoidarthritis classification criteria: an American College of Rheumatology / European League Against Rheumatism collaborative initiative. *Arthritis Rheum*. 2010; 62: 2569–81.

Angelis A, Kanavos P. Multiple criteria decision analysis (MCDA) for evaluating new medicines in health technology assessment and beyond: The Advance Value Framework. *Soc Sci Med*. 2017; 188: 137-156. doi: 10.1016/j.socscimed.2017.06.024.

Angelis A, Lange A, Kanavos P. Using health technology assessment to assess the value of new medicines: results of a systematic review and expert consultation across eight European countries. *Eur. J. Health Econ*. 2017a. [http:// dx.doi.org/10.1007/s10198-017-0871-0](http://dx.doi.org/10.1007/s10198-017-0871-0).

Danner M, Hummel JM, Volz F, et al. Integrating patients' views into health technology assessment: analytic hierarchy process (AHP) as a method to elicit patient preferences. *Int J Technol Assess Health Care*. 2011; 27: 369–75.

Department for Communities and Local Government. Multi-criteria analysis: a manual. Communities and Local Government Publications. London; 2009. ISBN:978-1-4098-1023-0.

Devlin NJ, Sussex J. Incorporating multiple criteria in HTA. *Methods and Processes*. Office of Health Economics, 2011.

-
- Diaby V, Campbell K, Goeree R. Multi-criteria decision analysis (MCDA) in health care: a bibliometric analysis. *Oper Res Health Care*. 2013; 2: 20-24.
-
- Diaby V, Goeree R. How to use multi-criteria decision analysis methods for reimbursement decision-making in healthcare: a step-by-step guide. *Expert Rev Pharmacoecon Outcomes Res*. 2014; 14(1): 81-99.
-
- Dionne F, Mitton C, MacDonald T, Miller C, Brennan M. The challenge of obtaining information necessary for multi-criteria decision analysis implementation: the case of physiotherapy implementation: the case of physiotherapy services in Canada. *Cost Eff Resour Alloc*. 2013; 11: 11.
-
- Dolan JG, Boohaker E, Allison J, Imperiale TF. Patients' preferences and priorities regarding colorectal cancer screening. *Med Decis Making*. 2013; 33: 59-70.
-
- Dolan JG. Medical decision making using the analytic hierarchy process: choice of initial antimicrobial therapy for acute pyelonephritis. *Med Decis Making*. 1989; 9: 51-6.
-
- Dolan JG. Shared decision-making—transferring research into practice: the analytic hierarchy process (AHP). *Patient Educ Couns*. 2008; 73: 418-25.
-
- Dolan JG. Multi-criteria clinical decision support. A primer on the use of multiple-criteria decision making methods to promote evidence-based, patient-centered healthcare. *Patient*. 2010; 3(4): 229-248.
-
- European Medicines Agency. Benefit-risk methodology project. Update on work package 5: effect stable pilot (phase I). 2015a. Available from: http://www.ema.europa.eu/docs/en_GB/document_library/Report/2014/02/WC500162036.pdf
-
- European Medicines Agency. Benefit-risk methodology project. Work package 3 report: field tests. 2015b. Available from: http://www.ema.europa.eu/docs/en_GB/document_library/Report/2011/09/WC500112088.pdf.
-
- European Medicines Agency. Benefit-risk methodology project. Work package 4 report: benefit-risk tool sand processes. 2015c. Available from: http://www.ema.europa.eu/docs/en_GB/document_library/Report/2012/03/WC500123819.pdf.
-
- EVIDEM Collaboration. 2015a. EVIDEM v. 3.0.
-
- EVIDEM Collaboration. Concept & definitions (v4.0). 10th Edition. 2017.
-
- Gass SI. Model world: When is a number a number? *Interfaces*. 2001; 31(5): 93-103.
-
- Goetghebeur MM, Wagner M, Khoury H, et al. Combining multicriteria decision analysis, ethics and health technology assessment: applying the EVIDEM decision-making framework to growth hormone for Turner syndrome patients. *Cost Eff Resou Alloc*. 2010; 8: 4.
-
- Goetghebeur MM, Wagner M, Khoury H, et al. Evidence and value: impact on Decision Making—the EVIDEM framework and potential applications. *BMC Health Serv Res*. 2008; 8: 270.
-
- Guindo LA, Wagner M, Baltussen R, Rindress D, van Til J, Kind P, et al. From efficacy to equity: Literature review of decision criteria for resource allocation and healthcare decision making. *Cost Eff Resour Alloc*. 2012; 10(1): 9. doi: 10.1186/1478-7547-10-9.
-
- Hansen P, Hendry A, Naden R, et al. A new process for creating points systems for prioritizing patients for elective health services. *Clin Govern Int J*. 2012; 17: 200-9.
-
- Husereau D, Boucher M, Noorani H. Priority setting for health technology assessment at CADTH. *Int J Technol Assess Health Care*. 2010; 26(3): 341-7.
-
- Ijzerman MJ, vanTil JA, Snoek GJ. Comparison of two multi-criteria decision techniques for eliciting treatment preferences in people with neurological disorders. *Patient*. 2008; 1: 265-72.
-
- Ijzerman MJ, vanTil JA, Bridges JFP. A comparison of analytic hierarchy process and conjoint analysis methods in assessing treatment alternatives for stroke rehabilitation. *Patient*. 2012; 5: 45-56.
-
- Johnson SR, Naden RP, Fransen J, et al. Multicriteria decision analysis methods with 1000 Minds for developing systemic sclerosis classification criteria. *J Clin Epidemiol*. 2014; 67: 706-14.
-
- Kamath PS, Kim W. The model for end-stage liver disease (MELD). *Hepatology*. 2007; 45: 797-805.
-

Liberatore MJ, Nydick RL. The analytic hierarchy process in medical and health care decision making: A literature review. *Eur J Oper Res.* 2008; 189(1): 194–207.

Marsh K, Dolan P, Kempster J, Lugon M. Prioritizing investments in public health: a multi-criteria decision analysis. *J Public Health.* 2013; 35: 460–6.

Marsh K, Ijzerman M, Thokala P, Baltussen R, Boysen M, Kalo Z, et al. Multiple criteria decision analysis for health care decision making – emerging good practices: Report 2 of the ISPOR MCDA emerging good practices task force. *Value Health.* 2016; 19: 125-137.

Marsh, K. Multiple criteria decision analysis for health care decision making – emerging good practices: Report 2 of the ISPOR MCDA emerging good practices task force. *Value & Outcomes Spotlight*, 2016; 25-26.

Mitton CR, Donaldson C. Setting priorities and allocating resources in health regions: lessons from a project evaluating program budgeting and marginal analysis (PBMA). *Health Policy.* 2003; 64: 335–48.

Mitton CR, Donaldson C, Waldner H, Eagle C. The evolution of PBMA: towards a macro-level priority setting framework for health regions. *Health Care Manag Sci.* 2003; 6: 263–9.

Peacock SJ, Richardson JR, Carter R, Edwards D. Priority setting in health care using multi-attribute utility theory and programme budgeting and marginal analysis (PBMA). *Soc Sci Med.* 2007; 64: 897–910.

Pecchia L, Martin JL, Ragozzino A, Vanzanella C, Scognamiglio A, Mirarchi L, Morgan SP. User needs elicitation via analytic hierarchy processs (AHP). A case study on a Computed Tomography (CT) scanner. *BMC Med Inform Decis Mak.* 2013; 13: 2.

Phillips LD, Banae CA. Transparent prioritisation, budgeting and resource allocation with multi-criteria decision analysis and decision conferencing. *Ann Oper Res.* 2007; 154: 51–68.

Phillips LD, Fasolo B, Zafiroopoulos N, Beyer A. Is quantitative benefit risk modeling of drugs desirable or possible? *Drug Discovery Today Technol.* 2011; 8: e3–10.

Radaelli G, Lettieri E, Masella C, et al. Implementation of EUnetHTA core models in Lombardia: the VTS framework. *Int J Technol Assess Health Care.* 2014; 30: 105–12.

Ramli A, Aljunid SM, Sulong S, Md Yusof FA. National drug formulary review of statin therapeutic group using the multiattribute sorting tool. *Ther Clin Risk Manag.* 2013; 9: 491-504.

Schey C, Krabbe PFM, Postma MJ, Connolly MP. Multi-criteria decision analysis (MCDA): testing a proposed MCDA framework for orphan drugs. *Orphanet J Rare Dis.* 2017; 12(10).

Sussex J, Rollet P, Garau M, Schmitt C, Kent A, Hutchings A. Multi-criteria decision analysis to value orphan medicines. Office of Health Economics. 2013; Research paper 13/03.

Thokala P, Devlin N, Marsh K, Baltussen R, Boysen M, Kalo Z, et al. Multiple Criteria Decision Analysis for Health Care Decision Making—An Introduction: Report 1 of the ISPOR MCDA emerging good practices task force. *Value Health.* 2016; 19: 1-13.

Thokala P, Duenas A. Multiple criteria decision analysis for health technology assessment. *Value Health.* 2012; 15(8): 1172-81.

Tony M, Wagner M, Khoury H, Rindress D, Papastavros T, Oh P, Goetghebeur MM. Bridging health technology assessment (HTA) with multicriteria decision analyses (MCDA): field testing of the EVIDEM framework for coverage decisions by a public payer in Canada. *BMC Health Serv Res.* 2011; 11: 329.

US Department of Health and Human Services. Health, United States with special feature on Emergency Care. Centers for Disease Control and Prevention. 2012. Available from: <https://www.cdc.gov/nchs/data/hus/hus12.pdf>.

van Til J, Groothuis-Oudshoorn C, Lieferink M, Dolan J, Goetghebeur M. Does technique matter? A pilot study exploring weighting techniques for a multi-criteria decision support framework. *Cost Eff Resou Alloc.* 2014; 12: 22.

van Til JA, Renzenbrink GJ, Dolan JG, Ijzerman MJ. The use of the analytic hierarchy process to aid decision making in acquired equinovarus deformity. *Arch Phys Med Rehabil.* 2008; 89: 457–62.

van Wijk BLG, Klungel OH, Heerdink ER, de Boer A. A Comparison of Two Multiple Characteristic Decision-Making Models for the Comparison of Antihypertensive Drug Classes Simple Additive Weighting (SAW) and Technique for Order Preference by Similarity to an Ideal Solution (TOPSIS). *Am J Cardiovasc Drugs*. 2006; 6(4): 251–8.

von Winterfeldt D, Fasolo B. Structuring decision problems: a case study and reflections for practitioners. *Eur J Oper Res*. 2009; 199: 857–66.

Youngkong S, Baltussen R, Tantivess S, et al. Multicriteria decision analysis for including health interventions in the universal health coverage benefit package in Thailand. *Value Health*. 2012; 15: 961–70.

Youngkong S, Teerawattananon Y, Tantivess S, Baltussen R. Multi-criteria decision analysis for setting priorities on HIV/AIDS interventions in Thailand. *Health Res Policy Syst*. 2012; 10: 6.

3rd Plenary session, ISPOR 16th Annual European Congress, 2-6 November 2013, Dublin, Ireland. Available from: <http://www.ispor.org/congresses/Dublin1113/presentations/Third-Plenary-Karl-Claxton-Slides.pdf>.



APPLICATION OF MULTI-CRITERIA DECISION ANALYSIS TO PREVALENT DISEASES

Javier Mar Medina

CHAPTER

7

7.1. INTRODUCTION TO MCDA IN PREVALENT DISEASES

The multi-criteria decision analysis can be defined as a systematic approach used to transform decisions taken in an opaque manner into explicit decisions taken through a systematic process based on a sequence of transparent steps (Howard, 1998). The potential applications of MCDA in healthcare decision-making are many and diverse, but can be broadly classified into two categories (Thokala et al., 2016). On the one hand, there is the general application, which is the distribution of resources by assigning priority in accordance with the result of the MCDA. This includes both the reimbursement/authorisation decisions (Danner et al., 2011) and the distribution of a budget (Airoldi et al., 2014). This approach has been the basis of its use in the evaluation of treatments for rare diseases or of orphan drugs (Wagner et al., 2016). And on the other hand, there is the use of MCDA in the classification of therapeutic alternatives for specific clinical problems, either for the preparation of clinical guidelines or to include the preferences of patients in shared decision-making.

This is not an exhaustive list, and MCDA can be useful in many decision contexts (Thokala et al., 2016). Nevertheless, it is useful to place its use in decision-making in the most frequent or prevalent diseases. This term is used exclusively to differentiate the latter from rare diseases (see Chapter 8). Whereas in the assessment of new treatments for rare diseases, the economic evaluation framework is questioned, and MCDA is proposed as an alternative (Wagner et al., 2016; Gilabert-Perramon et al., 2017), in prevalent diseases its use appears as complementary or at least not expressly substitutive (Thokala et al., 2016; Marsh et al., 2016). Garrison, describes the current situation as a turning point in the process of developing clinical practice guidelines, adding other criteria to cost-effectiveness in a decision-making process which is shared through MCDA (Garrison, 2016). Following this approach, the drug or therapeutic option which, according to the various criteria, best responds to the clinical problem posed is selected by taking into account several criteria jointly, such as risk or clinical benefit (Adunlin et al., 2015).

In this chapter we shall review the application of MCDA to prevalent diseases. The methodology used is the review and analysis of the bibliography related to the application of MCDA in frequent diseases, as detailed in the following section.

7.2. METHODS

In the search for practical applications, two strategies were used. First, a review of the literature was carried out in PubMed, associating the terms MCDA or multi-criteria with drugs, cancer, diabetes mellitus, COPD, stroke, infarction, hypertension and heart failure. The diseases were selected according to their prevalence and the burden they cause. The search took place in September, 2017. The search strategies are described in table 7.1. Except in the case of cancer, the number of bibliographical references obtained was small, which reinforces the assessment that MCDA is still at an early stage of application both in management and in clinical settings.

The second strategy consisted of searching for practical examples in books about MCDA. Although there are several manuals for the general use of MCDA (Marsh et al., 2017; Department for Communities and Local Government, 2009), only the one written by Marsh et al., 2017, includes MCDA applications both in the process of making decisions to evaluate new technologies and in the risk/benefit analysis of drugs.

TABLE 7.1. RESULT OF SEARCHES FOR ARTICLES ABOUT PREVALENT DISEASES

| DISEASE | SEARCH STRATEGY | No. OF PUBLISHED ARTICLES | |
|-----------------------|--|---------------------------|----------|
| | | Total | Included |
| Cancer | ("neoplasms"[MeSH Terms] OR "neoplasms"[All Fields] OR "oncology"[All Fields]) AND (multi-criteria[All Fields] OR "multi-criteria"[All Fields] OR "MCDA"[All Fields]) | 105 | 8 |
| COPD | ("pulmonary disease, chronic obstructive"[MeSH Terms] OR ("pulmonary"[All Fields] AND "disease"[All Fields] AND "chronic"[All Fields] AND "obstructive"[All Fields]) OR "chronic obstructive pulmonary disease"[All Fields] OR "copd"[All Fields]) AND (multi-criteria[All Fields] OR "multi-criteria"[All Fields] OR "MCDA"[All Fields]) | 2 | 1 |
| Myocardial infarction | ("myocardial infarction"[MeSH Terms] OR ("myocardial"[All Fields] AND "infarction"[All Fields]) OR "myocardial infarction"[All Fields]) AND (multi-criteria[All Fields] OR "multi-criteria"[All Fields] OR "MCDA"[All Fields]) | 3 | 1 |
| Stroke | ("stroke"[MeSH Terms] OR "stroke"[All Fields]) AND (multi-criteria[All Fields] OR "multi-criteria"[All Fields] OR "MCDA"[All Fields]) | 4 | 3 |
| Hypertension | ("hypertension"[MeSH Terms] OR "hypertension"[All Fields]) AND (multi-criteria[All Fields] OR "multi-criteria"[All Fields] OR "MCDA"[All Fields]) | 6 | 0 |
| Diabetes mellitus | ("diabetes mellitus"[MeSH Terms] OR ("diabetes"[All Fields] AND "mellitus"[All Fields]) OR "diabetes mellitus"[All Fields] OR "diabetes"[All Fields] OR "diabetes insipidus"[MeSH Terms] OR ("diabetes"[All Fields] AND "insipidus"[All Fields]) OR "diabetes insipidus"[All Fields]) AND (multi-criteria[All Fields] OR "multi-criteria"[All Fields] OR "MCDA"[All Fields]) | 6 | 1 |

Source: author's preparation.

7.3. RESULTS OF THE LITERATURE REVIEW ABOUT THE USE OF MCDA IN PREVALENT DISEASES

7.3.1. MCDA in oncology

The search related to oncology was the one which found the largest number of identified articles, with 105 references. Most of them consisted of applications of radiotherapy oncology (72.4%) and were aimed at improving the planning and administering of radiotherapy, which are not the subject of this book. Works focused on proposed methods appeared in 12 cases that were not included in this chapter because they are the subject of Chapters 3, 4, 5 and 6. Another 5 articles were not related to the use of MCDA in oncology. Finally, 7 articles relating to the process of drug selection and five articles analysing preventive practices in cancer, using MCDA, were reviewed. An outstanding article is the review carried out by Adunlin et al., in 2014, in which they analysed the application of MCDA in oncology based on a review of the literature published up to August 31, 2013 (Adunlin et al., 2015). As a result of the examination, only eight articles about methodology or applications were identified. At this stage, the work was aimed at showing the use of MCDA in shared decision-making processes related to oncological treatments. From the results of their study, they deduced that there is a lack of studies of practical application related to cancer. Their conclusion was that MCDA is a promising tool for choosing the alternative with the best balance between benefit and harm, and therefore for informing decision-making in oncology.

Several recently published articles discuss the concept of value in the selection of oncological treatments based on the proposal of value assessment frameworks by institutions such as the Memorial Sloan Kettering Cancer Center, the American Society of Clinical Oncology (ASCO), the National Comprehensive Cancer Network, the Institute for Clinical and Economic Review, and the European Society of Medical Oncology (ESMO) (Schnipper et al., 2015; Institute for Clinical and Economic Review, 2016; Yeung et al., 2017; Carlson and Jonasch, 2016; Cherny et al., 2015; Waldeck et al., 2017). Although they do not formally apply an MCDA methodology, all of the frameworks use several clinical and economic criteria to measure the value that each treatment provides for cancer. In this way, they also propose to include the opinion of the patient among the dimensions to be taken into account. When analysing the ASCO proposal, Waldeck et al., 2017, highlight several relevant points. First, they introduce the debate about the concept of value in microeconomic terms. They then underline the importance of the economic perspective adopted and the difference between the use of absolute and relative measurements to measure health outcomes. A novel aspect is the inclusion of the patient's opinion when comparing results of conventional treatments with those of immune-oncological treatments because of the long tail demonstrated by the latter in the survival curves. In this case, MCDA provides the patient's opinion when comparing better short-term results (conventional treatment) with better long-term results (immune-oncological treatment). Finally, they assess the inclusion of safety considerations in the quality-of-life measurements as a criterion with more weight for providing a more relevant dimension. Malone et al., 2016, also recognise, in general, the value of these new frameworks in addressing the challenges which doctors and patients face when deciding on new treatments whose costs are so great that they threaten the sustainability of cancer treatments. At the same time, they comment critically on the lack of explicit adoption of the ASCO framework for quality-adjusted life years (QALY), because of its difficulty in capturing all the relevant attributes of health outcomes, and on the lack of consensus on the subject of efficiency thresholds.

As a practical application of use of MCDA in oncology, we must highlight the article published by Wagner, which adapts the EVIDEM model of MCDA to three countries (France, Italy and Spain) in order to measure, in a global way, the value of lenvatinib in a specific type of differentiated thyroid cancer which is resistant to conventional treatment with radioactive iodine (Wagner et al., 2017). Lenvatinib was a new drug, not yet approved by the European Medicines Agency (EMA), and which in clinical trials had shown efficacy at the cost of greater side effects. The authors apply the EVIDEM model with 12 quantitative and 7 qualitative criteria to compare treatment with lenvatinib and a watchful, waiting attitude, with treatment with sorafenib. The latter is the only drug approved by the EMA for this type of thyroid cancer, resistant to radioactive iodine. The different components of the MCDA were appraised by panels of appropriate experts in each country. The weights of the criteria at individual and social level were obtained in this way, as were the value scores (from the evidence collected by the systematic review), the qualitative impacts of the contextual criteria and the verbal and written perceptions structured by criteria. Among the different criteria, the comparative efficacy, the quality of the evidence (Spain and Italy) and the severity of the disease (France) received the highest weights. The costs were included among the 12 criteria (intervention costs and other costs) but their weights were small standing only for 14.1% in France, 14.6% in Italy and 12.7% in Spain. The criteria relating to effectiveness were those that contributed more positively to the value of lenvatinib, whereas cost and safety penalised it. The overall quantitative assessment

of the MCDA found very similar overall benefits of lenvatinib in the three countries from the watchful waiting attitude and from treatment with sorafenib. Although the value of lenvatinib was consistently positive in various therapeutic contexts, the added value of the MCDA allowed one to identify explicitly the aspects that contributed most to this result. The authors do not explain why cost-effectiveness and budgetary impact were not included as criteria, but simply the comparative cost. It is possible that this reflects the low weight of economic evaluation in the appraisal of health technologies in southern Europe. Nevertheless, in another application of the EVIDEM model to compare tramadol with alternative treatments for non-oncological chronic pain, both cost-effectiveness and budgetary impact were included (Tony et al., 2011).

An example of the application of MCDA to inform interventions in cancer prevention can be found in the article by Hummel et al., 2013, which addresses the subject of the adherence of the population to the different techniques for screening for colorectal cancer in the Netherlands. To do this, they analysed how, depending on the screening technique used, the benefits and perceived risks influence the decision to participate through the analytic hierarchy process. Basically, the criteria included were diagnostic accuracy and side effects, and the study measured the preferences of the individuals and their impact on the decision to participate. A questionnaire based on a website was used to obtain the preferences of the target population. This consisted of Dutch men and women aged 55 to 75 years who weighted the relevance of the following attributes of the screening techniques: sensitivity, specificity, safety, inconvenience for the patient, and frequency of the test in four different screening techniques. The screening tests studied were the immunochemical fecal occult blood test (iFOBT), colonoscopy, sigmoidoscopy and colonography by computed tomography (CT). A five-point ordinal scale was used to gauge the intentions of the respondents to attend the screening, which were then correlated with the preferences for the screening techniques. The preferred screening method was CT colonography. Sensitivity (weight= 0.26) and safety (weight= 0.26) were the most important determinants of general preferences for the different screening techniques. However, the detection test with the greatest intention of attending was the iFOBT, and the reasons for that were the inconvenience, safety and frequency of the test. The MCDA showed the attributes of the screening techniques which should be taken into account in the Netherlands to increase adherence to the colorectal cancer screening programme.

7.3.2. MCDA in COPD

The search relating to COPD identified two articles. One of them was not related to the use of MCDA, so it was excluded (Monteiro et al., 2012). The article by Marsh et al., 2017, deals with the comparison of two pharmacological treatments of the same therapeutic group (aclidinium and tiotropium) in chronic obstructive pulmonary disease. The authors point out that the evaluation of these treatments usually focuses on measuring their effectiveness in avoiding exacerbations, without taking into account other aspects such as side effects and ease of use. So they apply the MCDA to assess the preferences of patients for each treatment, taking into account different attributes of the treatment (Table 7.2). Although it was intended to gather the perspective of the patients, the participants in the study were six clinicians who, based on data from the literature, identified the criteria and weighted them. In addition, they also assigned scores to the criteria for each alternative in an attempt to reflect the preferences of their patients.

TABLE 7.2. CRITERIA APPLIED IN THE APPRAISAL OF TREATMENTS FOR COPD USING MCDA

| CATEGORY | CRITERION |
|---------------|--|
| Effectiveness | Rescue medication use Symptom severity Early morning activity limitation Night-time awakening |
| Convenience | Exacerbations Confirmation of dose delivered Portability, preloading Dosing per day |
| Safety | Anticholinergic side effects Cardiovascular side effect Other side effects |

Source: Marsh et al. 2017.

Twelve criteria were identified, covering the clinical effectiveness, safety and convenience of COPD treatments. The exacerbations and preloading of the device were identified as the most important for the patients, while the use of rescue medications was the least important. The greater clinical effectiveness of tiotropium was counterbalanced by the better performance of aclidinium in safety and convenience. The overall result was that the net impact of the benefits over the risks of aclidinium (38.5) exceeded that of tiotropium (13.2). The authors themselves conclude that, based on the perceptions of physicians about the preferences of patients, MCDA suggests that aclidinium and tiotropium generate a similar value, considering the overall clinical benefits and risks. However, the tolerability and safety profile of aclidinium is better due to its lower anticholinergic side effects, combined with its greater ease of pre-loading and confirmation of dose administration. The authors conclude that these advantages outweigh the greater benefit of tiotropium in the prevention of exacerbations. The scenario and the sensitivity analyses suggest that the conclusion of the model is not sensitive to the structural and parameter uncertainty, since the probabilistic sensitivity analysis determined that patients preferred aclidinium in 80% of the simulations (Briggs et al., 2012).

The authors themselves point out two limitations of the study. In the first place, in order to know the preferences of the patients, the opinions of the clinicians were collected, whereas it would have been more appropriate to collect them directly from the patients, and not from the clinicians. The authors justify themselves by pointing out the pilot nature of their work. Nevertheless, it is a methodological aspect which greatly limits the relevance of their conclusions. In addition, it may be the origin of an unreasonable criteria weighting result. The clinicians gave the greatest weight to the criteria related to exacerbations and pre-loading of the device. The conclusion that avoiding an event which could increase the risk of death is only slightly more important than the pre-loading of the device does not seem reasonable. The consequence is that the MCDA places as the treatment of choice the one that is most convenient, even when in terms of effectiveness (exacerbations) there are no differences and in terms of safety (side effects) its profile is worse. One dimension totally absent from the criteria used is the cost, despite the fact that the price of aclidinium is higher than that of tiotropium. The greater convenience is accompanied by a greater cost, which could restrict its use. There is also no distinction between different patient profiles, since it is possible that for some of them the differ-

ences in convenience are irrelevant and co-payment represents a problem. Despite these limitations, it is a work which enables us to understand the process to be followed and the elements which must be incorporated into the use of MCDA to inform the clinical use of different treatments for COPD and its potential to integrate different dimensions of the decision into a single unit of measurement.

7.3.3. MCDA in myocardial infarction

Three articles about myocardial infarction were identified in the search, of which one was a generic study of a discrete choice experiment (Dolan et al., 2015), another was an MCDA in atrial fibrillation - which will be discussed in the section about stroke (Dogliotti and Giugliano, 2015) - and the third dealt with a study of the risk/benefit balance in the treatment with statins of myocardial infarction in Korea (Byun et al., 2016).

The study by Byun, analyses three statins (atorvastatin, cerivastatin and simvastatin) and placebo, comparing the risk/benefit balance at the time of approval with the evaluation carried out some time later, taking into account more available information. From a practical point of view, the situation had changed because cerivastatin had been withdrawn from the market due to muscle toxicity (rhabdomyolysis) and the authors wanted to assess whether the change in the risk/benefit balance justified that withdrawal. Based on the literature, they identified two criteria of benefit as the reduction in the plasma of LDL-cholesterol and the incidence of myocardial infarction, and two criteria of risk as hepatotoxicity and fatal rhabdomyolysis. To calculate the weights of the criteria, a discrete choice experiment was carried out in the Korean population. Finally, the scores in each criterion were based on a mixed treatment comparison from the literature. The combination of weights and scores showed that among the criteria of benefit LDL-cholesterol was the most important, with a weight of 37.50%, while the incidence of myocardial infarction reached a weight of 35.43%. Aspects related to safety were less valued, since hepatotoxicity obtained a weight of 16.28%, and rhabdomyolysis only 10.79%. Atorvastatin was the treatment which patients preferred in the overall result. Cervastatin and simvastatin changed their order after approval. Owing to new information about the additional risk of muscular type, cervastatin became the lowest rated option. Notable among the limitations of the study is the fact that the MCDA was carried out after the decision had been made. More than informing the process, its results helped to understand and corroborate the decision. Nevertheless, this work is interesting because the methodology applied was appropriate, the work was performed rigorously and the explanation in the section on methods is complete. In this way, criteria of both effectiveness and safety are incorporated into a single final measurement which compares the value contributed by different alternative treatments.

7.3.4. MCDA in stroke

On reviewing the literature, four studies were found, of which three analysed by means of MCDA the anticoagulant treatments in atrial fibrillation for the prevention of stroke (Dogliotti et al., 2015; Hsu et al., 2015; Hallgreen et al., 2014). The fourth reviews the approaches to the economic evaluation of stroke rehabilitation (Craig et al., 2014). Owing to the difficulty of measuring results such as rehabilitation, this latter study concludes by recommending the use of approaches such as MCDA, which are more comprehensive than cost-effectiveness analysis, to be able to make decisions in complex interventions. The limitation of the study is that it does not carry out any practical application.

Owing to the need to jointly assess the benefit of stroke prevention and the risk of bleeding which it entails, anticoagulant treatment has been the usual subject of decision-analysis methods such as economic appraisal (Teng et al., 2000) and, now, MCDA. The three studies identified apply a risk/benefit analysis to guide clinical practice in the treatment of atrial fibrillation. For Hallgreen et al., 2014, the comparison of warfarin with placebo is a case study which helps them to apply the BRAT framework methodology (Coplan et al., 2011). The results of the study indicate that the risk/benefit balance is favourable to warfarin when compared with placebo in the primary prevention of stroke in patients with atrial fibrillation. It is not really a clinical problem which poses uncertainty, since when compared with placebo, anticoagulant treatment always appears as the alternative of choice for the prevention of stroke. The interest of the article lies in the fact that its authors are members of the Protect Consortium group which carried out the IMI-PROTECT Benefit-Risk Integration and Representation project, coordinated by the EMA and financed by the Seventh Framework Programme of the European Union. This project is the expression of the interest of a regulatory agency such as the EMA in MCDA. Nevertheless, despite the support of the EMA, the BRAT tool has not become standard in MCDA.

The article by Hsu et al., 2015, assessed quantitatively, using MCDA, the effectiveness and safety of the new oral anticoagulants (dabigatran, rivaroxaban and apixaban) compared with those of warfarin in the treatment of non-valvular atrial fibrillation, under different risk scenarios. To do this, they used a methodology based on decision trees which include the different criteria as branches and finally combines them into a single measure of value. The incorporation of different risk scenarios, such as the general population, patients with a higher risk of stroke, and primary and secondary prevention of stroke, is of great interest, because it allows stratified measurement of the value of different drugs according to the risk scenario. It should be emphasised that in each scenario not only were the scores of the alternatives different, but the weightings of the criteria also varied. Normally, the new anticoagulants had higher value scores than warfarin, with similar results among them and with different results depending on the risk scenario. Thus, in the basic scenario dabigatran obtained the highest score (0.529), while in the population at risk it was apixaban which reached the highest yield value (0.686). Warfarin was always lower both in the general population (0.191) and in the population at risk (0.116). Although the comparison of the utility values obtained in the study is overwhelmingly favourable to the new anticoagulants, the authors' conclusion that they might be preferred to warfarin was timid. Possibly, they are aware that there are other criteria, such as cost and adherence, which have not been included in the MCDA. It is possible that those criteria not incorporated in the MCDA prove that warfarin is still alive as a treatment despite its disadvantages in terms of control and administration. Although the authors do not quote them, they take them into account when formulating the conclusion of the study.

Doglioti et al., 2015, analysed by MCDA the risk/benefit balance of the various kinds of treatment for atrial fibrillation, including both anticoagulants and platelet anti-aggregants. First, they carried out a network meta-analysis to obtain comparative results for the multiple alternatives analysed. Secondly, they used the SMAA-2 methodology to carry out an analysis of stochastic multi-criteria acceptability (Tervonen and Lahdelma, 2007). Applying the tool, they performed two analyses. The first had as a measure of the result a utility function based on a multi-attribute preference function, and the second included the healthcare costs of the complications, but not the cost of the treatment. The result is clearly favourable to the new anticoagulants, both in utility and in cost. The authors point out some of the limitations of the new drugs, such as the price and the lack of an antidote in case of

severe bleeding. If the objective were only the clinical risk/benefit balance, excluding the cost of the treatments could be justified. Nevertheless, quantifying the cost dimension by taking into account only the costs of the complications is difficult to justify, since it forgets the price and skews the result in favour of the new anticoagulants. Another weakness of this work is the lack of transparency in the presentation of the weighting of the criteria of effectiveness and safety. Although the SMAA-2 model has a relevant scientific trajectory, for the purposes of this work it is a black box in which we have to trust. It seems clear that the explicit presentation of the weight of each criterion is an element of good practice in the MCDA studies (Marsh et al., 2016).

7.3.5. MCDA in hypertension

Although six articles were found in the search, none of them contained any exercise which applied MCDA to the treatment of hypertension, so they are not included in the literature commented on. One of them showed a framework for the use of MCDA in rare diseases and three articles related to the classification of systemic sclerosis.

7.3.6. MCDA in diabetes

Of the six articles identified, the one by Mehrotra and Kim, 2011, was selected because it falls within the profile of applications in prevalent diseases. The other five articles did not have the required profile. The authors are researchers from an engineering school who apply MCDA to carry out a budget allocation in diabetes prevention programmes at the CDC in the United States, from the perspective of the decision-makers. They justify the use of MCDA because it is a multi-objective problem in which the parameters of prevalence, co-morbidities and mortality must be considered simultaneously, and no single criterion serves as a substitute for the others. They incorporate 21 clinical and social criteria (index of inequality) which they analyse using the methodology of the analysis of main components and the techniques of reverse linear programming. The exercise enables them to conclude that the resources of the CDC should be allocated to promote education about diabetes and to increase interactions between the patient and the healthcare provider. The detailed description of the quantitative methodology occupies most of the article, making it difficult to identify the criteria and their weights. The alternatives of the budget are not well described either, because, although the conclusions indicate a priority for specific interventions, the budget is distributed by the states.

7.3.7. Review of books

The book by Marsh et al., 2017, contains a chapter devoted to explaining three examples of the inclusion of MCDA and the evaluation of technologies for making financing decisions. In the world, various organisations with responsibility for setting priorities and allocating resources have explicitly proposed the use of MCDA. The examples from Colombia, Lombardy and Belgium are presented in detail in section 4.3 of this book.

In Colombia in 2002 and 2003 the Health Regulation Commission [*Comisión de Regulación en Salud (CRES)*] carried out a pilot application of MCDA using the EVIDEM method (Marsh et al., 2017; CRES, 2012; Castro et al., 2016). Four technologies were selected for which no reimbursement decisions had been made and for which technology evaluations carried out in Colombia were available. The decision to be taken by the focus group involved ordering, in relation to a possible reimbursement, the primary prophylaxis of haemophilia A, the

supply of zinc for the prevention of diarrhoea, the first-line treatment with anastrozole for postmenopausal women with metastatic breast cancer and positive hormone receptors, and the use of ticagrelor combined with aspirin in patients with acute coronary syndrome without ST elevation. A focus group was set up with decision-makers and experts to carry out the MCDA. They first weighted the different EVIDEM criteria and subsequently scored the four technologies in each criterion. The criteria which obtained greater weight were the severity of the disease (9.3), the size of the population (8.9) and the incremental effectiveness (8.7). By contrast, the attention to differential needs (4.3) and the validity of the evidence (5.0) were the least important criteria. The final result was that the supply of zinc obtained the highest score (0.904), followed by the use of anastrozole (0.822), the prophylaxis of haemophilia A (0.794) and lastly, the ticagrelor (0.708). The final consensus of the group was that a mixed method combining the MCDA and the budget impact analysis would be ideal for Colombia.

In the Italian region of Lombardy, MCDA was incorporated into a more general policy formulated in 2008 and aimed at improving the decision-making process related to health technologies. This enabled to define a specific framework that was applied in the evaluation of technologies from 2012 onwards, taking account of the methods developed by the European network for HTA (EUnetHTA) and the EVIDEM collaboration (Migliore et al., 2014). As a result of its use, each technology receives a report containing the index of priority or appropriate use and the qualitative analysis for each criterion. Based on this report, the committee rejects or accepts each technology. Most of the proposed technologies have been rejected. In some cases, the positive evaluation was accompanied by restrictions such as its exclusive use in specific centres, procedures for selecting patients and joint risk contracts (Regione de Lombardia, 2017). Of the list shown in the book (16 appraisals), only two were accepted without restrictions, two with restrictions, and the remaining 12 were rejected (Table 7.2). The same approach of the MCDA developed in Lombardy is currently being explored by other regions in Italy. The use of MCDA has recently been incorporated into a national law for the establishment of priorities in the area of medical devices (Marsh et al., 2017).

In Belgium, the use of MCDA by the Drug Reimbursement Committee has been proposed to decide about applications from pharmaceutical companies. In 2010, the Belgian Healthcare Knowledge Centre (KCE) proposed a possible framework for the application of MCDA (Le Polain et al., 2010). In contrast to the examples described in the literature, the Belgian framework of MCDA states that the weights of the criteria must come from the general public, because decision-making about the reimbursement of healthcare should be based on social preferences. Because these preferences for reimbursement criteria were not known, in 2014 the KCE conducted a large survey among the population to obtain these weights (see section 4.3). The reason is that, in a democratic system, the value judgements incorporated in the evaluation should ideally reflect social values and preferences. So it was recommended that a strict distinction be made between assessment and appraisal. Assessment involves the collection of evidence relating to the technology under consideration. Appraisal involves value judgments to determine the relative importance of each of the elements being evaluated. The result is that, unlike many other MCDA approaches, the Belgian approach does not combine the need for a new better treatment with the added value of new drugs in a single weighted score. The framework described is not yet being applied in practice, but it will be used in the context of early decisions about the reimbursement of drugs.

TABLE 7.3. LIST OF TECHNOLOGIES ASSESSED IN LOMBARDY AND RESULT OF THE EVALUATION

| AREA | TECHNOLOGY | DISEASE | RESULT |
|-----------|------------------------------------|----------------------------|------------------------|
| Devices | Therapy for psoriasis | Psoriasis | Rejected |
| | Vagal estimation | Resistant epilepsy | Conditional acceptance |
| | Extracorporeal shock wave therapy | Chronic wounds | Rejected |
| | Cochlear BAHA | Deafness | Conditional acceptance |
| | Intraocular lenses | Cataracts | Rejected |
| | LINX | Gastroesophageal reflux | Rejected |
| | Diaphragmatic stimulation | Severe tetraplegia | Accepted |
| Diagnoses | Optical coherence tomography | Coronary disease | Rejected |
| | Raman spectroscopy | Skin cancer | Rejected |
| | Presepsin | Sepsis | Rejected |
| | Wireless coronary guide for stents | Coronary disease | Rejected |
| | Laser endomicroscopy | Bladder cancer | Accepted |
| | Brevagen | Breast cancer | Rejected |
| | IMS monitoring service | Adherence to medication | Rejected |
| | Comfortscan | Breast cancer | Rejected |
| Radiology | 7 teslas NMR | Neurodegenerative diseases | Rejected |

Source: Migliore et al., 2014.

7.4. DISCUSSION

In accordance with the second ISPOR report (Marsh et al., 2016), MCDA aims to cover a wide range of decisions, including the optimisation of the portfolio of services, regulatory authorisation, the Health Technology Assessment (HTA), start-up decisions/priority frameworks, and decision-making in the hospital, prioritising patients' access to treatment and disease classification. However, the review of the literature about its application to prevalent diseases shows that MCDA has been used at clinical level in shared decision-making and to measure the risk/benefit balance. Except for the examples described in the book by Marsh et al., 2017, there is no literature about its application in the field of authorisation and reimbursement of new treatments for prevalent diseases. So the lack of experiences in the field of healthcare services forces Marsh et al., 2016, to use in their report methods and examples from the world of MCDA in general, and not only in the field of healthcare. In the field of reimbursement, authors such as Gilabert-Perramon state that the use of MCDA by the Catalan Healthcare Service is explicitly restricted to orphan drugs (Gilabert-Perramon et al., 2017). The second ISPOR document about MCDA, previously cited (Marsh et al., 2016), establishes a list of good practices in the use of MCDA, as detailed in Chapter 6. If we apply them to the reviewed literature about their use in prevalent diseases, the lack of compliance with them is striking, especially in relation to rigour in the explanation of the intermediate steps followed during the implementation of the study. Although the publication poses problems due to the limitation in the number of words in the manuscripts, technical appendices could be incorporated to reproduce the process followed by the authors. In some cases, the 'black box' effect resulting from the use of a specific piece of software

was particularly striking (Dogliotti and Giugliano, 2015). In the same ISPOR report, they identify the participants involved in an MCDA, and point out several types: decision-makers, stakeholders, analysts and experts. Although the term 'stakeholder' is generic, since it can include several categories of participants, this classification, when applied to the examples analysed, indicates that, basically, the participants identified in this type of analysis are analysts and experts.

The criterion of efficiency is systematically absent from the examples which have been reviewed. The comment made by a group of experts from ISPOR (Malone et al., 2016) about the lack of inclusion of the quality-adjusted life-years produced by each alternative, as a criterion for decision-making within the framework of the ASCO (Schnipper et al., 2015), could also be applied to the examples analysed outside the oncological context. The dimension of cost (budgetary impact) and efficiency (cost-effectiveness) may not be the only one to be taken into account, but the entire body of scientific knowledge about health economics supports its use in decision-making under conditions of allocation of scarce resources (López-Bastida et al., 2010). Thus the creators of the EVIDEM framework include cost-effectiveness and budget impact in an exercise of integration of the assessment of health technologies into the MCDA, in a work of 2011 aimed at testing the method (Tony et al., 2011). However, when they published their application to treatment alternatives in 2017, they excluded these criteria (Hummel et al., 2013). In general, the objectives of these examples are more in line with those formulated by experts such as those nominated by ASCO, who, at the beginning of their document updating the framework, point out that it is aimed at developing a framework that enables doctors and their patients to assess the value of a particular cancer treatment taking into account the patients' individual preferences and circumstances (Schnipper et al., 2015). In this scenario, other participants, such as managers, have little room. As in the economic appraisal, the methods matter and the credibility of the MCDA has to be sustained by the participation of all the interlocutors involved and by an accredited methodological support.

7.5. CONCLUSIONS

As the main conclusion of this chapter, it can be said that MCDA is still beginning its journey as a tool for decision-making in prevalent diseases. The standardisation of its methods will be a key element in its consolidation as a scientific tool, since the transparency of the process and the objectivity in the selection of the criteria are critical points of the process that cannot be left to the subjective selection of those responsible for the study. Despite this, the set of methods included in MCDA can provide a formal framework for assessing treatment alternatives and assigning priorities in decision-making in prevalent diseases in a transparent way and applying different criteria in a quantitative manner. The application of these methods would help to avoid variability and inconsistencies in the decisions of clinicians and managers (Thokala et al., 2016).

REFERENCES

Adunlin G, Diaby V, Montero AJ, Xiao H. Multicriteria decision analysis in oncology. *Health Expect*. 2015;18(6): 1812-26.

Airoldi M, Morton A, Smith JA, Bevan G. STAR--people-powered prioritization: a 21st-century solution to allocation headaches. *Med Decis Making*. 2014;34(8): 965-75.

Briggs AH, Weinstein MC, Fenwick EA, et al. ISPOR-SMDM Modeling Good Research Practices Task Force. Model Parameter Estimation and Uncertainty: A Report of the ISPOR-SMDM Modeling Good Research Practices Task Force-6. *Med Decis Making*. 2012; 32(5): 722-32.

Byun JH, Kwon SH, Ha JH, Lee EK. A benefit-risk assessment model for statins using multicriteria decision analysis based on a discrete choice experiment in Korean patients. *Ther Clin Risk Manag*. 2016 Jun 13; 12: 965-74.

Carlson R, Jonasch E. NCCN Evidence Blocks. *J Natl Compr Canc Netw*. 2016; 14 (5S): 616-19.

Castro Jaramillo HE, Goetghebeur M, Moreno-Mattar O. Testing multi-criteria decision analysis for more transparent resource-allocation decision making in Colombia. *Int J Technol Assess Health Care*. 2016 Jan; 3(4): 307-314.

Cherny N, Sullivan R, Dafni U, et al. A standardised, generic, validated approach to stratify the magnitude of clinical benefit that can be anticipated from anti-cancer therapies: The European Society for Medical Oncology Magnitude of Clinical Benefit Scale (ESMO-MCBS). *Ann Oncol*. 2015; 26(8): 1547-73.

Comisión de Regulación en Salud (CRES) Unidad Administrativa Especial. Metodología para la determinación de los criterios y categorías para la priorización de tecnologías en salud en el proceso de actualización del POS. Documento técnico proyecto POS – UPC 2012 – 2013. Subdirección técnica, Bogotá. 2012. Accessed: november 18, 2017. Available from: http://www.med-informatica.com/OBSERVAMED/EvaluacionesFT/CRESaDRBCyTAS_InformeTecnicoCriteriosPriorizacion2012-12-13.pdf.

Coplan PM, Noel RA, Levitan BS, Ferguson J, Mussen F. Development of a framework for enhancing the transparency, reproducibility and communication of the benefit–risk balance of medicines. *Clin Pharmacol Ther*. 2011; 89(2): 312–315.

Craig LE, Wu O, Bernhardt J, Langhorne P. Approaches to economic evaluations of stroke rehabilitation. *Int J Stroke*. 2014 Jan; 9(1): 88-100.

Danner M, Hummel JM, Volz F, van Manen JG, Wiegand B, Dintsios CM, Bastian H, Gerber A, Ijzerman MJ. Integrating patients' views into health technology assessment: Analytic hierarchy process (AHP) as a method to elicit patient preferences. *Int J Technol Assess Health Care*. 2011; 27(4): 369-75.

Department for Communities and Local Government. Multi-criteria analysis: a manual. 2009. London.

Dogliotti A, Giugliano RP. A novel approach indirectly comparing benefit-risk balance across anti-thrombotic therapies in patients with atrial fibrillation. *Eur Heart J Cardiovasc Pharmacother*. 2015 Jan;1(1): 15-28.

Dolan JG, Cherkasky OA, Chin N, Veazie PJ. Decision Aids: The Effect of Labeling Options on Patient Choices and Decision Making. *Med Decis Making*. 2015 Nov; 35(8): 979-86.

Garrison LP Jr. Cost-Effectiveness and Clinical Practice Guidelines: Have We Reached a Tipping Point?-An Overview. *Value Health*. 2016; 19(5): 512-5.

Gilbert-Perramon A, Torrent-Farnell J, Catalan A, Prat A, Fontanet M, Puig-Peiró R, Merino-Montero S, Khoury H, Goetghebeur MM, Badia X. Drug evaluation and decision making in Catalonia: development and validation of a methodological framework based on multi-criteria decision analysis (MCDA) for orphan drugs. *Int J Technol Assess Health Care*. 2017; 33(1): 111-120

Hallgreen CE, van den Ham HA, Mt-Isa S, Ashworth S, Hermann R, Hobbiger S, Luciani D, Micallef A, Thomson A, Wang N, van Staa TP, Downey G, Hirsch I, Hockley K, Juhaeri J, Metcalf M, Mwangi J, Nixon R, Peters R, Stoeckert I, Waddingham E, Tzoulaki I, Ashby D, Wise L. Benefit-risk assessment in a post-market setting: a case study integrating real-life experience into benefit-risk methodology. *Pharmacoepidemiol Drug Saf.* 2014 Sep; 23(9): 974-83.

Howard RA. Decision analysis: practice and promise. *Management Science*, 1988; 34: 679–695.

Hsu JC, Hsieh CY, Yang YH, Lu CY. Net clinical benefit of oral anticoagulants: a multiple criteria decision analysis. *PLoS One.* 2015 Apr 21; 10(4): e0124806.

Hummel JM, Steuten LG, Groothuis-Oudshoorn CJ, Mulder N, Ijzerman MJ. Preferences for colorectal cancer screening techniques and intention to attend: a multi-criteria decision analysis. *Appl Health Econ Health Policy.* 2013; 11(5): 499-507.

Institute for Clinical and Economic Review. ICER value assessment framework: 1.0 to 2.0. ICER webinar. Accessed: July 29, 2016. Available from: <http://icer-review.org/wp-content/uploads/2016/02/Value-Assessment-Frameworkslides-for-July-29-webinar-FINAL.pdf>.

Le Polain M, Franken M, Koopmanschap M, Cleemput I. 2010. Drug reimbursement systems: international comparison and policy recommendations. Health Services Research (HSR). Belgian Healthcare Knowledge Centre (KCE). KCE reports 147C, Brussels.

López-Bastida J, Oliva J, Antoñanzas F, García-Altés A, Gisbert R, Mar J, Puig-Junoy J. Spanish recommendations on economic evaluation of health technologies. *Eur J Health Econ.* 2010 Oct; 11(5): 513-20.

Malone DC, Berg NS, Claxton K, Garrison LP Jr, IJzerman M, Marsh K, Neumann PJ, Sculpher M, Towse A, Uyl-de Groot C, Weinstein MC. International Society for Pharmacoeconomics and Outcomes Research Comments on the American Society of Clinical Oncology Value Framework. *J Clin Oncol.* 2016; 34(24): 2936-7.

Marsh K, Goetghebeur M, Thokala P, Baltussen R. Multi-Criteria Decision Analysis to support healthcare decisions. Springer. 2017. Cham.

Marsh K, Ijzerman M, Thokala P, Baltussen R, Boysen M, Kaló Z, Lönnngren T, Mussen F, Peacock S, Watkins J, Devlin N; ISPOR Task Force. Multiple Criteria Decision Analysis for Health Care Decision Making—Emerging Good Practices: Report 2 of the ISPOR MCDA Emerging Good Practices Task Force. *Value Health.* 2016; 19(2): 125-37.

Marsh K, Zaiser E, Orfanos P, Salverda S, Wilcox T, Sun S, Dixit S. Evaluation of COPD Treatments: A Multicriteria Decision Analysis of Acclidinium and Tiotropium in the United States. *Value Health.* 2017; 20(1): 132-140.

Mehrotra S, Kim K. Outcome based state budget allocation for diabetes prevention programs using multi-criteria optimization with robust weights. *Health Care Manag Sci.* 2011 Dec; 14(4): 324-37.

Migliore A, Tringali M, Fortino I, Cerbo M. The impact of EAA activities: from assessment to decision-making in Lombardia. *EuroScan Newsletter* 2014; 16: 7.

Monteiro A, Carvalho V, Velho S, Sousa C. Assessing and monitoring urban resilience using COPD in Porto. *Sci Total Environ.* 2012; 414: 113-9.

Regione de Lombardia. Programma regionale di HTA delle tecnologie sanitarie. Available from: <https://htadm-lombardia.ats-pavia.it/>. Accessed el 19 de noviembre de 2017.

Schnipper LE, Davidson NE, Wollins DS, Tyne C, Blayney DW, Blum D et al American Society of Clinical Oncology. American Society of Clinical Oncology Statement: A Conceptual Framework to Assess the Value of Cancer Treatment Options. *J Clin Oncol.* 2015 Aug 10; 33(23): 2563-77.

Teng MP, Catherwood LE, Melby DP. Cost effectiveness of therapies for atrial fibrillation. A review. *Pharmacoeconomics.* 2000 Oct; 18(4): 317-33.

Tervonen T, Lahdelma R. Implementing stochastic multicriteria acceptability analysis. *Eur J Oper Res* 2007; 178: 500–513.

Thokala P, Devlin N, Marsh K, Baltussen R, Boysen M, Kalo Z, Longrenn T, Mussen F, Peacock S, Watkins J, Ijzerman M. Multiple Criteria Decision Analysis for Health Care Decision Making--An Introduction: Report 1 of the ISPOR MCDA Emerging Good Practices Task Force. *Value Health*. 2016; 19(1): 1-13.

Tony M, Wagner M, Khoury H, Rindress D, Papastavros T, Oh P, Goetghebeur MM. Bridging health technology assessment (HTA) with multicriteria decision analyses (MCDA): field testing of the EVIDEM framework for coverage decisions by a public payer in Canada. *BMC Health Serv Res*. 2011; 11: 329.

Wagner M, Khoury H, Bennetts L, Berto P, Ehreth J, Badia X, Goetghebeur M. Appraising the holistic value of Lenvatinib for radio-iodine refractory differentiated thyroid cancer: A multi-country study applying pragmatic MCDA. *BMC Cancer*. 2017; 17(1): 272.

Wagner M, Khoury H, Willet J, Rindress D, Goetghebeur M. Can the EVIDEM Framework Tackle Issues Raised by Evaluating Treatments for Rare Diseases: Analysis of Issues and Policies, and Context-Specific Adaptation. *Pharmacoeconomics*. 2016; 34(3): 285-301.

Waldeck AR, Botteman MF, White RE, van Hout BA. The Importance of Economic Perspective and Quantitative Approaches in Oncology Value Frameworks of Drug Selection and Shared Decision Making. *J Manag Care Spec Pharm*. 2017 Jun; 23(6-a Suppl): S6-S12.

Yeung K, Basu A, Hansen R, Watkins J, Sullivan S. Impact of a value based formulary on medication utilization, health services utilization, and expenditures. *Med Care*. 2017; 55(2): 191-98.



APPLICATION OF MULTI-CRITERIA DECISION ANALYSIS TO ORPHAN DRUGS

Carlos Campillo-Artero

CHAPTER

8

8.1. INTRODUCTION

The low frequency of rare diseases (RDs) and, in many cases, their severity, their progressive, chronic and disabling nature, the scarcity of effective treatments, the uncertainty associated with their evolution and treatments, as well as their medical, psychological, economic and social consequences, make both the assessment of efficacy and safety in authorising Orphan Drugs (ODs) and the decisions about their price, coverage and reimbursement (appraisal) complex and controversial.

Although their harmonisation has been sought for some time, assessment and appraisal do not go hand in hand, and this takes on special relevance and intensity in the sphere of ODs. The regulation of assessment has been made without considering the processes of pricing and reimbursement. Notwithstanding their authorisation, the uncertainty about the effectiveness of ODs, the heterogeneity of methods and the international variability of authorisation decisions and their economic, social and political consequences, their high price and limited budgets, their joint budgetary impact, the pressure to reduce spending and the discrepancies between the modalities of fixing prices and of having the same prices between countries, have unleashed an avalanche of studies and publications in the past decade (Tsoi et al., 2013; Wagner et al., 2016; Hughes-Wilson et al., 2012; Hutchings et al., 2012; Ethgen et al., 2012; Fedyaeva et al., 2014; Kawalec et al., 2016; Drummond et al., 2009; Campillo-Artero, 2014; Drummond et al., 2013; Solon and Kanavos, 2015; Mincarone et al., 2017; Hughes et al., 2005; Kanavos and Nicod, 2012; McCabe et al., 2007).

All of this is further complicated if we take into account that there are ODs with more than one indication, that a drug used for a non-rare disease can be authorised in the second instance to treat an OD (Orphan Drug) and viceversa, that there can be more than one drug to treat the same RD, and that achieving with different manoeuvres the declaration and authorisation of the OD can serve as a fast-entry way of entering the market and offering additional benefits, such as those resulting from market exclusivity with respect to the patent (Hughes-Wilson, et al., 2012; Campillo-Artero and Peiró, 2009).

Numerous authors affirm that the current conventional methods of Health Technology Assessment (HTA), including the economic appraisal – understood here as the fourth barrier – cannot capture, combine or reflect the multiplicity of dimensions, criteria and implications of an ethical, political, social and economic nature which should be considered in the appraisal of ODs nor assign relative weights to them in accordance with their importance for the affected individuals and for society. Decisions about prioritisation and allocation of limited resources require, in addition to good information, people's value judgements and social judgements regarding multiple principles and criteria. Even with the best information, it is necessary to make judgements because information cannot replace them. To adequately combine all relevant variables in appraisal decisions requires deliberation, and the field of ODs is a fertile ground where the inconsistencies of criteria and the contrast of values and positions are intensified (Wagner et al., 2016; Hughes-Wilson et al., 2012; Hutchings et al., 2012; Ethgen et al., 2012; Fedyaeva et al., 2014; Kawalec et al., 2016; Drummond et al., 2009; Drummond et al., 2013; Solon and Kanavos, 2016; Mincarone et al., 2017; Hughes et al., 2005; Kanavos and Nicod, 2012; Wahlster et al., 2015; Angelis and Kanavos, 2016). To what extent the HTA and the economic appraisal in particular, and especially that relating to ODs, manage to incorporate in the reimbursement decisions the crucial components of the deliberation will be discussed in the last sections of this chap-

ter, in the light of all the information which is presented (Culyer and Lomas, 2006; Culyer, 2012; Culyer, 2014; Culyer, 2017).

The MCDA (Multi-Criteria Decision Analysis) is one of the methods that are proposed to overcome some of these pitfalls. The strong point of these proposals is that the MCDA is one of the available methods of a taxonomy of methodological tools for combining or adding multiple criteria, in a single value parameter of the drug, which can help to prioritise alternatives and allocate limited resources (Wagner et al., 2016; Hughes-Wilson et al., 2012; Hutchings et al., 2012; Ethgen et al., 2012; Fedyeva et al., 2014; Paulden et al., 2015; Sussex et al., 2013a; Mincarone et al., 2015; Iskrov et al., 2016; Kolasa et al., 2016; Devlin and Sussex, 2011; Wagner et al., 2017; Garau et al., 2016; Schlander et al., 2014; Schlander et al., 2016; Annemans et al., 2017; Gutierrez et al., 2015; Drummond et al., 2007; Solon and Kanavos, 2016; Mincarone et al., 2017; Hughes et al., 2005; Culyer, 2014).

The general purpose of this chapter is to describe the use of MCDA in the evaluation of ODs. In addition, as more specific purposes, the advances which MCDA has achieved to date in the appraisal of ODs, together with its limitations, are analysed with a critical look at the arguments used to justify its use, and perspectives are drawn from MCDA in the short and medium term in the field of ODs and RDs.

To produce it, the bibliography was reviewed to find: 1) applications of MCDA in the appraisal of ODs, 2) information about methodological aspects of MCDA, and 3) recommendations and consensus documents about the role and utility of MCDA in the context of the appraisal of medical technologies for treating RDs. Use was also made of published guides and references which offer criteria for assessing the quality of these analyses and recommendations for improving their notification: the two ISPOR reports about MCDA methods and their potential for supporting decisions in health care (Thokala et al., 2016; Marsh et al., 2016), the general description of the EVIDEM framework (Goetghebeur et al., 2008) and its versions 3.0 and 3.1 (The EVIDEM Collaboration, 2015), the CHEERS guide about standards of notification of economic appraisals (Husereau et al., 2013), a Spanish correlate (López-Bastida et al., 2010), and a specific one for ODs in Europe (Schuller et al., 2015). Finally, published articles and documents were used in which relevant aspects of regulation (assessment and appraisal), and especially of HTA and economic evaluation, as well as of health economics in the fields of RDs and ODs, are analysed.

8.2. REASONS FOR USING MCDA IN DECISIONS ABOUT PRICE AND REIMBURSEMENT OF ORPHAN DRUGS

Several arguments are advanced about the limitations of HTA in estimating and making decisions about the value of ODs, and about their coverage and reimbursement. This section is included because, based on these arguments, some propose MCDA as a complement to HTA, and others, formulating more severe criticisms of the limitations indicated, as a candidate to replace them and establish itself as a new paradigm of appraisal (Clarke et al., 2009; Winquist et al., 2012; Sussex, et al., 2013c; Mincarone et al., 2017; Iskrov et al., 2016; Kolasa et al., 2016; Schey and Conolly, 2014; Schey et al., 2017; Schlander et al., 2014; Schlander et al., 2016; Gutierrez et al., 2015; Hughes et al., 2005; Lloyd et al., 2015; Drummond et al., 2007; Drummond et al., 2009; Drummond et al., 2013). For the sake of brevity, and because of their relevance, these arguments are summarised in table 8.1, and in a later section are compared

with competing arguments from other authors, especially in those aspects where the coefficient of friction between positions is higher.

TABLE 8.1. ARGUMENTS ABOUT THE LIMITATIONS OF HTA AND ECONOMIC EVALUATION IN MAKING DECISIONS ABOUT THE VALUE OF ORPHAN DRUGS AND ABOUT THEIR COVERAGE AND REIMBURSEMENT

| | |
|---|--|
| In the set of studies reviewed it is stated that, for all the reasons summarised below, the wide range of normative and positive elements which underlie the extra-welfarist logic of cost-effectiveness for prioritising interventions in situations of limited resources (as a criterion of allocative efficiency) should be reviewed in depth. | |
| 1 | Rarity, unmet medical needs, seriousness and endangerment of life are criteria which per se confer a status or special consideration on RDs. Unlike the MCDA, neither the HTA, nor the economic evaluation in particular, consider or adequately combine the opinion and preferences of all the stakeholders who should participate in decisions about authorisation, coverage and reimbursement. |
| 2 | HTA and economic evaluation omit these and other criteria, and do not add all the elements which define the value of a new drug, and in particular those of the ODs. Economic evaluation, based on utilitarian and extra-welfarist principles, by denying coverage and reimbursement leads to the abandonment of these patients and to inequity. Because of their peculiarities, ODs cannot exceed the usual regulatory standards of efficacy, safety, quality and cost-effectiveness, and their maintenance is unfair and inequitable for patients suffering from these diseases. |
| 3 | Social value does not equate to utility as it is understood in economic evaluation. QALYs do not express social preferences or the set of values that translate the social value of ODs and RDs. Their value is not the same in all patients, those of patients are not consistent with those of the general public, and the measurement scales do not adequately reflect the patients' experience, nor are they weighted according to variables such as the prevalence of the disease, life expectancy or severity. Not all stakeholders accord equal importance to economic valuation criteria. Accepting the distributive neutrality of QALYs limits the expression of social and patients' preferences. |
| 4 | The origin and use of an incremental cost-effectiveness ratio (ICER) threshold is arbitrary and inadequate and leads to unfair decisions, because its essence depends on the validity of the QALY maximisation hypothesis, a weak assumption because it does not capture social preferences or benefits, which transcend the principle of quasi-utilitarian maximization, which, by design, is blind to distributive considerations. If other objectives, in addition to the maximisation of health, were incorporated, different ICERs would be obtained. The uncertainty associated with estimates of the efficacy and safety of ODs can render the ICER threshold meaningless. By using only costs, efficacy and the ICER threshold, potentially beneficial therapeutic alternatives are excluded. The incremental cost-value ratio (ICVR) is proposed as an alternative to the ICER. |
| 5 | Value-based pricing models omit the genuine dimensions of the value of ODs and RDs. The results of the use of these models and of the economic evaluation with ODs, together with the coverage and reimbursement decisions adopted through their application, vary among countries and impair equity and access. |
| 6 | HTA, and specifically economic evaluation, are not transparent methods. |

HTA: Health Technology Assessment. ICER: Incremental Cost-Effectiveness Ratio.

Source: author's preparation from Clarke et al., 2009. Winquist et al., 2012. Sussex et al., 2013c. Mincarone et al., 2017. Iskov et al., 2016. Kolasa et al., 2016. Schey and Conolly, 2014. Schey et al., 2017. Schlander et al., 2014. Schlander et al., 2016. Gutierrez et al., 2015. Hughes et al., 2005. Lloyd et al., 2015. Drummond et al., 2007. Drummond et al., 2009. Drummond et al., 2013. Wahlser et al., 2015. Angelis and Kanavos, 2016.

8.3. MODELS OF COMBINATION OF CRITERIA APPLIED TO ORPHAN DRUGS

In the published studies there are two groups of models which combine multiple criteria, specifically in the appraisal of ODs. The first consists of theoretical models of MCDA, within which one can distinguish applications of these models with simulated ODs and applications under real

conditions of assessment with ODs and real data. The second group is made up of combination models of multiple criteria different from the MCDA, which have also been applied to ODs.

8.3.1. Theoretical MCDA models

The first one consists of the adaptation of the EVIDEM framework to incorporate in the evaluation of ODs the factors derived from the singularities of RDs and of these drugs (Wagner et al., 2016). By means of a literature review, and taking into account a wide range of ethical principles, as well as the foundations of the design of the MCDA, the factors which affect the appraisal were identified, as well as the current policies or proposals for the regulation of ODs in general, and for their reimbursement in particular. Each of the identified factors and criteria was compared with those of the third version of the EVIDEM framework to decide whether they needed to be modified in order to increase their usefulness in the appraisal of these medicines with respect to that general framework.

Table 8.2 shows this EVIDEM framework adapted to ODs. In essence, the adapted framework integrates social and individual values, and combines competing ethical dilemmas. Its added value depends on the following modifications: the addition of sub-criteria to the criteria of disease severity, comparative effectiveness, comparative cost-consequence relationship, and cost and consequences/non-medical costs relationship, the incorporation of negative values in scales for all the comparative criteria, and the incorporation of a method to add to the framework specific policies and priorities of the context where it will be used.

In addition, in this new framework, attention is paid to the uncertainty which is especially relevant in the field of RDs and ODs. The degree of uncertainty in the judgements that are pronounced about the evidence can be evaluated using ranges in the scoring scales; that about the weights, by using different weighting techniques from among those available; that which is due to the variability of individual perspectives and judgements, by means of dispersion measurements; and that which is attributable to the structure of the model, by modifying its structure, for example, by eliminating the weighted criterion of lesser magnitude.

The peculiarity of the second model is the evaluation criteria which it includes, because some of them are not considered in the other models which combine multiple proposed criteria (Hughes-Wilson et al., 2012). As shown in table 8.2, this model does not include the same criteria as the previous one and, as a particularity, it combines new criteria, such as the complexity of the OD's manufacture, additional post-authorisation measures to gather more information about effectiveness, severity, available therapeutic alternatives/unsatisfied medical needs, level of impact of the OD on the disease/extent to which it modifies the course of the disease, and use in a single indication or not, and with respect to the latter, in addition, whether the first authorised indication was as an OD and the next for a frequent illness or vice versa. Several studies have poured criticism on the inclusion of some of these criteria, which will be discussed in the penultimate section of this chapter. In addition, with this model, by stratifying these criteria on a scale with three categories (low, medium and high) of different parameters, it is possible to guide, for example, the setting of price based on value, understood as the result of the inclusion of the selected criteria.

The other three theoretical models published cannot be evaluated in sufficient detail because they have only been reported as abstracts (Hutchings et al., 2012; Ethgen et al., 2012; Fedyaeva et al., 2014).

TABLE 8.2. MAIN CRITERIA INCLUDED IN THE MCDA OF SOME OF THE STUDIES REVIEWED

| Criterion | Wagner (2016) | Hughes-Wilson (2012) | Hutchings (2012) | Iskrov (2016) | Sussex (2013) | Garau (2016) | Goetghebeur (2010) | Wagner (2017) | Kolasa (2016) | Gilbert (2017) |
|---|--|----------------------|------------------|---------------|---------------|--------------|--------------------|---------------|---------------|----------------|
| Need for intervention | Severity | ✓ | | ✓ | | ✓ | ✓ | ✓ | ✓ | ✓ |
| | Size of the affected population | | | ✓ | | ✓ | ✓ | ✓ | ✓ | ✓ |
| | Unmet needs/alternative available treatments | ✓ | | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Comparative results of the intervention | Comparative effectiveness | | | ✓ | ✓ | ✓ | ✓ | ✓ | | ✓ |
| | Comparative safety/tolerability | | | ✓ | ✓ | ✓ | ✓ | ✓ | | ✓ |
| | Comparative patient's perceived outcomes/PROs | | | | | ✓ | ✓ | ✓ | | ✓ |
| Type of benefit of the intervention | Type of preventive benefit | | | ✓ | | ✓ | | ✓ | | ✓ |
| | Type of therapeutic benefit | ✓ | | ✓ | | ✓ | ✓ | ✓ | | ✓ |
| Economic consequences of the intervention | Cost consequences compared: cost of the intervention | ✓ | | | | ✓ | | ✓ | | ✓ |
| | Cost consequences compared: other medical costs | | | | | ✓ | | ✓ | | ✓ |
| | Cost consequences compared: non-medical costs | | | | | ✓ | | ✓ | | ✓ |
| Knowledge about the intervention | Quality/Strength of evidence | | | ✓ | | ✓ | ✓ | ✓ | | ✓ |
| | Expert consensus/Clinical practice guidelines | | | | | ✓ | ✓ | ✓ | | ✓ |
| Priorities of the population | Rarity | ✓ | | | | ✓ | | | ✓ | |
| | Other priorities/Vulnerable groups | | | ✓ | | ✓ | ✓ | ✓ | | |
| Contextual regulations | Alignment with the mandates of the healthcare system | | | | | ✓ | ✓ | ✓ | | ✓ |
| | Alignment with common goals | | | | | ✓ | | | | ✓ |
| | Environmental sustainability | | | | | ✓ | ✓ | | | ✓ |

| Criterion | Wagner (2016) | Hughes-Wilson (2012) | Hutchings (2012) | Iskrov (2016) | Sussex (2013) | Garau (2016) | Goetghebeur (2010) | Wagner (2017) | Kolasa (2016) | Gilabert (2017) |
|---------------------------------------|--|----------------------|------------------|---------------|---------------|--------------|--------------------|---------------|---------------|-----------------|
| Contextual feasibility | Payment capacity and opportunity costs System capacity and proper use Political, historical and cultural context | | | | | ✓ ✓ | ✓ ✓ ✓ | ✓ ✓ ✓ | | ✓ ✓ ✓ |
| Level of research undertaken | | ✓ | | | | | | | | |
| Uncertainty about effectiveness | | ✓ | | | | | | | ✓ | |
| Manufacturing complexity | | ✓ | | | | | | | ✓ | |
| Follow-up measures | | ✓ | | | | | | | | |
| Modification of course of the disease | | ✓ | | | ✓ | | | | | |
| Survival | | | | | ✓ | | | | | |
| Use in a single indication | | ✓ | | | | | | | ✓ | |
| Burden of disease | | | ✓ | ✓ | ✓ | | | | | |
| Family and social impact | | | ✓ | | ✓ | | | | | |
| Budgetary impact | | | ✓ | ✓ | | | ✓ | | ✓ | |
| Cost-effectiveness | | | | ✓ | | | ✓ | | ✓ | |
| Life-saving | | | | ✓ | | | | | | |
| Scientific innovation | | | ✓ | | ✓ | | | | ✓ | |
| Public health interest | | | | | | | ✓ | | | |
| Stakeholders pressures | | | | | | | ✓ | ✓ | | |

Source: author's preparation from Wagner, 2016; Hughes-Wilson, 2012; Hutchings, 2012; Iskrov, 2016; Sussex, 2013; Garau, 2016; Goetghebeur, 2010; Wagner, 2017; Kolasa, 2016; Gilabert, 2017.

8.3.2 Application of MCDA to simulated orphan drugs

In the context of the healthcare system in Bulgaria, an MCDA was carried out for the assessment and appraisal of two simulated ODs: A (prevalence < 1/10,000 inhabitants (ultra rare), beginning in childhood, great disability and great need for care) and B (1-5/10,000 inhabitants, variable onset, intermediate disability and need for intermediate care) (Iskrov et al., 2016). The selection of criteria and weights was carried out by means of a survey of 143 participants (47% of answers) representing four preferences: doctors, healthcare authorities, patients and industry representatives. In the model, additive and linear, a scale of 100 points distributed among criteria grouped into three categories was used: drug (existence of alternative, safety, effectiveness, health benefits and whether it saves life), illness (burden and severity) and public health considerations (preventive effect, vulnerable groups, strength of evidence, cost-effectiveness and budgetary impact) (Table 8.2). The individual weights of each criterion were normalised to add up the total weight of each category.

The was the category with the highest score, although there were no differences between the weights assigned to the three categories by the four groups consulted. Regarding the OD, there was a general consensus, except for the criterion of whether it saves life, which was more highly valued by the authorities. Regarding the disease, the severity was scored a little higher than the burden of disease, although without differences between the groups. Regarding public healthcare considerations, the criterion most valued by all was the strength of the evidence, and the least valued were the cost-effectiveness and the budgetary impact, without differences between the scores of each. Overall, the OD B received a higher score than the A, although the difference was very small in both the best and the worst scenarios delineated for each drug. An additional contribution of this model – whose suitability and degree of arbitrariness should be assessed by the reader in the absence of explanation by the authors – was the establishment of three thresholds for reimbursement: < 50 points on the scale (no reimbursement), 51-69 (conditional reimbursement) and ≥ 70 (unconditional reimbursement).

8.3.3. Application of MCDA to real orphan drugs

In the first of the six studies in this category, use was made of information about two orphan drugs which were not identified as being in the final stages of the authorisation process (Sussex et al., 2013b): the orphan drug A, to treat an immunodeficiency, between 250 and 600 affected patients, there are other treatments available, it aims to increase survival and the evidence comes from an uncontrolled open clinical trial; and orphan drug B, for a neuromuscular RD, between 3,000 and 8,000 affected patients, there are no alternative treatments, it aims to slow the progression of the disease, and the evidence comes from a controlled double-blind clinical trial.

In this study, 8 criteria were selected, 4 relating to the impact of the RD and the associated unmet medical need (availability of treatment, survival with available treatment, morbidity and disability with such treatment, and social impact of the disease in the daily life of the patients and their caregivers) (Table 8.2). The cost-effectiveness was omitted to preserve the principle of redundancy (avoiding double counting), which will be discussed in another section. Adopting the perspective of society, an MCDA to measure the combined value of the

ODs was performed by two groups of stakeholders: patients and experts (doctors and health economists), who assigned weights to the criteria in a linear manner, on a scale from 1 to 7.

The patients assigned higher weights to the two ODs and more homogeneous weights to all the criteria, and gave more value to the impact of the disease and the OD than did the experts. The absence of treatment and the evidence of the clinical efficacy of the OD were less important for patients, although the latter criterion was the one which was most highly valued by the two groups. Experts and patients rated drug A more highly than drug B. The reason was the same in both: with treatment A, better results are achieved. In accordance with the sensitivity analysis (modifying scales and weights), only large-scale and unlikely changes in the weights of the criteria could modify the choice of OD.

In the second study, which was the first performance of MCDA in Italy, the version 3.0 of the EVIDEM framework was applied for the appraisal of obinutuzumab, indicated for the treatment of non-Hodgkin lymphoma refractory to rituximab (Garau et al., 2016). 9 patients, 5 physicians and 5 financiers were consulted for the appraisal of obinutuzumab combined with bendamustine, compared with bendamustine in monotherapy. The patients and physicians expressed their greater preference for ODs used for serious diseases and a lower preference for economic criteria, and the financiers awarded the highest value to the economic criteria and the quality of the evidence. According to all stakeholders, the main endpoints of obinutuzumab were those of the disease: severity, type of therapeutic benefit and unmet medical needs. The three main limitations of the study were the possible low representativeness of the sample of respondents (it was small and convenient), the parties' difficulties in understanding the EVIDEM framework, and the non-compliance with the preferences independence requirements (evaluating the value assigned to a criterion while not knowing that assigned to the other criteria of the model). In addition, when assigning points in two stages (first to the domains, and then to the criteria of each domain), it was observed that the domains with two criteria received higher scores than those with three. To overcome this distortion, especially when larger scales are used, it is recommended to find out the preferences of the participants using methods of solid theoretical basis, such as discrete-choice experiments (explained in previous chapters) or the PAPRIKA method.

The third is an appraisal performed as a proof of concept with the intention of complementing the HTA in Canada of the growth hormone for the treatment of Turner's syndrome, with the alternative of not administering treatment (Goetghebeur et al., 2010). 4 pediatric endocrinologists, 1 ethics expert, 1 nurse expert in Turner's syndrome, 1 patient and 2 health economists were consulted. In this study, the expanded EVIDEM framework was applied with an extrinsic assessment tool which includes an ethical framework (with the healthcare system's target criteria, opportunity cost, priority and access) and other components (system capacity and appropriate use of the intervention, pressures from the agents, and historical and political context) (Table 8.2). The final hormone score was 41 points out of 100 with an acceptable conformity at individual level (CCI= 0.687), although with a wide variation between stakeholders. The discrepancies in the test and re-test were greater in the weights than in the scores, and at the individual level the greatest discrepancy was due to the cost-effectiveness assessment. The most valued criteria were improvement in efficacy, severity and quality of the evidence. With the scores of the extrinsic tool criteria, the assessments of the perceived value of the growth hormone were disparate.

In the fourth study, the EVIDEM framework was used to estimate in France, Italy and Spain the contribution of different criteria to the value of lenvatinib for the treatment of differentiated thyroid carcinoma refractory to radioactive iodine, compared with sorafenib and with controlled waiting (Wagner et al., 2017). The information about the value of lenvatinib was obtained in meetings with physicians, patients, methodologists and decision-makers from the three countries. The highest weights were given to the comparative effectiveness and the quality of the evidence (in Spain and Italy) and to the severity (in France). The criteria which most contributed to the value of lenvatinib were comparative effectiveness, severity, unmet needs, and evidence. The superiority of this drug compared with the comparators was robust in all the analyses performed. The contributions of the criteria to the value varied with each comparator, between countries and between individual participants. The results were reproducible at group level, and the impact of the contextual criteria varied between countries. In addition to the potential expected utility of the MCDA, this study highlights the importance and influence of the context, the healthcare systems of each country and their cultural values in the results.

In this category there are also two studies aimed at estimating the ability of the MCDA to estimate the value of ODs in relation to their price. In the MCDA of the first study (Schey et al., 2017), the Hughes-Wilson model was used (Hughes-Wilson et al., 2012) (Table 8.2) with three scenarios: 1) assignment of the same weight to all the criteria, 2) assignment after exclusion of the level of research carried out, uncertainty about effectiveness, complexity of manufacturing and unique indication or not, and 3) assignment after exclusion of manufacturing complexity and unique indication or not. The sensitivity analysis was carried out because the authors believed that the model omitted some relevant criteria, since it had been constructed from the perspective of the manufacturer, and because some are not criteria which are used in the HTA when making reimbursement decisions. Six ODs were chosen to cover a broad range of their annual average costs per patient (indicated for primary pulmonary hypertension, mucopolysaccharidosis II and IV, paroxysmal nocturnal haemoglobinuria, Lennox-Gastaut syndrome, and myelodysplastic syndromes). The R^2 of the regression between the global score of the value of each drug and the price showed a strong linear relationship (0.7869).

The second study is another practical application of the model of Hughes-Wilson et al. (Schey and Conolly, 2014), which aimed to find out whether there is a linear correlation between the individual added value of a group of MCDA and their respective indications for various RDs and their price. After reviewing the bibliography, new criteria were selected which were added to the nine of the original Hughes-Wilson model: convenience of administration, age of the target population, quality of life and degree of innovation of the drug (Table 8.2). For each criterion, weights from 1 to 3 were assigned and subjected to a sensitivity analysis, in which the inclusion and exclusion of the price of each drug were also evaluated (R^2 were 0.808 and 0.704 respectively). As it was published in summary form, there is no more information to allow a thorough evaluation, but it can be considered as an example of the potentially useful applications of MCDA.

8.3.4. Other models of combination of multiple criteria

The first model found among other combinations was constructed in the context of the 'Ontario Public Drug Programs of Canada' (Winqvist et al., 2012; Clarke et al., 2009). In it,

a framework was designed to inform decisions about the reimbursement of ODs from the perspective of the healthcare system. The evaluation framework was structured in seven steps: 1) confirm that the disease is really rare, 2) understand it, 3) understand the potential value of the OD, 4) model the effectiveness of the OD, 5) estimate the cost and issue a recommendation of funding, 6) review the evaluation carried out with experts and other stakeholders, and 7) re-evaluate the results as more information is obtained. This framework was validated by evaluating the OD idursulfase for the treatment of Hunter's syndrome (mucopolysaccharidosis type II). Although this drug had not been considered cost-effective (its ICER threshold ranges between \$40,000 and \$60,000), the results of the model show that the increase in life expectancy which is achieved, although reduced, was highly valued. In the light of this result, the review of the Markov model and funding recommendations, together with further negotiations with the pharmaceutical company, led to the approval of its selective public funding in the subgroup of patients with the best response (6 or more years of age and without neurocognitive symptoms).

An added advantage of this experience is that information about the incidence and natural history of the disease was available, which does not always occur with RDs. However, the validation was performed with a Markov model based on a Bayesian approach, which is considered highly speculative. It is, therefore, an experience which, although limited, provides as an innovation a framework for informing iterative reimbursement decisions whose methodological limitations must be reduced to increase their robustness.

Another distinct, novel and unique model for combining criteria brings together factors associated with the social value of ODs, which are linked to their reimbursement in a context of limited healthcare resources (Paulden et al., 2015). The model is composed of three elements. First, 19 of the factors identified in a review of the literature, which are grouped into three categories: 1) those that determine the opportunity cost of the coverage of the ODs and its comparators (cost and budgetary impact of treatment), 2) those related to the value assigned to the ODs, their comparators and the opportunity cost of each one (linked with the disease, with the drug, with the target population and with socioeconomic and other factors), and 3) factors which neither determine the opportunity cost nor are associated with the value, but which are relevant in the coverage decision (feasibility of the diagnosis and treatment of the RD and cost-effectiveness of the treatment). Second, the preferences of patients, doctors, industry, decision-makers and society in general among the factors identified and how they should be implemented in a reimbursement decision. And, third, the value proposals, the views about how limited resources should be allocated among competing groups of patients or how these factors should be incorporated into a decision framework: the rule of rescue, the equity principle and the rights-based approach.

This theoretical contribution, essentially methodological and complex, includes the functions of each of these three elements in a framework of decision: the value which each stakeholder assigns to each treatment, the value which each value proposal assigns to each treatment, and the assessment of the treatment and of the buyer that each decision-maker makes together with the opportunity cost of the treatment and its comparators. When making a coverage decision, the decision-maker compares their valuations of the OD, its comparators and the opportunity cost of each one.

8.4. THE APPRAISAL OF HEALTH TECHNOLOGIES AND THE MCDA WITH ORPHAN DRUGS: PRACTICAL CASES

In previous chapters we analysed the role, the utility, the advantages and the disadvantages that the incorporation of MCDA can have as a complement to conventional HTA processes. Two empirical studies and a reflection have been published on this subject, in which the feasibility and usefulness of adapting and incorporating MCDA in the appraisal of ODs are evaluated, as well as the congruence of the decisions which arise from its application with those produced by the appraisal methods that have been used to date, one in the Hungarian Health System (Endrei et al., 2014), another in the Health System of Poland (Kolasa et al., 2016), and the third in the Health Service of Catalonia in Spain (Gilabert-Perramon et al., 2017).

In Hungary, the MCDA was incorporated in 2010 but only for the appraisal of new technologies other than drugs, with six criteria: healthcare priorities, severity, equity, cost-effectiveness, quality of life, added budgetary impact and national and international recommendations. In the decisions about coverage and reimbursement of all drugs, the cost-effectiveness analysis is used. The authors affirm that MCDA has been useful in their decisions to date and that its scope should be broadened (Endrei et al., 2014).

In the Polish case, the impact of the use of MCDA in decisions about the price and reimbursement of ODs was evaluated, and the results were compared with the recommendations (positive and negative) of the HTA of 27 ODs with the indication for each one (Kolasa et al., 2016). For the MCDA, two sets of main criteria, extracted and selected by consensus, were used by two reviews of the consulted bibliography: one with 8 criteria and another with 10, adding cost-effectiveness and budget impact to the first group (Table 8.2). The same weight was assigned to each criterion and they were combined using a simple additive linear model. In summary, with the 27 ODs there were 12 discrepancies between the recommendations of the HTA and the results of the MCDA (most of them involved positive recommendations from the HTA and negative ones from the MCDA). All the ODs with negative recommendations from the HTA received positive recommendations with the MCDA. Economic information was available for only 12 ODs, of which 2 were not recommended by the HTA and were by the MCDA, and 2 the opposite when excluding the two economic criteria. The authors conclude that decisions from the two methods may differ because the consideration with the MCDA of more decision-making criteria than in the HTA may involve either a greater scrutiny of these drugs, leading to their rejection, or the revelation of greater value, thus increasing the probability of approval.

The study is not exempt from limitations which reduce both its internal and its external validity. Among them, apart from having constructed a simple and linear model and not describing in enough detail the HTA process currently in use in the country: the subjectivity of the reviewers' selection of criteria stands out clearly; only the researchers participated; the quality of the primary data was not analysed and the economic data were very incomplete, together with the double counting problems which the inclusion in the cost-effectiveness model (with three categories) could create; the analysis of the uncertainty of the available information was omitted, and the only sensitivity analysis carried out consisted of assessing the variability of the MCDA results with three positive recommendation thresholds: more than 50% of the total points of the maximum added value (base case), more than 25%, and

more than 75%, which showed wide variations in the criteria which received higher and lower scores and, consequently, in the drugs included and excluded in each case.

In the Catalan case (Gilabert-Perramon et al., 2017), 16 of the 27 members of the appraisal programme [*Programa d'avaluació, seguiment i finançament dels tractaments d'alta complexitat (PASFTAC)*] who agreed to participate (evaluators, a representative of patients and another of the decision-makers) adapted and validated the EVIDEM framework, in accordance with its standardised procedures. All the domains and quantitative criteria were maintained and from the contextual tool three of them were excluded and one was adapted (mandate and scope of the healthcare system, priorities and access of the population, environmental impact, and political, historical and cultural context) (Table 8.2). The corresponding weights were assigned using two techniques: a 5-point scale and a hierarchical point assignment. The test and re-test showed that the tendency of both was similar, and the intra-class correlation coefficient (< 35) showed that the agreement was weak in 10 of the 13 criteria selected in the final model. The criteria with the greatest weight were severity, unmet needs, comparative effectiveness, type of therapeutic benefit, and quality of evidence, which is consistent with those considered in other studies (Wagner et al., 2016; Sussex et al., 2013b). The one with the least weight was the size of the affected population.

The procedures of the PASFTAC and of the EVIDEM framework are very similar, since in both cases the same main criteria are considered to a large extent. In this attempt to apply the MCDA, areas of improvement were identified in the notification of all the criteria used and in their scales of scoring and weighting, as was done in the Lombardy experience in Italy with the EUnetHTA model, which incorporates elements of the MCDA, although, as in another application of the EVIDEM framework in the Canadian context, they are not applied to the ODs either (Radaelli et al., 2014; Tony et al., 2011). The authors point out that, with good training of the participants and an increase in the participation of actors other than the evaluators, the EVIDEM framework offers a standardised deliberative environment which could be useful for developing the MCDA adapted to the HTA and to the context and appraisal procedures specific to each country, because it promotes the systematic and explicit consideration of all elements of decision-making, including ethical considerations, and can serve as a basis for structured appraisals. They also state that for a procedure such as the PASFTAC, which in coordination with the '*Comisió d'Avaluació Econòmica i d'Impacte Pressupostari*' includes cost-effectiveness and budget impact in its decisions, all costs must be included in the deliberations, and that MCDA is a useful tool for structuring and reporting qualitative and quantitative information about the use and impact of economic resources, a recommendation which is not exempt from criticism, as will be explained later.

8.5. RECOMMENDATIONS AND INTERNATIONAL CONSENSUS ABOUT ASSESSMENT AND APPRAISAL OF ORPHAN DRUGS AND THE INCORPORATION OF MCDA

The inclusion in this chapter of the two most recent consensus documents, which detail some of the discrepancies between assessment and appraisal, the indicated limitations of the HTA and the economic evaluation of the ODs, is justified because both of them issue a set of recommendations aimed at reducing these limitations, one of them being to incorporate several (multiple) criteria in the determination of the value of the ODs (Gutierrez et al., 2015; Annemans et al., 2017).

The recommendations which they include encompass the following dimensions: decision criteria, decision-making processes, sustainable financing and coordination among the countries of the European Union. Table 8.3 lists these recommendations.

TABLE 8.3. RECOMMENDATIONS FOR INCREASING CONGRUENCE BETWEEN THE CRITERIA IN ASSESSMENT AND THOSE IN APPRAISAL OF ORPHAN DRUGS

| European Working Group for Value Assessment and Funding Processes in Rare Diseases (ORPH-VAL) | Principles for a coherent assessment of the value and sustainable financing of orphan drugs in Europe |
|---|--|
| COMMON PRINCIPLES | |
| The assessment of ODs should consider all relevant elements of product value in an appropriate multi-dimensional framework | The assessment must consider all the value elements of the ODs |
| Pricing and reimbursement decisions should be founded on the assessment of the OD's value and adjusted to reflect other elements beyond that value | The consideration of the value of these drugs must apply multiple criteria |
| Those making P&R decisions at national level should take into account the relevant European regulations (COMP, EMA, INAHTA) | Price and reimbursement decisions must recognise the assessment of the EMA |
| Experts in RDs should participate in them, including the perspectives of professionals, healthcare staff and patients | The national authorities must involve experts in RDs in their local decision-making |
| To respond to uncertainty, the assessment and P&R decisions should be adaptive to the availability of new information | Mechanisms of value assessment must be flexible in order to accommodate uncertainty at time of ODs approval |
| Funding should be established at national level to guarantee access for all | Funding must be adequate to ensure optimal access for all and to encourage research |
| Evidence-based funding mechanisms should be developed to ensure long-term sustainability | Evidence-based funding mechanisms must be developed to ensure long-term sustainability |
| If used, the ICER thresholds should be modulated to reflect the singularities of the RDs (rarity, unmet therapeutic needs and social preferences) and the need to maintain innovation | If used, the ICER thresholds should be modulated to reflect the specificities of RDs and ODs |
| In the future, the coordination of assessment processes at European level should increase | The processes of pricing and reimbursement of the countries should recognise the assessment of therapeutic benefit made by the COMP of the EMA |
| All affected patients should be taken into account in price and reimbursement decisions at national level, although some coverage decisions only include sub-groups | The representation of the patients should be systematic in all meetings for the assessment and appraisal of ODs |
| PRINCIPLES NOT SHARED | |
| | National authorities should develop adaptive and efficient processes in order to optimise the use of real-life data obtained before and after authorisation Reimbursement decisions have to be made with references of the value and price of treatments with similar characteristics |

OD: Orphan Drug. ICER: Incremental Cost-Effectiveness Ratio. RD: Rare Disease. P&R: Pricing and Reimbursement. COMP: Committee for Orphan Medicinal Products. INAHTA: International Network of Agencies for Health Technology Assessment. EMA: European Medicine Agency.

Source: Annemans et al., 2017. Gutierrez et al., 2015.

8.6. CRITICAL APPRAISAL AND LINES OF PROGRESS

8.6.1. General comments

The joint analysis of all the reviewed studies about the application of MCDA to ODs reveals several important facts. First, the applications of MCDA to these drugs are scarce and very recent. Only six applications with real ODs have been found in real evaluation conditions; the others are theoretical developments of this analysis or applications with simulated data about ODs, and the majority were published between 2012 and today. As indicated in many of the reviewed studies, knowledge about them is still very limited (Wagner et al., 2016; Hughes-Wilson et al., 2012; Hutchings et al., 2012; Fedyaeva et al., 2014; Iskov et al., 2016; Sussex et al., 2013c; Garau et al., 2016; Goetghebeur et al., 2010; Wagner et al., 2017; Schey and Conolly, 2014; Schey et al., 2017; Winquist et al., 2012; Clarke et al., 2009; Kolasa et al., 2016; Gilabert-Perramon et al., 2017; Endrei et al., 2014). And it is noteworthy that, of the few studies and applications published, a large proportion comes from Bulgaria, Poland, Russia and Hungary - something which this author cannot explain.

As it is a question of an incipient stage of its development and application, we can expect the small number, the remarkable variability of constructed analysis models and the methodological limitations compared with some of the methodological principles of their design, running and analysis which still do not enjoy general consensus among researchers, and the virtual absence of consolidated guidelines for their design, implementation and analysis. In fact, several studies reiterate that in this initial phase of its development the main principles on which its proposal is based as a method of combining criteria are still not observed: inclusion of the dimensions and the main and relevant criteria of a decision, non-overlapping, principle of independence (redundancy), allocation of weights and participation of all relevant stakeholders.

8.6.2. Methodological elements

As the methodological elements of MCDA have been explained in detail in previous chapters, here they are only analysed together in the included studies, most of which are applications of MCDA to estimate the value of ODs for reimbursement decisions (Thokala et al., 2016; Marsch et al., 2016; Goetghebeur et al., 2008; Angelis and Kanavos, 2016; Wahlster et al., 2015; Marsh et al., 2014; van Til et al., 2014; Mühlbacher and Kaczynski, 2016; Broekhuizen et al., 2015; Ram et al., 2011; Golan et al., 2011; Baltussen and Niessen, 2006; Tromp and Baltussen, 2012).

Although there are several types of MCDA and their formal classification cannot yet be considered closed, only variants of value-measurement models have been applied to ODs, not those known as outranking or goal programming. For the MCDAs used with ODs, the main sources of information and selection of the dimensions, criteria and, in some models, sub-criteria, are reduced to literature reviews, expert opinions and other models which are used as references or starting points and subjects for adaptation. To date, the model which in relative terms seems to be the most used, or used as a reference, is the EVIDEM.

Regarding the dimensions and the criteria which include the models which have so far been applied with ODs, there is evidence that not all of them include a reasoned justification of all the chosen ones. The smaller amount of available information about the criteria of interest of

both the ODs and the RDs could explain the fact that the evidence obtained about each criterion and its weights varies between studies, in some cases is scarce and, for these reasons, the fact that the associated level of uncertainty which all this incorporates into the models is high, including the variability of opinions and preferences expressed by the participating stakeholders (patients, doctors and other healthcare staff, decision-makers, health economists and other experts).

It has also been noticed that some criteria which reflect the same concept or dimension have different names in different studies. Furthermore, failure to comply with the properties of the chosen criteria (integrity, non-redundancy, non-overlapping, independence of preference) is evident in many applied models. Also, the range of criteria selected in several applications does not cover the entire process about which decisions are supposed to be taken, for example, coverage and reimbursement of ODs [structure-inputs, process (e.g., access or use)] and results. This could be due both to the lack of formal training with MCDA and to the fact that the perspective adopted is exclusive to one of the parties, as shown, for example, by the inclusion as a criterion of the manufacturing process of the OD, that which has one or more orphan indications or the degree of innovation of the OD, which belong to the manufacturer's perspective. The scant development achieved by these analyses could also explain that no applications to ODs have been found in which techniques have been used to maximise the best selection of attributes and criteria among individuals and among groups, which becomes more relevant in a reviewed case in which MCDA is used simultaneously in several countries with different comparators of the evaluated drug (Wagner et al., 2017).

Despite this variability between studies, partial compliance with methodological principles and insufficient reporting of studies, in all the applications of MCDA the models are simple linear and additive with scales from 0 to 2, from 1 to 7 or from 1 to 10. In one study it was found that domains with two criteria tend to receive higher scores than those with three. To overcome this distortion, especially with larger scales, it is recommended to find out the preferences of the participants with methods with solid theoretical foundation, such as discrete-choice experiments or the PAPRIKA method (Thokala et al., 2016). If the weights are used for group decisions, the technique for revealing them has no impact on the weights of each criterion or on the overall value of what is evaluated, although when they are used to support individual decisions, the revelation technique does influence the final results. Nevertheless, the studies published to date have not analysed or reported on the efficiency, construction, accuracy, implementation, or discriminatory capacity of the scales and weights, nor the aggregation method (for example, measuring the severity or impact on health or the degree of innovation in different ways). In very few studies have there been estimates of the concordance of assessments of individual participants and those of group test/re-test in the selection of criteria and assignments of weights, and no use of the performance matrix has been found.

Variability in the selection of agents participating in MCDA is also observed. In the theoretical applications, and with real data, not all the stakeholders are represented in those in which the preferences of all of them are supposed to be combined. In one it is indicated that it is a proof-of-concept study, but in the others account is not taken of the omission of necessary preferences. In all of them it is recognised that the samples are small and con-

venient and with unequal representation of the relevant agents. As most of the cases were pilot studies, it is also recognised that the samples could be biased because participants with higher levels of education than those of the respective groups of the general population were included.

The results are disparate in terms of strong preferences in the area of RDs and ODs. In some studies, agents with different preferences rated the treatment (the OD) more highly than the disease (Fedyayeva et al., 2014; Iskrov et al., 2016), and in another, by contrast (Garau et al., 2016), the experts preferred the treatment to the disease, and with patients it was the opposite (Sussex et al., 2013). These disparities are also seen between studies in relation to the inclusion of the criteria through which RDs are attributed a special status compared with other diseases: rarity, severity, unmet needs, lack of alternative treatment and the OD's life-saving capacity (Table 8.2).

Two proposals have been found – without any known further development – of models for combining multiple criteria different from the MCDA. One provides a decision framework for reimbursement (with idursulfase) in several stages and iterative (Winqvist et al., 2012; Clark et al., 2009), and the other, consistent with the principles of cost-utility analysis (CUA), brings together functions of the social value of the ODs, of the preferences of the groups of agents and of the proposals or theories of the value of the allocation of resources, which is discussed in more detail in the following paragraph (Paulden et al., 2015).

The inclusion of economic criteria (costs, cost-effectiveness) is variable. Consensus about them has not been reached in the published models and there is often an equivocal definition of criteria, overlapping and double counting of costs. In only four of the models found was the cost-effectiveness ratio included as a dimension or criterion (Iskrov et al., 2016; Goetghebeur et al., 2010; Endrei et al., 2014; Kolasa et al., 2016), and its omission is only justified in two of those which exclude it. In the first one, which analyses the application of the EVIDEM framework to ODs, it is indicated that, as the framework already includes the criteria captured in this ratio's numerator and denominator, the ratio must be excluded in order to comply with the principle of non-redundancy or double counting (Wagner et al., 2016). In the second, one of the models of combination different from the MCDA, it is indicated, in addition, that effectiveness is a function of the other criteria which intervene in the decisions about prioritisation or reimbursement and which capture, in the estimation of the QALYs gained, the severity and the magnitude of the benefit of the treatment, and in the ICER threshold, the opportunity cost (Paulden et al., 2015). In fact, this new model is an alternative to the others, consistent with the principles of the CUA, and which preserves the nuclear consideration in this analysis of the opportunity cost, and, in addition, allows one to incorporate, without restriction, more elements in addition to the value of the OD, the value proposals and the perspectives.

In the studies reviewed, uncertainty was not analysed with any of the available methods (based on probability, deterministic analyses, Bayesian methods, those based on the fuzzy theory, on the grey theory, on risk, or on scenario analyses). This analysis was reduced to such simple scenarios as the inclusion and exclusion of one criterion and the modification of the values of the scales and weights.

It should also be emphasised that in several applications the participants express their difficulties in understanding the methods fully. The absence of consolidated design guidelines, for the selection of the most appropriate method for each context, including the cultural one, and for each situation (determined by the RD and the OD) of conducting, analysing, interpreting and reporting the results of the MCDA, and preferably on an international scale, is one of the pending subjects to be advanced in its development and in its applications. The ISPOR reports and the generic EVIDEM framework can be good starting points, because they are the ones that have advanced the most in this direction, as discussed in Chapter 6 of this book.

On the subject of value-based pricing models, three studies have been found. Two of them develop mathematical functions which link value with price (Ethgen et al., 2012; Schey et al., 2017), and the other one establishes three differential price thresholds based on the scores of the criteria included in the model of combination of multiple criteria (Hughes-Wilson et al., 2012). All have been carried out from the perspective of the manufacturer and are initial models from whose subsequent application and validation information is lacking.

8.6.3. HTA and MCDA

With regard to the potential complementarity of HTA and MCDA, it is convenient to underline here some key reflections in the bibliography, although they are of a general nature and do not relate specifically to ODs, because they are also valid, and are even more important, in decisions about the reimbursement of these drugs.

First, from a perspective which is more conceptual than pragmatic, the current practice of HTA does not include all the dimensions of different kinds that it deals with, nor does it treat them systematically. One of its main problems is the combination of the wide range of variables which must be considered in its mission: those related to the costs and consequences of the new technologies, to ethics, justice, culture, and the social preferences linked to decisions. These variables often pose problems of interpretation, which are not limited to the understanding of preferences and of the benefits deriving from them. As indicated in the introduction, decisions about prioritisation and the allocation of resources also involve the judgements and values of people and society in relation to multiple principles and criteria. To properly combine all the relevant variables in appraisal decisions requires quantitative methods and, in addition, deliberation, and the area of ODs is one in which these facts acquire more prominence (Claxton and Culyer, 2006; Culyer and Lomas, 2006; Culyer, 2012; Culyer, 2014).

The pending subject for the HTA is not so much the metrics as the design of the processes of deliberation which allow the relevant dimensions to be combined, especially those which are more difficult to apprehend and incorporate in decision-making, as well as individual and social preferences – those of a moral, ethical or even religious nature (Claxton and Culyer, 2006; Culyer and Lomas, 2006; Culyer, 2012; Culyer, 2014; Culyer, 2016).

Secondly, from the most practical and metric perspective, in many of the publications reviewed in this chapter, and in one which can be considered as a synthesis, there is a debate

about the usefulness of MCDA as a complement to HTA. It is argued that MCDA provides a method to structure and combine these variables, as well as to test these judgements, include them in deliberations and decisions, and notify the deliberative process as a whole with transparency and consistency, and on behalf of society, and in doing so, maintaining a balance between the flexibility and congruence required in the decision-making processes, fleeing from opaque deliberations (Devlin and Sussex, 2011). Some of the studies reviewed indicate that MCDA is not intended to replace decision-making, but to complement and facilitate it (Devlin and Sussex, 2011).

8.6.4. Criticisms concerning the ineptitude of economic evaluation for the appraisal of orphan drugs

Table 8.4 summarises the main published replies to the arguments which justify MCDA as an alternative which overcomes the limitations of HTA and economic evaluation in making decisions about the value of orphan drugs, and about their coverage and reimbursement. Some people regard MCDA as a complement to HTA and economic evaluation, and others regard it as a tool for a new evaluation paradigm. The resolution of the conflict between these opposing arguments contains, among other things, an important deliberative component, because when comparing them, one notices different conceptions and erroneous interpretations of some concepts, and biases about others. The original reasons are different in the two cases. In good scientific logic, the unacceptable ones are those which make judgements without providing the reasons which unequivocally demonstrate their veracity, as when, for example, something is described as being arbitrary.

TABLE 8.4. REPLIES TO THE ARGUMENTS ABOUT THE LIMITATIONS OF THE HTA AND THE ECONOMIC APPRAISAL IN MAKING DECISIONS ABOUT THE VALUE OF ORPHAN DRUGS AND ABOUT THEIR COVERAGE AND REIMBURSEMENT

| | |
|----------|---|
| 1 | It is dangerous to say that some diseases (rare diseases, some types of tumours) have more social value than others. Severity or endangering life are criteria common to many diseases, and are independent of the opportunity cost of financing ineffective, not very effective, or non-cost-effective treatments, especially considering that the threshold of inclusion of technologies for financing is or should be the one with the lowest ICER. Horizontal justice would force one to do the same with other diseases. There are empirical studies which show variations in social preferences with respect to these criteria. The ability to add or combine multiple criteria and the problem of deliberations have been dealt with in sections 8.1 and 8.6.3. |
| 2 | The definition of inequitable depends on that of its opposite, and equity, in addition to having been defined as equanimity in the access to treatments, is also conceptualised as absence of biases and favouritism. The use of the ICER or the incremental cost-utility ratio (ICUR) seems to be more equitable than other methods which assign greater value to certain selected diseases, because by using these ratios, one allocates scarce resources to maximize social benefit, and the health gains of people with different diseases are valued in the same way. That certain drugs do not exceed the authorisation and reimbursement standards may be due to several reasons. First, they are not effective or safe according to regulatory standards. Second, the evaluation which society makes of their benefits does not exceed their costs at current prices. The fact that they only provide some benefit, and that their relative effectiveness is low or limited because it was estimated with substitute variables is not a sufficient condition to finance them, or, some argue, to authorise them either, given that in this case their financing would involve, due to their high price and few benefits, the loss of greater benefits for other patients due to their opportunity cost. |

OD: Orphan Drugs. ICER: Incremental Cost-Effectiveness Ratio. ICUR: Incremental Cost-Utility Ratio. QALYs: Quality Adjusted of Life Years. CEA: Cost-Effectivity Analysis.

Continue

Continued

| | |
|---|---|
| 3 | <p>The QALYs do not depend on a utilitarian or welfarist approach, because these approaches reject the information which is not limited to utilities because it is considered irrelevant in judgements about efficiency and justice. As the QALYs include concepts and characteristics of individuals which affect their well-being, they transcend the welfarist concept of utility and do not accept that it is the only result of interest in an evaluation. This makes them extra-welfarist. In a system with limited budgets, the cost of providing health services is the benefits which other patients fail to perceive. A service should not be financed when the magnitude of its benefits does not socially justify that of its cost, when the benefits gained are less than those lost by other patients; all patients count, as do those who will be patients in the future. The comparison between health gained and lost is inherent in the cost-effectivity analysis (CEA). QALYs have been adopted because they are generic and not specific to the procedure being evaluated (procedures are evaluated, not people), their advantages and limitations are known, and the deviations observed from their assumptions because they are considered inadequate are submitted to deliberation. The value judgements subsumed under the QALYs come from empirical research conducted to detect biases among the patients affected by the decisions. QALYs offer the common outcome measurement which is needed to compare, and they have the same social value for all patients. Their limitations are not greater than those of other measurements. The main difference between QALYs and similar measurements is that the former identify the need to deliberate about value judgements. The NICE Appraisal Committee is a good example of this. The lack of consensus about whether any characteristic or dimension of individuals should receive more or less weight than others is not resolved by QALYs, because it is matter for deliberation. The accepted limitations of QALYs are common to all technologies or processes which are evaluated, and are not exclusive to any of them, a position which is consistent with equity.</p> |
| 4 | <p>Like the ICER, QALYs only provide part of the information necessary to make decisions, they do not replace them. The choice of different ICER thresholds for the treatment of different diseases should be justified in such a way that any doubt about bias or conflicts of interest introduced into their justification by interested parties is dispelled. Although it is in the industry's interests to have more drugs on the market, that of any healthcare system is to submit them to a fair test of cost-effectiveness, because the costs of providing healthcare services are the benefits which other patients fail to perceive. If it is decided that a service will not be financed, this means that the benefits gained with it would be much lower than those lost by other patients; all patients count. What justification supports the inclusion of a non-cost-effective service, whose benefits lost by other patients exceed those provided by the included treatment? Increasing the ICER threshold more increases the loss of health of other patients because it allows the introduction of services which are less cost-effective than the least cost-effective services already included. There is no reason to suppose that these losses imposed on the healthcare system by the inclusion of an OD, or of a treatment for cancer, or of another disease, are greater or less than the losses of benefits of the same magnitude imposed by any other type of treatment. The principle of opportunity cost is the same for all types of decisions about the allocation of limited resources. If the budget is increased for a certain group of technologies, this is done at the expense of reducing another. The incremental cost-value ratio as a candidate to replace the ICER has not yet passed the proposal stage. Efficiency must be accepted from a moral point of view because its opposite, inefficiency, means that some patients do not receive the care which a more efficient system would afford them with the same resources.</p> |
| 5 | <p>In the reviewed performances of MCDA, accounting costs, opportunity costs, prices and expenses are confused and, in fact, in almost all of them the opportunity cost is omitted. Added to this limitation is that of double counting, as occurs with costs, imprecise definitions of criteria, such as complexity of manufacturing, interest for public health, overlapping and non-compliance with the principle of mutual independence, as is the case with the criteria of level of research done, knowledge of the intervention, level of uncertainty about effectiveness, survival and saving lives.</p> |
| 6 | <p>In most healthcare systems the price does not correspond to the added therapeutic value of a drug or to the degree of innovation. The price should not be guided by the incentives which have been given previously because they are sunk costs with respect to the decision about whether to finance a new treatment or not. Both the incentives and the price should depend on the social value given to the drugs, not the reverse. In the end, the socially desirable type of innovation is not that of drugs of very high price and exiguous benefits, but the opposite. Spending of any magnitude on a drug should never be at the expense of a greater health gain or less uncertainty than could have been obtained in an alternative use of that spending. The price as a signal should reflect the standards of efficiency (cost and effectiveness or utility) which a medicine must satisfy. The funder does not buy drugs but results.</p> |
| 7 | <p>When, in countries such as the United Kingdom or Sweden, appraisal decisions are submitted to the deliberation of all the agents, and both that process and its results are made public, are subject to the free scrutiny of society and are modified accordingly, in those countries – and not in others which are less transparent – the criticisms of lack of transparency are baseless.</p> |

Source: author's preparation from Claxton and Culyer, 2006. Culyer, 2006. McCabe et al., 2007. Laupacis, 2009. Campillo and Peiró, 2009. Culyer, 2012. Culyer, 2014. Culyer, 2016. Culyer, 2017. Palaska and Hutchings, 2015.

If the origin of the debate is erroneous interpretation, the resolution of the controversies is simple: dispel doubts by explaining and clarifying concepts and theories again. If the objective, on the contrary, is to introduce more technologies into the market in addition to all the considerations taken into account, the resolution of disputes is more complex, because, in addition, it requires one at the same time to reveal sophisms and declare conflicts of interest. Whether simple or complex, resolution is necessary, since part of well-being, in any way that this is understood, depends on it.

8.6.5. Recommendations and international consensus

With regard to the international recommendations and consensus presented (Table 8.3), it should be emphasised that the development and application of MCDA to both ODs and other drugs should not be considered regardless of the place and function that this analysis occupies in the whole of the regulatory framework. The MCDA is nothing more than the proposal of an incipient methodological tool of an intricate regulatory maze. The consequences of decisions about the approval, price, coverage and reimbursement of these drugs depend on the set of regulatory mechanisms in force and on uncontrolled exogenous variables, not only on the intervention of only one of its constituent elements.

Taken together, the principles and recommendations set forth constitute a structured and consensual proposal to improve various aspects of the regulatory framework for ODs in Europe. Although the ORPH-VAL recommendations are subsequent to the list of principles – in fact, they take them into account – the two are congruent both in relation to the elements and facets chosen for improvement and in relation to their content. Although the MCDA is not mentioned explicitly, it is directly referred to when affirming that in the evaluation of the ODs, multiple criteria and attributes must be considered, and also those which capture values which transcend those included in the conventional HTA.

In these recommendations, the combination of multiple criteria is proposed as a complement to conventional HTA methods, not as an alternative to replace them, as is defended in other published studies. The recommendations specifically advocate modulating the ICER thresholds where the fourth barrier is used in decisions about coverage and reimbursement of ODs. This is a proposal which, although imprecise, betrays its intentionality: to modulate is to modify the factors which intervene in a process in order to modify the results obtained. In this case, and knowing the modifications of the ICER threshold proposed in several studies, everything suggests that they relate to increases in the value of the threshold with the ODs as an exception, although this is also defended with oncological drugs, a measure criticised by others, as summarised in table 8.4.

8.7. CONCLUSIONS

From the information analysed in this chapter, the following conclusions are drawn:

- The number of studies which have used MCDA with ODs to date is very low and most of them have been published since 2012.
- Almost all the published applications are theoretical developments of these analyses, most of them with simulated ODs. Those which were carried out with real ODs under real conditions of evaluation correspond to pilot or proof-of-concept studies.

- The few examples of studies aimed at assessing the extent to which MCDA can complement the current HTA system show, although with important limitations, both concordances and discrepancies in the decisions which derive from each one, and also potential means of enrichment.
- There is a notable variability in the use of the methodological tools of the MCDA, for example, in the methods to select main criteria, to reveal and assign weights, estimate individual and group reproducibility, or deal with uncertainty.
- Although one of the objectives of MCDA is to combine multiple criteria, different preferences and different value proposals, in the applications carried out to date there is variability in the stakeholders that have participated in these analyses. The samples of representatives of the different parties are of small size and convenience, and in some that adopt the social perspective, relevant agents are excluded, so their external validity is still limited.
- Standardised guidelines with international consensus for designing, applying, analysing and notifying MCDA are not yet available, in contrast to those which exist for other types of studies, such as clinical trials or economic appraisals.
- Accordingly, MCDA applied to ODs are currently in the initial development phase. This fact, first, could explain the variability observed in all the aspects discussed and in the breach of essential methodological principles which could be foreseen (imprecise definition of criteria, independence between them, redundancy or double counting, omission of the opportunity cost). Second, it should guide the lines of improvement which should be introduced into its development in the short and medium terms.
- No approach or method by itself provides the answers or the final solutions. The MCDA, if the mentioned limitations are overcome, typical of any method in its conception and initial applications, could offer an additional adequate methodological tool which serves as a useful complement in the appraisal of ODs.
- In the context of the HTA, the fundamental actions to improve the appraisal of all medical technologies, including ODs, go through a proper structuring and enriching of the framework to deliberate and design empirical studies which answer the relevant questions in the decision-making process.
- Contradictory views (value proposals) about how limited resources should be allocated among agents or how these factors should be incorporated into a decision framework (the rule of rescue, the principle of equity and the rights-based approach) have to correct erroneous interpretations of concepts of economic theory in general and of economic evaluation in particular, and have to divest themselves of appraisals without theoretical or empirical foundation and of conflicts of interest, since improvement in the appraisal of ODs is largely dependent on that.

REFERENCES

- Angelis A, Kanavos P. Value-based assessment of new medical technologies: Towards a robust methodological framework for the application of multiple criteria decision analysis in the context of health technology assessment. *Pharmacoeconomics*. 2016; 34(5): 435-446. doi:10.1007/s40273-015-0370-z.
- Annemans L, Aymé S, Le Cam Y, Facey K, Gunther P, Nicod E, Reni M, Roux JL, Schlander M, Taylor D, Tomino C, Torrent-Farnell J, Upadhyaya S, Hutchings A, Le Dez L. Recommendations from the European Working Group for Value Assessment and Funding Processes in Rare Diseases (ORPH-VAL). *Orphan J Rare Diseases*. 2017 Mar 10;12(1):50. doi: 10.1186/s13023-017-0601-9.
- Baltussen R, Niessen L. Priority setting of health interventions: the need for multi-criteria decision analysis. *Cost Eff Resour Alloc*. 2006; 4(1): 14.
- Broekhuizen H, Groothuis-Oudshoorn C.G.M, van Til J.A, Hummel J.M, and Ijzerman M.J. A review and classification of approaches for dealing with uncertainty in multi-criteria decision analysis for healthcare decisions. *Pharmacoeconomics*. 2015 May;33(5):445-55. doi: 10.1007/s40273-014-0251-x.
- Campillo-Artero C, Peiró S. Enfermedades raras, medicamentos huérfanos: el valor de la orfandad. *Gest Clín San*. 2009; 11-(4): 119-126.
- Campillo-Artero C. Regaining Health Technology Assessment from Oblivion: Improving and Integrating Regulation of Drugs, Medical Devices, Diagnostic Tests and Surgical Innovations. En: del Llano, J., and Campillo-Artero, C, eds. *Health Technology Assessment and Health Policy Today: A Multifaceted View of Their Unstable Crossroads*. Springer Healthcare. 2014.
- Clarke J, Bell C, Coyle D, Stevenson H, Evans G, Martin M.J, Sabharwal M, Gadhok A, Winqvist E. A policy framework for funding drugs for rare diseases. *Value Health*. 2009. 12,7.
- Claxton K, Culyer AJ. Wickedness of folly? The ethics of NICE's decisions. *J Med Ethics*. 2006; 32(7): 373-7.
- Culyer AJ. The humble economist. University of York, Office of Health Economics. 2012.
- Culyer, AJ. Where are the limits of cost-effectiveness analysis and health technology assessment? *J Med Assoc Thai*. 2014; 97(5): S1-2.
- Culyer, J. Cost-effectiveness thresholds in health care: a bookshelf guide to their meaning and use. *Health Econ Policy Law*. 2016; 11(4): 415-32. doi: 10.1017/S1744133116000049.
- Culyer AJ. Ethics, priorities and cancer. *J Cancer Policy*. 2016; 12-3.
- Devlin N, Sussex J. Incorporating multiple criteria in HTA: Methods and processes. Office of Health Economics. 2011.
- Drummond MF, Wilson DA, Kanavos P, Ubel P, Rovira J. Assessing the economic challenges posed by orphan drugs. *Int J Tech Asses Health Care*. 2007; 23-1.
- Drummond M, Evans B, Leloir J, Karakiewicz P, Martin D, Tugwell P, MacLeod S. Evidence and values: Requirements for public reimbursement of drugs for rare diseases – A case study in oncology. *Can J Clin Pharmacol*. 2009; 16, 2.
- Drummond M, Tarricone R, Torbica A. Assessing the added value of health technologies: Reconciling different perspectives. *Value Health*. 2013; 16, S7-S13.
- Endrei D, Molics B, Ágoston I. Multi-criteria decision analysis in the reimbursement of new medical technologies: real-world experiences from Hungary. *Value Health*. 2014; 17-4.
- Ethgen O, Hutchings A, Rollet P, Standaert B. From value to price: What should be the path for orphan drugs. *Value Health*. 2012; 15, A1-A256.
- Fedyeva VK, Omelyanovsky VV, Rebrova O, Khan N, Petrovskaya EV. MCDA approach to ranking rare diseases in Russia: Preliminary results. *Value Health*. 2014; 17, A323-A686.
- Garau M, Marsden G, Devlin N, Mazzanti NA, Profico A. Applying a Multi-criteria decision analysis (MCDA) approach to elicit stakeholders' preferences in Italy. The case of obinituzumab for rituximab-refractory indolent non-hodgkin lymphoma (iNHL). Office of Health Economics. 2016; 16-8.

Gilabert-Perramon A, Torrent-Farnell J, Catalan A, Prat A, Fontanet M, Puig-Peiró R, Merino-Montero S, Khoury H, Goetghebeur MM, Badia X. Drug evaluation and decision making in Catalonia: development and validation of a methodological framework based on multi-criteria decision analysis (MCDA) for orphan drugs. *Int J Tech Assess Health Care*. 2017; 1–10.

Goetghebeur M, Wagner M, Khoury H, Levitt RJ, Erickson LI, Rindress D. Evidence and value: Impact on Decision Making – the EVIDEM framework and potential applications. *BMC Health Serv Res*. 2008; 8:210.

Goetghebeur M, Wagner M, Khoury H, Rindress D, Grégoire JP, Deal Ch. Combining multi-criteria decision analysis, ethics and health technology assessment: applying the EVIDEM framework decision making framework to growth hormone for Turner syndrome patients. *Cost Effec Resour Alloc*. 2010; 8, 4.

Golan O, Hansen P, Kaplan G, Tal O. Health technology prioritization: Which criteria for prioritizing new technologies and what are their relative weights? *Health Policy*. 2011; 102, 2-3.

Gutierrez L, Patris J, Hutchings A, and Cowell W. Principles for consistent value assessment and sustainable funding of orphan drugs in Europe. *Orphan J Rare Diseases*. 2015; 10:53.

Hughes DA, Tunnage B, Yeo ST. Drugs for exceptionally rare diseases: do they deserve special status for funding? *Qual J Med*. 2005; 98-11.

Hughes-Wilson W, Palma A, Shuurman A, and Simoens S. Paying for the orphan drug system: break or bend? It is time for a new evaluation system for payers in Europe to take account of new rare disease treatments? *Orphan J Rare Diseases*. 2012; 7-74.

Husereau D, Drummond M, Petrou S, Carswell CH, Moher D, Greenberg D, Augustovski F, Briggs AH, Mauskopf J, Loder J. On behalf of the ISPOR Health Economic Evaluation Publication Guidelines - CHEERS Good Reporting Practices Task Force. Consolidated Health Economic Evaluation Reporting Standards (CHEERS)—Explanation and Elaboration: A Report of the ISPOR Health Economic Evaluation Publication Guidelines Good Reporting Practices Task Force. *Value Health*. 2013; 16-2.

Hutchings A, Ethgen O, Schmitt C, Rollet P. Defining elements of value for rare disease treatments. *Value Health*. 2012; 15, A1-A256.

Iskrov G, Miteva-Katrandzhieva T, Stefanov R. Multi-criteria decision analysis for assessment and appraisal of orphan drugs. *Front Public Health*. 2016; 4.

Kanavos P, Nicod E. What is wrong with orphan drug policies? Suggestions for ways forward. *Value Health*. 2012; 15-8.

Kawalec P, Sagan A, Pilc A. The correlation between HTA recommendations and reimbursement status of orphan drugs in Europe. *Orphan J Rare Diseases*. 2016; 11-122.

Kolasa K, Zwolinski KM, Hermanowski T. Potential impact of the implementation of multiple-criteria decision analysis (MCDA) on the Polish pricing and reimbursement process of orphan drugs. *Orphan J Rare Diseases*. 2016; 11-23.

Laupacis A. Evidence and values: Requirements for public reimbursement of drugs for rare diseases –A case study in oncology. *Commentary Can J Clin Pharmacol*. 2009; 16-2.

López-Bastida J, Oliva J, Antoñanzas F, García-Altés A, Gisbert R, Mar J, Puig-Junoy J. Propuesta de guía para la evaluación económica aplicada a las tecnologías sanitarias. *Gaceta Sanitaria*. 2010; 24-2.

Lloyd A, Gallop K, Hutchings A, Acaster S. How do we estimate quality adjusted life years (QALYS) in rare diseases? A case study in hypophosphatasia. *Value Health*. 2015; 18(7): A651.

Marsh K, Lanitis T, Neasham D, Orfanos P, Caro J. Assessing the value of healthcare using multi-criteria decision analysis: A review of the literature. *Pharmacoeconomics*. 2014 ; 32(4): 345-65.

Marsh K, Ijzerman M, Thokala P, Baltussen R, Boysen M, Kaló Z, Lönngren, T., Mussen, F., Peacock, S., Watkins, J., and Devlin, N. Multiple Criteria Decision Analysis for Health Care Decision Making—Emerging Good Practices: Report 2 of the ISPOR MCDA Emerging Good Practices Task Force. *Value Health*. 2016; 19(1): 1-13.

- McCabe C, Tsuchiya A, Claxton K, Raftery J. Assessing the economic challenges posed by orphan drugs: A comment on Drummond et al., *Int J Technol Assess Health Care*. 2007 Summer;23(3):397-401.
- Mincarone P, Leo CG, Sabina S, Sarriá-Santamera A, Taruscio D, Serrano-Aguilar PG, Kanavos P. Reimbursed price of orphan drugs: Current strategies and potential improvements. *Public Health Genomics*. 2017; 20(1): 1-8.
- Mühlbacher AC, Kaczynski A. Making good decisions in healthcare with multi-criteria decision analysis: The use, current research and future development of MCDA. *Appl Health Econ Health Policy*. 2016; 14(1): 29-40.
- Palaska C, Hutchings A. Value assessment and pricing frameworks for rare diseases treatments: New approaches from the literature. *Value Health*. 2015; 18(7): A678.
- Paulden M, Stafinski T, Menon D, McCabe C. Value-based reimbursement decisions for orphan drugs: a scoping review and decision framework. *Pharmacoeconomics*. 2015; 33(3): 255-69.
- Radaelli G, Lettieri E, Masella C, Merlino L, Strada A, Tringali M. Implementation of EUNETHTA core model in Lombardia: The VTS framework. *Int J Technol Assess Health Care*. 2014; 30(1): 105-12.
- Ram C, Montibeller G, Morton A. Extending the use of scenario planning and MCDA for the evaluation of strategic options. *J Oper Res Soc*. 2011; 62(5): 817–829.
- Schey C, Conolly M. Multi-criteria decision analysis (MCDA): testing a proposed MCDA model for orphan drugs. *Value Health*. 2014; 17(7): A323-A686.
- Schey C, Krabbe PFM, Postma MJ, Connolly MP. Multi-criteria decision analysis (MCDA): testing a proposed MCDA framework for orphan drugs. *Orphanet J Rare Dis*. 2017; 12: 10.
- Schlander M, Garattini S, Holm S, Kolominsky-Rabas P, Nord E, Persson U, Richardson J, Simoens S, de Solà Morales O, Tolley K, Toumi M. Incremental cost per quality-adjusted life year gained? The need for alternative methods to evaluate interventions for ultra-rare disorders. *J Comp Eff Res*. 2014; 3(4): 399-422.
- Schlander M, Garattini S, Kolominsky-Rabas P, Nord E, Persson U, Postma M, Richardson J, Simoens S, de Solà Morales O, Tolley K, Toumi M. Determining the value of medical technologies to treat ultra-rare disorders: a consensus statement. *J Mark Access Health Policy*. 2016; 27: 4.
- Schuller Y, Hollak C.E.M, Biegstraaten M. The quality of economic evaluations of ultra-orphan drugs in Europe – a systematic review. *Orphanet J Rare Dis*. 2015; 10: 92.
- Solon C, Kanavos P. An analysis of HTA decisions for orphan drugs in Canada and Australia. London School of Economics. 2015.
- Sussex J, Towse A, Devlin N. Operationalising value based pricing of medicines: A taxonomy of approaches. *Pharmacoeconomics*. 2013; 31(1): 1-10.
- Sussex J, Rollet P, Garau M, Schmitt C, Kent A, and Hutchings, A. A pilot study of multi-criteria decision analysis for valuing orphan medicines. *Value Health*. 2013; 16(8): 1163-9.
- Sussex J, Rollet P, Garau M, Schmitt C, Kent A, Hutchings A. Multi-criteria decision analysis to value orphan medicines. Office of Health Economics. 2013c; 13-3.
- The EVIDEM Collaboration. 2015a. EVIDEM v3.0.
- The EVIDEM Collaboration. 2015b. EVIDEM v3.1.
- Thokala P, Devlin N, Marsh K, Baltussen R, Boysen M, Kalo Z, Longrenn T, Mussen F, Peacock S, Watkins J, Ijzerman M. Multiple Criteria Decision Analysis for Health Care Decision Making—An Introduction: Report 1 of the ISPOR MCDA Emerging Good Practices Task Force. *Value Health*. 2016; 19(1): 1–13.
- Tony M, Wagner M, Khoury H, Rindress D, Papastavros T, Oh P, Goetghebeur MM. Bridging health technology assessment (HTA) with multi-criteria decision analyses (MDCA): field testing of the EVIDEM framework for coverage decisions by a public payer in Canada. *BMC Health Services Research*. 2011; 11-329.

Tromp N, Baltussen R. Mapping of multiple criteria for priority setting of health interventions: an aid for decision makers. *BMC Health Serv Res.* 2012; 13(12): 454.

Tsoi B, Masucci L, Campbell K, Drummond M, O'Reilly D, Goree, R. (2013). Harmonization of reimbursement and regulatory approval processes: a systematic review of international experiences. *Expert Rev Pharmacoecon Outcomes Res.* 2013; 13(4): 497-511.

van Til J, Groothuis-Oudshoorn C, Lieferink M, Dolan J, Goetghebeur M. Does technique matter; a pilot study exploring weighting techniques for a multi-criteria decision support framework. *Cost Eff Resour Alloc.* 2014; 18;12:22.

Wagner M, Khoury H, Willet J, Rindress D, Goetghebeur M. Can the EVIDEM Framework tackle issues raised by evaluating treatments for rare diseases: Analysis of issues and policies, and context-specific adaptation. *Pharmacoeconomics.* 2016; 34(3): 285-301.

Wagner M, Khoury H, Bennetts L, Berto P, Ehreth J, Badia X, Goetghebeur M. Appraising the holistic value of lenvatinib for radio-iodine refractory differentiated thyroid cancer: A multi-country study applying pragmatic MCDA. *BMC Cancer.* 2017 17;17(1):272.

Wahlster P, Goetghebeur M, Kriza Ch, Niederländer Ch, Kolominsky-Rabas P. Balancing costs and benefits at different stages of medical innovation: A systematic review of Multi-criteria decision analysis (MCDA). *BMC Health Serv Res.* 2015; 9(15): 262.

Winqvist E, Bell CM, Clarke JTR, Evans G, Martin J, Sabharwal M, Gadhok A, Stevenson H, Coyle D. An evaluation framework for funding drugs for rare diseases. *Value Health.* 2012; 15(6): 982-6.



MCDA IN DECISION-MAKING: EXPERIENCES IN THE SPANISH HEALTHCARE FRAMEWORK

Xavier Badia Llach

Antoni Gilabert-Perramon

CHAPTER

9

9.1. THE NEED TO MAKE BETTER DECISIONS IN THE SELECTION, EVALUATION AND POSITIONING OF DRUGS

Decision-making is an activity which is inherent in the daily lives of healthcare professionals, whether in the field of clinical assistance or in management. Making decisions means having to choose between different options. Having to choose one option or another is one of the most difficult tasks which a manager faces, since few things are white or black, and one does not always have all the information to be able to decide. So the uncertainty surrounding any decision is the main cause of blockage or inhibition when one is faced with a problem which calls for a solution. Not having all the knowledge or evidence about what we have to decide greatly complicates any decision, especially if it has effects on In the field of drugs, decisions about whether or not to gain access to public financing, and under what conditions this access is given, are of great importance, not only for the manufacturing companies but also for the target population, both in terms of health and economically. And there, all the management structures, from the 'macro' at ministerial level, the 'meso' at regional level, and the 'micro' at the level of the healthcare centre or healthcare professional have a great responsibility when it comes to deciding what is financed and what is not, and at what price, as well as to whom it is given and under what circumstances.

Making these types of decisions is not easy, since the information and evidence necessary to take them, with all the certainty that we would want, is not always fully available. The greater the uncertainty, the more difficult it is to face the situation and make the right decision. Faced with this type of problem there are two possible types of approach, one which should be avoided, and another one which would be the most appropriate.

The first thing to avoid is the concentration of decisions in an environment of little transparency, leaving high levels of freedom to decision-makers, be they macro, meso or micro managers, since transparency is necessary for good public management (United Nations, 2016). This first situation, which, written in this way, may scandalise many people, is not difficult to find in the field of regulated models with legal norms which are either not developed properly or not strictly complied with (CNMC, 2015; AES, 2016), which results in a kind of decision-making with certain amount of variability in the best of the cases, or with certain degrees of arbitrariness in the worst cases.

The second approach consists of arming oneself with an evaluation methodology which, in a standardised way, assures us of internal and external validity when making decisions. This is where multi-criteria decision analysis (MCDA) can help decision-makers to make the most suitable decisions, which are, at the same time, the most coherent and consistent with other evaluations of other drugs made separately or in different periods.

It is logical to think that this can be solved simply with the traditional evaluation of efficacy, safety and costs, which is also correct. However, in the traditional evaluation many of the criteria used in the MCDA are implicit, but they are not used explicitly, and not always in the same way. For example, when this type of question is presented to an evaluating group, the most common response is to say that "of course, concepts such as severity or unmet need have been taken into account," even when there are no explicit criteria. The definition of these criteria and their inclusion in the evaluation in a standardised way is an opportunity for improvement which enables one to make decisions with the assurance that all the relevant

aspects have been taken into account in the same way in each evaluation.

So the use of the structured methodology of MCDA can help us, as a minimum:

- To decide, on a consensual basis, which explicit criteria will be used to evaluate a particular drug in order to make decisions about its financing or use.
- To define each criterion in a clear, precise and transparent way to help to apply it in the same way in each and every evaluation.
- To generate a debate among the evaluators about the relative weights of each criterion.
- To establish the weight which will be given to each criterion. A weight which does not need to make much sense in its absolute value, but, rather, as regards its relative position with respect to the other criteria.
- To make decisions, after taking account of each and every criterion of a standardised form for the different evaluations, allowing not only the external validity of the decision, but also the internal validity in terms of homogeneity and reproducibility.

9.2. THE MULTI-CRITERIA DECISION ANALYSIS AND THE NEED TO MAKE DECISIONS BASED ON A REFLECTIVE FRAMEWORK

The MCDA is not only an evaluation tool, it is also used to support the culture of appropriate decision-making by promoting a standardised and legitimate procedure. This includes the representative selection of decision-makers, the importance of the reasoning for the decision, transparency, and the degree of interest and implementation, and all this should be based on an ethical framework which serves to render account for the actions and decisions of healthcare services (Goetghebeur et al., 2008).

To ensure that decisions are based on important reasons (substantive legitimacy), some frameworks provide a series of generic decision criteria derived from ethical imperatives which are implicit in the objective of health care (EVIDEM, 2018; FIFARMA, 2016). This represents a generic interpretative framework (reflective MCDA) which can be used either to obtain individual values and facilitate the exchange of diverse perspectives during the deliberations of committees which try to analyse the benefits in the National Health System, or for other purposes (e.g. shared decisions between patient and doctor). These generic criteria can continue to be specified to reflect specific therapeutic areas such as cancer, rare diseases or specific interventions. The EVIDEM framework also provides a common structure, so that all the members who have to contribute to the decision-making can express their interpretation of the scientific evidence for each criterion and, in this way, share their reasoning with other evaluators or decision-makers. These interpretations can be expressed quantitatively through interpretative scorings, qualitatively through impacts (qualitative criteria), as well as narratively through comments (all criteria). All information is limited to make fairer decisions based on the available information and common sense.

So the ideas on which MCDA is based are that the current processes must be changed and the limitations of the reflection and decision processes must be reduced, ensuring that all relevant criteria are included in the decision-making process (regardless of whether they are considered qualitatively or quantitatively); that scientific and relevant non-scientific evidence should be available through an efficient process of synthesis; that the apparent validity (common sense) should be reviewed at each step of the process (corresponding weights, scores and narratives, aggregate measures) to ensure that the visual representations of the quantitative results reflect the reasoning of individuals or, for the taking of decisions at system level (e.g. national, regional or local), of the commission, committee or panel within and throughout the evaluation process.

9.2.1. The importance of the criteria used in the MCDA appraisal according to the decision-makers

The criteria included in an MCDA may vary depending on the different objectives of the institutions or committees, and may also be different in terms of their content and importance (Thokala 2016). For example, a national commission will use criteria for price and financing at the national level; in a regional committee the criteria have to be as wide as possible in order to analyse, in a holistic manner, the value for the population covered by the region; and in a pharmaco-therapeutic committee, the most relevant criteria for the selection and therapeutic positioning of a drug in the hospital setting should be used (Badia et al., 2017).

In Spain, there have been some recent experiences in view of the possibility of incorporating the MCDA methodology in the evaluation of drugs, and which we review below.

For example, the Spanish Agency for Medicines and Health Products [*Agencia Española de Medicamentos y Productos Sanitarios (AEMPS)*] produces therapeutic positioning reports (TPRs) to inform about the national therapeutic positioning of a new drug, and to support price and financing decisions (AEMPS, 2013). The reports mainly use the criteria of comparative effectiveness, comparative safety, and use and monitoring of the drug. In a recent study carried out among the evaluators of the AEMPS who produced the TPRs (Hernández et al., 2017), the importance of the criteria included in the EVIDEM framework was rated from 1 (not important at all) to 5 (very important).

The most important criteria were comparative efficacy and safety, and the severity of the disease. The size of the population, the non-medical costs, and the consensus of experts or clinical guidelines (Table 9.1) were considered as being of little importance. It was recognised that some criteria that were also considered important, such as the severity of the disease, the quality of evidence and the needs not covered by the new drug, were not explicitly considered in the TPR. Subsequently, a scoring exercise was carried out for alternative drugs in the treatment of psoriasis, and it was found that the evidence matrix was very useful to explain the basis of each criterion. A high value was given to the discussion generated by the scores awarded to each drug for each of the criteria.

TABLE 9.1. DATA RESULTING FROM THE WEIGHTING OF THE QUANTITATIVE CRITERIA OF THE EVIDEM FRAMEWORK BY THE EVALUATORS OF THE SPANISH AGENCY FOR MEDICINES AND HEALTH PRODUCTS

| DIMENSION | CRITERIA | MEAN | STANDARD DEVIATION | MINIMUM -MAXIMUM |
|---------------------------------------|---|------|--------------------|------------------|
| Need for intervention | Disease severity | 4.5 | 0.6 | 3-5 |
| | Size of affected population | 2.9 | 1.4 | 1-5 |
| | Unmet needs | 4.1 | 1.0 | 2-5 |
| Comparative out-comes of intervention | Comparative effectiveness | 4.5 | 0.7 | 3-5 |
| | Comparative safety/tolerability | 4.3 | 0.9 | 2-5 |
| | Comparative patient-perceived outcomes | 3.0 | 0.9 | 1-4 |
| Type of benefit of intervention | Type of therapeutic benefit | 4.1 | 0.6 | 3-5 |
| Economic consequences of intervention | Comparative cost of the intervention | 3.4 | 1.0 | 2-5 |
| | Comparative other medical costs | 3.2 | 0.9 | 2-5 |
| | Comparative non-medical costs | 3.0 | 1.0 | 1-4 |
| Knowledge about the intervention | Quality of evidence | 4.0 | 0.7 | 3-5 |
| | Expert consensus / Clinical practice guidelines | 2.8 | 1.1 | 1-5 |

Source: Hernández et al., 2017.

In relation to decision-making at regional level, table 9.2 shows the average of the ratings given by the decision-makers in the regional committees of Andalusia, the Basque Country and Catalonia (in Catalonia the committee included members of PASFTAC) (Gilabert-Perramon et al., 2016). The criteria most valued by the regional evaluators were the severity of the disease, the comparative efficacy/effectiveness, the quality of the evidence, the comparative safety/tolerability and the type of therapeutic benefit. The least valued criteria were other non-medical costs, the size of the affected population and other medical costs, but all the criteria were regarded as important because their average scores were all above 3.

Figure 9.1 shows the consistency in the value given to the different quantitative criteria of the EVIDEM framework by the different evaluating committees, with the relevant exception of the criterion of size of the affected population. Account should be taken of the fact that in Catalonia the participants were the members of the Committee for the Evaluation of Orphan Drugs, so this criterion was not considered relevant since orphan drugs are intended to treat diseases with a low prevalence (<50/100,000 inhabitants), and even a high weight in this

criterion could have a negative impact on ultra-rare diseases (prevalence $<2/100,000$ inhabitants) for which unmet needs are high.

TABLE 9.2. DATA RESULTING FROM THE WEIGHTING OF THE QUANTITATIVE CRITERIA OF THE EVIDEM FRAMEWORK AT THE REGIONAL LEVEL IN SPAIN

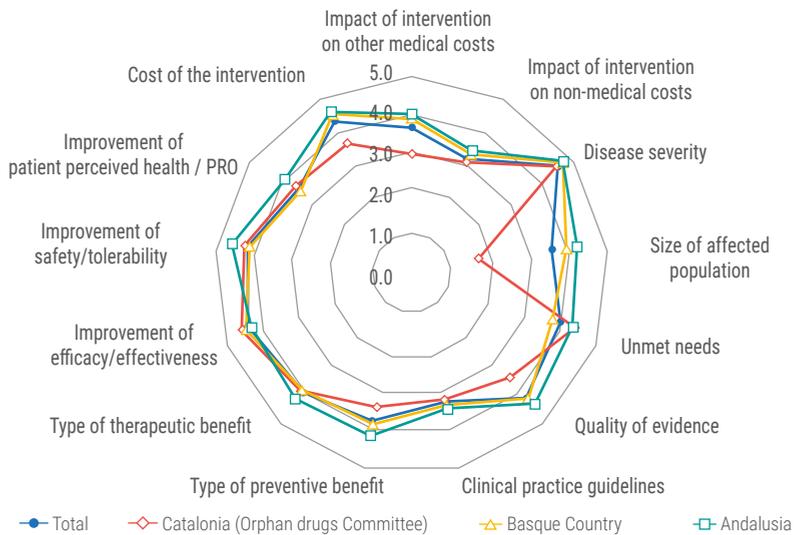
| DIMENSION | CRITERIA | MEAN | STANDARD DEVIATION | MINIMUM -MAXIMUM |
|---------------------------------------|---|------|--------------------|------------------|
| Need for intervention | Disease severity | 4.60 | 0.53 | 3-5 |
| | Size of affected population | 3.42 | 1.37 | 1-5 |
| | Unmet needs | 4.20 | 0.81 | 2-5 |
| Comparative outcomes of intervention | Comparative effectiveness | 4.50 | 0.74 | 2-5 |
| | Comparative safety/tolerability | 4.34 | 0.72 | 2-5 |
| | Comparative patient-perceived outcomes | 3.64 | 0.94 | 1-5 |
| Type of benefit of intervention | Type of preventive benefit | 4.13 | 0.91 | 2-5 |
| | Type of therapeutic benefit | 4.32 | 0.59 | 3-5 |
| Economic consequences of intervention | Comparative cost of the intervention | 4.14 | 0.99 | 1-5 |
| | Comparative other medical costs | 3.52 | 1.01 | 1-5 |
| | Comparative non-medical costs | 3.19 | 1.04 | 1-5 |
| Knowledge about the intervention | Quality of evidence | 4.34 | 0.80 | 2-5 |
| | Expert consensus / Clinical practice guidelines | 3.63 | 1.13 | 1-5 |

Source: Gilabert-Perramon et al., 2016.

Nevertheless, a recent study in oncology conducted by a group of presidents of patients' associations (Aguarón et al., 2017) showed that for other groups the importance of the criteria may be different. The most important criteria for the patients were the impact on the patient's quality of life and well-being, the comparative effectiveness through, for example, the correct and validated integration of systems for analysing the results perceived by patients, and the unmet medical needs. There was also a lot of discussion about other criteria such as prevention and risk reduction, with the cost of treatment or budgetary impact being valued as having little importance.

Regarding the MCDA, it was considered that it provides a well structured framework, sufficiently specific and open to the participation of all the parties involved to be seriously taken into account.

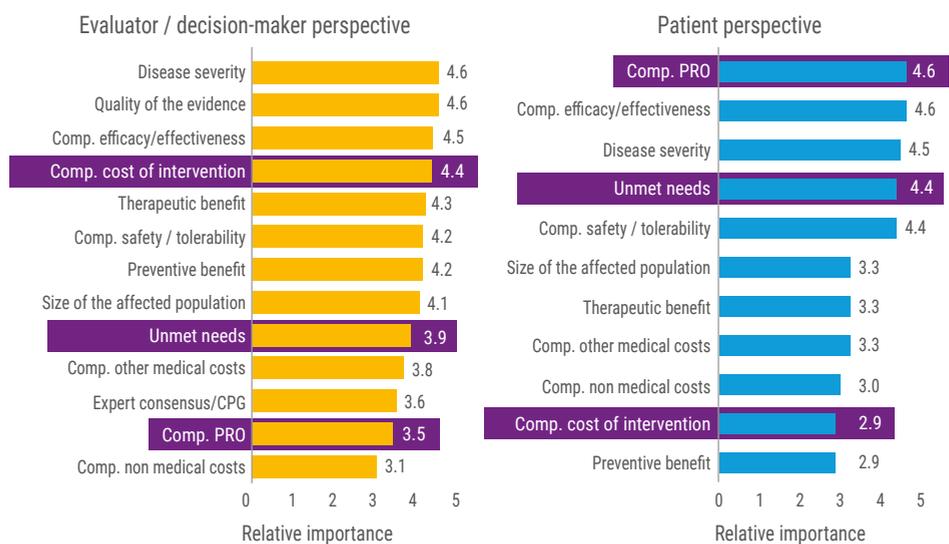
FIGURE 9.1. WEIGHTING OF QUANTITATIVE CRITERIA OF THE EVIDEM FRAMEWORK AT REGIONAL LEVEL IN SPAIN, ACCORDING TO THE EVALUATION COMMITTEE IN ANDALUSIA, THE BASQUE COUNTRY AND CATALONIA



Source: Gilabert-Perramon et al., 2016.

In figure 9.2 there is a comparison of the importance of the criteria from the perspective of patients with their importance from the perspective of the evaluators/decision-makers at national

FIGURE 9.2. IMPORTANCE OF THE CRITERIA FROM THE PERSPECTIVE OF THE PATIENT AND FROM THE PERSPECTIVE OF NATIONAL AND REGIONAL EVALUATORS IN SPAIN



Source: Aguaron et al., 2017.

and regional level in Spain. The evaluators and decision-makers gave greater importance to the cost of the intervention and a lower importance to the outcomes reported by the patient and to the unmet needs, whereas the patients scored the criteria in the opposite way, giving the greatest importance to the health reported by patients and less importance to the cost of the intervention (Aguaron et al., 2017).

This comparison shows the differences between the perspectives of the different stakeholders of the system when it comes to giving value to criteria related to making decisions about healthcare, and demonstrates the importance of having a tool which enables us to take account of all points of view and to make decisions from a holistic point of view, without excluding any group from the process.

9.3. EVIDENCE-BASED INFORMATION: THE REFLECTIVE SCORING MATRIX TO INFORM THE DISCUSSION OF THE DECISION

One of the most valued aspects of MCDA in decision-making is that it requires the systematic explanation of each committee member's reasoning which is behind the scores which they assign to the information about the value of a drug during its evaluation (Goetghebeur et al., 2008). It also makes explicit and shares the subjective conditioning factors which often influence decision-making.

In the two experiences initiated in Spain, there are remarkable coincidences. Within the framework of the CatSalut Drug-Therapeutic Committee (Gilbert-Perramon et al., 2017), it can be seen that this type of methodology helps one to understand the reason why a certain decision is taken, providing contextualisation, solidity and coherence, and improving traceability. The EVIDEM method which has been implemented by CatSalut (Generalitat de Catalunya, 2016) since September, 2017, "is easy to apply, being a good means of communication between the members of the Pharmacotherapeutic Committee, and helps to see why a certain decision is made", and "it is working well because it focuses the discussion, and we see what points of conflict there are after the technical evaluation."

A recent experience with the multi-criteria analysis by the Pharmacotherapeutic Committee of the University Hospital Virgen de la Macarena, not yet published, regards the methodology of the MCDA as positive because it allows us to detail the criteria, which are not always clear, at the same time as we perform the evaluation, and to understand the reason why a certain decision is made. In addition, this methodology allows us to include the opinions of all members of the Pharmacotherapeutic Committee, so all opinions are taken into account in the decision-making.

MCDA is an option which satisfies the need for all the parties involved to participate in healthcare decisions (funders, clinicians, pharmacists and patients, among others).

9.4. THE RELATION OF THE MCDA EVALUATION TO DECISION-MAKING

Making decisions about healthcare is becoming more complex every day due to the change which are taking place in the healthcare sector, with the continuous progress in health technologies, new, innovative drugs, new therapeutic areas, combinations of drugs, and the appearance of health technologies which represent a disruptive innovation. This makes it

necessary for the way in which decisions are made to evolve as the complexity of our environment increases.

MCDCA represents a methodology and a general framework to support decision-making, by considering in a standardised, orderly, transparent and explicit way the different and multiple criteria that are relevant for decision-making in healthcare, adapting to the needs of the current environment for evaluation and decision-making, taking into account other criteria besides traditional ones, such as health outcomes in real-life situations of therapeutic effectiveness and the results reported by the patient. This can be particularly relevant in specific fields such as oncology and rare diseases, where continuous advances involve novel diagnostic methods, such as the early detection of various types of cancer by means of liquid biopsy, thanks to the study of immuno-oncology, or a change in the clinical prognosis of many diseases which used to be fatal but have now become chronic (Badia et al., 2017).

MCDCA offers an essential input to decision-making, ensuring that all the relevant factors, including the always important contextual factors, such as the ability of the system to incorporate innovation or opportunity cost, are taken into account in decision-making and communication of the value contributed by a drug. And this compels one to structure and make explicit the reasoning which underlies the decision-making, and imposes a dialogue between the different parties involved in it.

One question to ask is whether or not the MCDCA should end with a quantitative score that is directly applied in the decision-making. In our opinion, the MCDCA should be used to inform decision-making, and the quantification of the value of a drug is only a numerical way which is used to know and provoke discussion of the reasoning behind the number. In current decision-making, different forms of classification are used, involving the decision about a new drug or health technology, depending on the evaluation group. For example, the recommendations or criteria for the use of, access to, and supply of the hospital medicines of CatSalut (CatSalut, 2012), in which the drugs are classified under "individualised authorisation, direct verification of the clinical criteria of indication by professionals and healthcare centres, follow-up and therapeutic response according to reports and opinions of committees of experts, and drugs with indications of exceptional use"; or the classification system of the Joint Committee for the Evaluation of New Drugs (CMENM, 2007), in which drugs are classified into the following five categories: 'non-assessable', 'not a therapeutic advance', 'contributes in concrete situations', 'moderate therapeutic improvement', and 'important therapeutic improvement.'

The MCDCA could standardise the criteria according to which healthcare decisions are made, providing a better positioning of the drug in the form, or attributing a higher price to the technologies of greater value, a value attributed quantitatively and qualitatively.

9.5. CONCLUSIONS

Although it is still early to know the exact relationship between the input given by MCDCA to decision-making and the making of those decisions, what is clear is that the reflective MCDCA methodology is designed to support the culture of informed decision-making by promoting a procedure with the necessary legitimacy. This includes the representative selection of the decision-makers taking into account the perspectives of all the parties involved in healthcare decision-making, the importance of the reasoning for the decision taken, and the

transparency. It provides not only the external validity of the decision but also its internal validity in terms of homogeneity and reproducibility

REFERENCES

Agencia Española de Medicamentos y Productos Sanitarios (AEMPS). Propuesta de colaboración para la elaboración de los informes de posicionamiento terapéutico de los medicamentos. 2013.

Aguarón A, Fernández, A, Gimón A, Nafria B, Gaspar B, Guarga L, Gálvez M, Fuentes M, Saldaña R, and Badia, X. (2017). Patient Empowerment for health care decision making assessing oncology innovation through Multi-Criteria Decision Analysis in Spain. *Value Health*. 2017, 20: A451.

Asociación de Economía de la Salud. Boletín Economía y Salud, num 87, November 2016.

Badia X, Poveda-Andrés JL, Merino-Montero S. Análisis descriptivo de los criterios usados en los informes de evaluación de fármacos en España. *Rev Esp Econ Salud*. 2017; 12(1).

Badia X, Pontes C, Fontanet M, Obach M, Vayano A, Torrent J, Guarga L. Development of an specific evaluation framework for orphan drugs based on MCDA for health care decision making in Catalonia. *Value Health*. 2017; 20: A674.

CatSalut. Servei Català de la Salut. Instrucció 01/2012, de 10 de gener 2012. Disponible en: http://catsalut.gencat.cat/web/.content/minisite/catsalut/proveidors_professionals/normatives_instruccions/any_2012/instruccio_01_2012_10_gener/instruccio_01_2012.pdf.

Comisión Nacional de los Mercados y la Competencia. Proyecto de RD por el que se regula la financiación y fijación de precios de medicamentos y productos sanitarios y su inclusión en la prestación farmacéutica del SNS. Available from: https://www.cnmc.es/sites/default/files/757377_12.pdf.

Comité Mixto de Evaluación de Nuevos Medicamentos (CMENM). Categorías de calificación de los nuevos medicamentos según su innovación terapéutica. 2007.

EVIDEM: Evidence and Values: Impact of Decision Making. A Reflective framework to advance Health. Available from: www.evidem.org.

FIFARMA. Utilización del ADMC para apoyar la toma de decisiones.2016. Available from: <http://www.fifarma.org>.

Generalitat de Catalunya, Departament de Salut. Pla de Salut de Catalunya 2016-2020: Un sistema centrat en la persona: públic, universal i just.

Gilbert-Perramon A, Betolaza J, March J, Espín J, Merino-Montero S, Gil A, Badia X. Multi-Criteria Decision Analysis (MCDA): common tools for different needs supporting healthcare decision making in Spain. ISPOR 19th Annual European Congress 29 October-2 November 2016, Vienna, Austria.

Gilbert-Perramon A, Torrent-Farnell J, Catalan A, Prat A, Fontanet M, Puig-Peiró R, Merino-Montero S, Khoury H, Goetghebeur MM, Badia X. Drug evaluation and decision making in Catalonia: development and validation of a methodological framework based on multi-criteria decision analysis (MCDA) for orphan drugs. *Int J Technol Assess Health Care*. 2017; 33(1): 111-120.

Goetghebeur MM, Wagner M, Khoury H, et al. Evidence and value: impact on Decision Making—the EVIDEM framework and potential applications. *BMC Health Serv Res*. 2008; 8: 270.

Hernández C, Blázquez A, Gil A, Badia X. Value of Reflective multi-Criteria Decision Analysis (MCDA) drug evaluation among therapeutic positioning report evaluators from the Spanish Agency of medicines (AEMPS). *Value Health*. 2017; 20: A699.

United Nations. Report of the United Nations Secretary-General's high-level panel on access to medicines. September 2016. Available from: <http://bit.ly/2fVxnll>.

Thokala P, Devlin N, Marsh K, Baltussen R, Boysen M, Kalo Z, et al. Multiple Criteria Decision Analysis for Health Care Decision Making—An Introduction: Report 1 of the ISPOR MCDA Emerging Good Practices Task Force. *Value Health*. 2016; 19(1): 1–13.



A PRACTICAL CASE OF MCDA APPLIED TO RARE AND PREVALENT DISEASES

Bleric Alcalá Revilla

Néboa Zozaya González

CHAPTER

10

10.1. INTRODUCTION AND OBJECTIVES

Following on from the review of the theoretical framework of the MCDA and its application in different pathologies, this chapter includes a practical case applied to a prevalent disease and to a rare disease, as an example. Specifically, an Expert Committee coordinated by Weber carried out an MCDA to evaluate three fictitious drugs aimed at two different pathologies in Spain.

The objective of the exercise was dual. On the one hand, we tried to compare, as far as possible, the final value obtained for the evaluated drugs, validating any differences between a prevalent pathology and a rare disease, as well as the degree of sensitivity of the MCDA to changes in the proposed scenarios. On the other hand, we were looking to the experts to go through all the phases of the process, so that we could gather, at first hand, their reflections on the methodology and its implications, making the debate richer and more substantiated.

Throughout this chapter we review the methodology followed in the case study, we present the results obtained and we propose some reflections arising from it.

10.2. METHODOLOGY

The case study was proposed to contrast the results of MCDA in a pathology of high prevalence, such as chronic obstructive pulmonary disease (COPD), with the results in a rare or infrequent disease, represented in this case by mucopolysaccharidosis (MPS). The real characteristics of these diseases were taken into account, while the drugs to be evaluated were presented under different fictitious assumptions, and in several alternative scenarios.

The exercise was performed following the EVIDEM framework, which, as mentioned in previous chapters, is one of the MCDA methodologies most used in the healthcare sector, and has a validated procedure and standard decision criteria. For practical purposes, only the 13 quantitative criteria covered by the EVIDEM 4.0 (EVIDEM Collaboration, 2017) methodology were considered, excluding the qualitative or reflective ones of the analysis.

In this section, the characteristics of the selected diseases and the fictitious drugs are described, and the steps followed in carrying out the exercise are described.

10.2.1. Selected diseases

The pathologies selected for the practical exercise were, on the one hand, COPD, in its serious or very serious condition, as a prevalent disease, and mucopolysaccharidosis as a rare disease, evaluating one of its most attenuated forms (MPS I: Scheie) and one of its most serious forms (MPS VII: Sly). Two extreme variants of the MPS were selected to try to collect the greatest possible spectrum of variability. Next, the main characteristics of the pathologies in the exercise are briefly described.

Severe COPD

COPD is one of the diseases with the highest prevalence and morbi-mortality. According to the WHO, approximately 64 million people worldwide have COPD (WHO, 2004). In Spain, it affects 10.2% of the population aged 40-80 years (mild COPD 16%, moderate 46%, severe

28% and very severe 11%) (Soriano et al., 2015). The disease is characterised by a chronic inflammatory response of the respiratory tract and a persistent airflow limitation. In 2012, COPD caused 3 million deaths worldwide and by 2030 it is expected to have become the third cause of death (OMS, 2004). In Spain, the annual mortality per 1,000 patients was 21.1 for mild COPD, 39.3 for moderate COPD, 73.2 for severe, and 121.1 for very severe (Soriano et al., 2015). Mortality from COPD is linked to exacerbations, and is higher the more frequent and/or severe the exacerbations, especially if they require hospitalisation (Soler-Cataluna et al., 2005; Granda-Orive et al., 2016; Müllerova et al., 2015).

Currently, there are different pharmacological treatments for COPD, such as long-acting β 2 agonists (LABA), long-acting muscarinic antagonists (LAMA) and inhaled corticosteroids (IC). None of them is yet able to cure the disease, so the objectives of the current treatment are based on reducing symptoms (improving the state of health, quality of life and tolerance to exercise) and the risk of future occurrences (preventing the progression of the disease, preventing and treating exacerbations and reducing mortality). There are different clinical practice guidelines that are updated regularly with pharmacological innovations (Miravittles et al., 2017; Molina París et al., 2017; Global Initiative for Chronic Obstructive Lung Disease, 2017).

Finally, owing to its chronic and progressive nature, severe COPD implies a high cost in terms of use of healthcare resources. In Spain, the mean estimated direct annual healthcare cost of managing one patient with severe COPD is approximately €9,850 (of which €900 corresponds to drugs), to which should be added more than €5,000 per year for non-healthcare (informal care) direct costs, in addition to the labour losses involved (estimated at about €300) (Izquierdo, 2003).

Mucopolysaccharidosis

Mucopolysaccharidoses are a group of rare diseases (global incidence of one in 22,500 inhabitants) which are hereditary, multisystemic and progressively metabolic. They are included among the lysosomal or depositional diseases (Muenzer, 2011; Lampe et al., 2015) caused by the absence or malfunction of certain enzymes necessary for the processing of molecules called glycosaminoglycans (AECOM and SEPEAP, 2015). MPS progress as the storage of glycosaminoglycans affects the bones, the skeletal structure, the connective tissues and other organs. Neurological complications may include neuronal damage, as well as pain and impairment of the motor function (Khan et al., 2017; AECOM and SEPEAP, 2015).

There are six types of MPS depending on the enzymatic defect, with different forms of clinical presentation, and which present severity levels. Thus, more severe forms and more attenuated forms are distinguished mainly by the age of presentation, the progression of the disease, the degree of neurological involvement and the associated life expectancy. For the case study, one of the most attenuated forms of MPS (the Scheie syndrome, which is within MPS type I), and one of the most severe forms (the type VII, or Sly syndrome) were selected (Table 10.1). Despite its severity, the Scheie syndrome presents mild neurological involvement, and patients reach adulthood, whereas patients with the Sly syndrome have a life expectancy of 15 years and often present facial dysmorphism, cardiomyopathy and cognitive impairment, among other disorders.

TABLE 10.1. TYPES OF MPS

| TYPE | NAME | CLASSIFICATION | SEVERITY | LIFE EXPECTANCY |
|---------|--------------------------|--|--|--|
| MPS I | Hurler | Classic storage disease | Progressive and severe neurological involvement. Cardiorespiratory failure. Mild neurological involvement. | Between 7 and 10 years. |
| | Hurler-Scheie | | | ≈ 20 years. |
| | Scheie | Mild or without neurological involvement | Adulthood | |
| MPS II | Hunter | Classic storage disease | Neurological damage and cardiorespiratory failure. | ≈ 20 years. Mild: Adulthood. |
| MPS III | Sanfilipo (type A,B,C,D) | Neurodegenerative disease | Progressive and limiting disease with predominant involvement of the central nervous system. | At 20-3-0 years old, vegetative slate with total dependence. |
| MPS IV | Morquio (type A and B) | Bone disease | Multisystemic disease, major progressive skeletal-joint abnormality with normal intelligence. | Severe: Between 20 and 40 years. Mild: Between 60 and 70 years. |
| MPS VI | Maroteaux-Lamy | Classic storage disease | Progressive disease, with involvement of the skeletal and cardiopulmonary system, cornea, liver, spleen, brain and meninges. | Severe: ≈ 15 years Mild: between 50 and 60 years. |
| MPS VII | Sly | Classic storage disease | Early fatality: <i>Hydrops fetal</i> , facial dysmorphism, myocardiopathy, organomegaly. Intermediate: "Hurler" phenotype, dysostosis, cognitive impairment, organomenaly, hepatic cholestasis. | Fatal: foetal death/early life. |
| | | | Mild: Dysostosis. | Intermediate: ≈ 15 years. Mild: ≈ 37 years. |

Source: author's preparation from Clinical Practice Guidelines for MPS (AECOM and SEPEAP, 2015).

The different types of MPS are also distinguished by the types of therapeutic alternatives that they have. Currently, there is the possibility of substitute enzyme treatment (SET) only for MPS I, II, IV and VI. There is no treatment for MPS VII, and its only option is bone marrow transplant. In the next few years, some advances are expected in this regard, together with intrathecal therapies for MPS I, II and IIIA (in the trial phase) and gene therapy (Noh and Lee, 2014; Valayannopoulos and Wijburg, 2011). At the moment, multidisciplinary support is still the most important treatment to ensure that the patient has the best quality of life possible. According to clinical practice guidelines, these measures must address both the different aspects of the condition and the patient's general needs for integration into daily life (AECOM and SEPEAP, 2015).

Additionally, MPS are associated with very high costs, in terms of both healthcare resources (the replacement enzyme therapy alone costs more than €75,000 per patient per year) and non-healthcare resources, derived from the personal care that they require (the cost of which has been calculated at between €40,000-55,000 per year) (Péntek et al., 2016).

10.2.2 Steps followed to carry out the implementation of the MCDA

The steps followed in the implementation of MCDA are described below, and are governed by the principles laid down by the EVIDEM framework and the best practice guidelines (EVIDEM Collaboration, 2017).

- 1. Selection of the stakeholders involved.** The evaluation committee was made up of nine experts in health economics, with academic, clinical and management profiles, and most of them were familiar with MCDA. Specifically, the members of the Expert Committee were (in alphabetical order): Xavier Badía Llach, Carlos Campillo Artero, Jaime

Espín Balbino, Antonio J. García Ruiz, Antoni Gilibert Perramon, Javier Mar Medina, Carlos Martín Saborido, Jaume Puig Junoy and Marta Traperó Bertran. WEBER was in charge of the technical part of the project, which included the methodological and evidence review, the preparation of materials, the coordination of the Committee and the analysis and presentation of the results.

- 2. Contextualisation of interventions and compiling of available evidence.** WEBER compiled information about the current situation in severe COPD, MPS type I (Scheie syndrome) and MPS type VII (Sly syndrome) in the Spanish context, based on the evidence available in the literature. The main bibliographical sources used for the case study were those referred to at table 10.2. With this information, the case study for the evaluation of fictitious drugs was performed (Table 10.2), serving as a guide for the discussion among the experts during the face-to-face meeting. Three fictitious drugs were evaluated that improved the current situation in the three pathologies, being more effective than the available treatments in terms of efficacy and outcomes perceived by patients, although at a higher cost.

TABLE 10.2. CURRENT SITUATION OF COPD, MPS I AND MPS VII IN THE 13 CRITERIA TO BE EVALUATED

| | | CURRENT SITUATION | | |
|---------------------------------------|--|--|---|--|
| | | COPD | MUCOPOLYSACCHARIDOSIS | |
| | | Severe | MPS-I (less severe) | MPS-VII (very severe) |
| Need for intervention | Criterion | | | |
| | Severity disease | MR : 121 x 1,000 High risk of exacerbation | They can live to adult hood | Survival for a few months or until adolescence |
| | Unmet needs | Very severe : 38 x 10,000 inhabitants Severe : 132 x 10,000 inhabitants | 0.2 x 100,000 inhabitants (93 cases) | < 0.1 x 100,000 inhabitants (100 cases in the world) |
| Size of the affected population | Several Tx with room for improvement | Substitute enzyme therapy | None (bone marrow transplant) | |
| Comparative outcomes of intervention | Comparative effectiveness/efficacy | FEV: 0.5-0.10 L Reduction of exacerbations: $\geq 20\%$ | FVC: +5.6 % 6MWT: +38.1 % | -- |
| | Comparative safety/tolerability | Cardiovascular AEs | Severe AEs in patients with respiratory problems | -- |
| | Comparative patient-perceived outcomes/PRO | HRQoL: Improvement in the SGRQ total score: \geq points from baseline | CHAQ: +57% Pain: passes from 0.93 to 0.56 | -- |
| Type benefit of intervention | Type of preventive benefit | No | No | No |
| | Type of therapeutic benefit | It does not cure it. Reduction of symptoms and risk of future events | It does not cure it SET stops the progression of the disease | -- |
| Economic consequences of intervention | Comparative cost of intervention | LAMA+LABA+ICS: €852 year | TSET (Iaronidase): €78,000/year | -- |
| | Comparative other medical costs | < €5.500 | N/A | Bone marrow transplant: €60,000 |
| | Comparative non-medical costs | €300 labour productivity losses | DNHC: €44,000-€55,000; IC: €8,500 | |

Continue

Continued

| | | | | |
|------------------------------|---|---|-------------------------------|----------------------|
| Knowledge about intervention | Quality of the evidence | Numerous CTs with good quality evidence | 2 CTS (n=45) | There is no evidence |
| | Expert consensus/ clinical practice guidelines | Gas EPOC in Spain; GOLD 2017 | Management Guidelines for MPS | |

FEV: Forced Expiratory Volume. FVC: Forced Vital Capacity. 6MWT: 6-Minutes Walk Test. MR: Mortality Rate. PROs: Patient-Reported Outcomes. Tx: Treatments. HRQoL: Health-Related Quality of Life. SGRQ: Saint George's Respiratory Questionnaire. CHAQ: Child Health Assessment Questionnaire. DNHC: Direct Non-Healthcare Costs. IC: Indirect Costs. SET: Substitute Enzyme Therapy. CTS: Clinical Trials. AEs: Adverse Effects. N/A: Not Available.

Source: authors' analysis based on existing literature. Global Initiative for Chronic Obstructive Lung Disease, 2017. Miravittles et al., 2017. Mannino y Kiri, 2006. Soler-Cataluña et al., 2010. Granda-Orive et al., 2016. Müllerova et al., 2015. Soler-Cataluña, 2005. Barrecheguren et al., 2016. Santibáñez et al., 2016. Soriano et al., 2012. Cazzola y Matera, 2014. Gruffydd-Jones y Jones, 2011. Khakban et al., 2016. Martín Escudero, 2009. AECOM y SEPEAP, 2015. Clarke et al., 2009. Muenzer, 2011. Péntek et al., 2016. Muenzer, 2014. Pérez-López et al., 2017. Khan et al., 2017. Pinto et al., 2003. Tomatsu et al., 2015. Verma et al., 2017. Tylki-Szymanska et al., 2010. Wyatt et al., 2012. BotPlus 2017a, 2017b. Wraith et al., 2004. MSSSI, 2006.

- 3. Weighting of the criteria.** Each expert, from their individual perspective, assigned a weight to each criterion, depending on the relative importance that they attributed to it. The weighting was done in two different ways: using a direct rating scale method, with a scale of between 1 and 5 points (1 having the least relevance and 5 the most), and using the hierarchical point allocation method of sharing out 100 points (distributing 100 points among the 5 domains or groups of criteria, and then 100 points among the criteria which make up each domain – the greater the importance, the more points). This step was performed on line before the results of the evaluated interventions were known, in order not to be influenced by them (Goetghebeur et al., 2008). In this way, the same weighting was applied to each of the three interventions. The results of this stage were shown during the face-to-face meeting.
- 4. Scoring of the interventions.** For each of the three fictitious drugs evaluated, each expert assigned a score to each criterion, individually and confidentially. This exercise was carried out during the face-to-face meeting, based on the evidence provided, the assumptions made (Table 10.3, Table 10.4, Table 10.5 and Table 10.6), the exchange of opinions between experts and the experience and individual perception of each one. For the absolute criteria (those that do not consider comparisons between interventions), the possible scores range from 0 to 5, with 0 being the lowest value and 5 being the highest. For the relative criteria (those that are compared with an alternative), the scale ranges from -5 to 5 to reflect the full range of comparative effects (improvements or worsenings) (EVIDEM Collaboration, 2017).
- 5. Calculation of the total estimated value of the interventions.** The overall value of the interventions evaluated was estimated by combining the relative weighting of each criterion with the score obtained for each of them, and was transformed into a 0-1 scale to facilitate its interpretation. Specifically, an additive linear model was applied from the following formula, as it appears in the EVIDEM methodology (Goetghebeur et al., 2008):

$$V = \sum_{x=1}^n V_x = \sum_{x=1}^n \left(\frac{W_x}{\sum W_n} S_x \right)$$

Where V is the total estimated value. V_x is the value contribution of each criterion x. W_x is the weighting of each criterion (obtained in step 3). $\sum W_n$ is the sum of all the weights; and S_x is the score of each criterion (obtained in step 4).

6. Sensitivity analysis. To verify the degree of stability of the results obtained in the application of the practical case, a univariate deterministic sensitivity analysis was carried out, using 4 different scenarios (A, B, C and D) to analyse how the results would vary with changes in some of the parameters, *ceteris paribus* the others. The changes incorporated implied a better/worse efficacy and a lower/higher price of the drugs compared with the base scenario (Table 10.3, table 10.4 and table 10.5). For example, scenario A assumes that the drug has a better efficacy than in the base case (in the case of COPD, a FEV of 0.15 L instead of 0.13 L; in the case of MPS I an increase in FVC of 8.4% instead of 7%; and in the case of MPS VII an increase in survival of 7 years instead of 5 years).

TABLE 10.3. EVIDENCE AND ASSUMPTIONS FOR THE APPRAISAL OF THE FICTITIOUS DRUG IN SEVERE COPD

| | | SEVERE COPD | | | | |
|---------------------------------------|---|-------------|--|--|---|---|
| | | Criterion | Current situation | Fictional base case | Sensitivity analysis 1 | Sensitivity analysis 2 |
| Need for intervention | Severity disease | | MR: 121 x 1,000 High risk of exacerbation | Same premises | <i>Ceteris paribus</i> | <i>Ceteris paribus</i> |
| | Size of affected population | | Very severe: 38 x 10,000 inhabitants Severe: 132 x 10,000 inhabitants | | | |
| | Unmet needs | | Several Tx with room for improvement | | | |
| Comparative outcomes of intervention | Comparative Effectiveness/Efficacy | | FEV: 0.5 - 0.10 L Reduction of exacerbations: ≥20% | FEV: 0.13 L Reduction of exacerbations: +5% |  A FEV: 0.15 L |  B FEV: 0.11 L |
| | Comparative Safety/Tolerability | | Cardiovascular AEs | Similar | | |
| | Comparative patient-perceived health / PRO | | HRQoL: Improvement in the SGRQ total score: ≥4 points from baseline | Greater comfort | | |
| Type of benefit of intervention | Type of preventive benefit | | No | Same premises | | |
| | Type of therapeutic benefit | | It does not cure it. Reduction of symptoms and risk of future events | | | |
| Economic consequences of intervention | Comparative cost of intervention | | LAMA+LABA+ICS: €852/year | €1,050/year (+€200) |  C €680/year |  D €1,200/year |
| | Comparative other medical costs | | < €5,500 | Savings of €150 | | |
| | Comparative non-medical costs | | €300 labour productivity losses | Similar or slightly lower | | |
| Knowledge about intervention | Quality of the evidence | | Numerous CTs with good quality evidence | Same premises | | |
| | Expert consensus/clinical practice guidelines | | GesEPOC in Spain; GOLD 2017 | | | |

FEV: Forced Expiratory Volume. AEs: Adverse Effects. MR: Mortality Rate. PROs: Patient-Reported Outcomes. Tx: Treatments. HRQoL: Health-Related Quality of Life. SGRQ: Saint George's Respiratory Questionnaire. CHAQ: Child Health Assessment Questionnaire. CTs: Clinical Trials.

Source: author's preparation based on the available evidence and performed cases.

TABLE 10.4. EVIDENCE AND ASSUMPTIONS FOR THE APPRAISAL OF THE FICTITIOUS DRUG IN MPS-I (SCHEIE SYNDROME)

| MPS-I | | | | | |
|---------------------------------------|---|---|------------------------|------------------------|------------------------|
| | Criterion | Current situation | Fictional base case | Sensitivity analysis 1 | Sensitivity analysis 2 |
| Needs for intervention | Severity disease | They can live to adulthood | Same premises | Ceteris paribus | Ceteris paribus |
| | Size of affected population | 0.2 x 100,000 inhabitants (93 cases) | | | |
| | Unmet needs | Substitute enzyme therapy | | | |
| Comparative outcomes of intervention | Comparative Effectiveness/Efficacy | FVC: +5.6% 6MWT: +38.1% | FVC: +7% 6MWT: +45% | | |
| | Comparative Safety/Tolerability | Severe AEs in patients with respiratory problems | Similar | | |
| | Comparative patient-perceived health/PRO | CHAQ: +57% Pain: passes from 0.93 to 0.56 | 20% improvement | | |
| Type of benefit intervention | Type of preventive benefit | No | Same premises | | |
| | Type of therapeutic benefit | It does not cure it. SET stops the progression of the disease | | | |
| Economic consequences of intervention | Comparative cost of intervention | TSE (Laronidasa): €78,000/year | €100,000/year | | |
| | Comparative other medical costs | N/A | Potential savings | | |
| | Comparative non-medical costs | DNHC: €44,000 - €55,000; IC: €8,500 | Similar | | |
| Knowledge about intervention | Quality of the evidence | 2 CTs (n=45) | Same premises | | |
| | Expert consensus/clinical practice guidelines | Management Guidelines for MPS | | | |

FVC: Forced Vital Capacity. 6MWT: 6-Minutes Walk Test. PROs: Patient-Reported Outcomes. SET: Substitute Enzyme Therapy. CHAQ: Child Health Assessment Questionnaire. DNHC: Direct Non-Healthcare Costs. IC: Indirect Costs. CTs: Clinical Trials. N/A: Not Available.

Source: author's preparation based on the available evidence and performed cases.

- 7. Re-test.** In order to evaluate the degree of consistency and replicability of the analysis, the experts again weighted and scored the criteria for the evaluated interventions, both in their base cases and in their respective sensitivity analyses. This re-evaluation was done online, three weeks after the face-to-face meeting. The degree of agreement or consistency between the responses made on the two occasions was evaluated through the intraclass correlation index (ICC) resulting from a model of mixed effects (Thokala et al., 2016).
- 8. Presentation of results.** WEBER prepared the draft of the results report, which was submitted for the Committee's comments, for final validation. In the discussion, the

conceptual, methodological and other considerations which were collected during all phases of the case study are included.

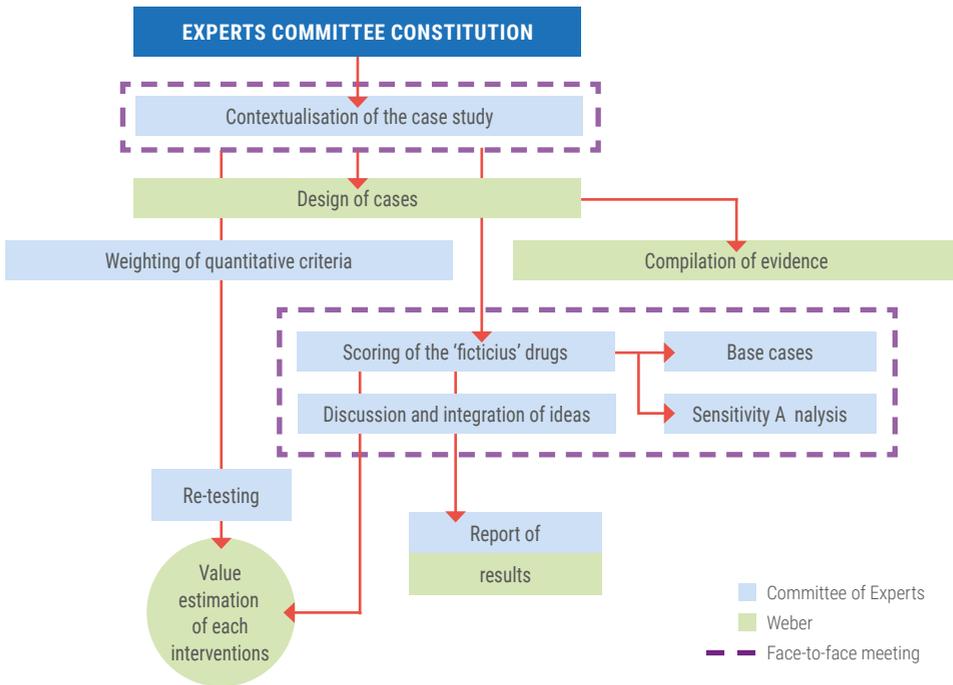
TABLE 10.5. EVIDENCE AND ASSUMPTIONS FOR THE APPRAISAL OF THE FICTITIOUS DRUG IN MPS-VII (SLY SYNDROME)

| | | MPS-VII | | | | |
|---------------------------------------|---|-----------|---|--------------------------------|--|--|
| | | Criterion | Current situation | Fictional base case | Sensitivity analysis 1 | Sensitivity analysis 2 |
| Needs for intervention | Severity disease | | Survival for a few months or until adolescence | Same premises | <i>Ceteris paribus</i> | <i>Ceteris paribus</i> |
| | Size of affected population | | <0.1 x 100,000 inhabitants (100 cases in the world) | | | |
| | Unmet needs | | None (bone marrow transplant) | | | |
| Comparative outcomes of intervention | Comparative effectiveness /efficacy | | -- | Survival: +5 years FVC: +2% |  +7 years |  +3 years |
| | Comparative safety/ tolerability | | -- | Moderate and severe AEs | | |
| | Comparative patient-perceived health/ PRO | | -- | 20% improvement | | |
| Type of benefit intervention | Type of preventive benefit | | No | Same premise | | |
| | Type of therapeutic benefit | | -- | Improvement | | |
| Economic consequences of intervention | Comparative cost of intervention | | -- | €350,000/year |  |  |
| | Comparative other medical costs | | Bone marrow transplant: €60,000 | Similar | | |
| | Comparative non-medical costs | | DNHC: €44,000 - €55,000; IC: €8,500 | Similar | | |
| Knowledge about intervention | Quality of the evidence | | There is no evidence | Same premises | | |
| | Expert consensus / clinical practice guidelines | | Management Guidelines for MPS | | | |

FVC: Forced Vital Capacity. PROs: Patient-Reported Outcomes. AEs: Adverse Effects. DNHC: Direct Non-Healthcare Costs. IC: Indirect Costs. CTs: Clinical Trials. N/A: Not Available.

Source: author's preparation based on the available evidence and performed cases.

By way of summary, Figure 10.1 reflects the different steps taken to perform the case study, specifying who took them and what tasks were carried out in person during the meeting (of 4 hours) that took place in November 2017.

FIGURE 10.1. STEPS CARRIED OUT FOR THE MCDA, THE CASE STUDY

Source: author's preparation

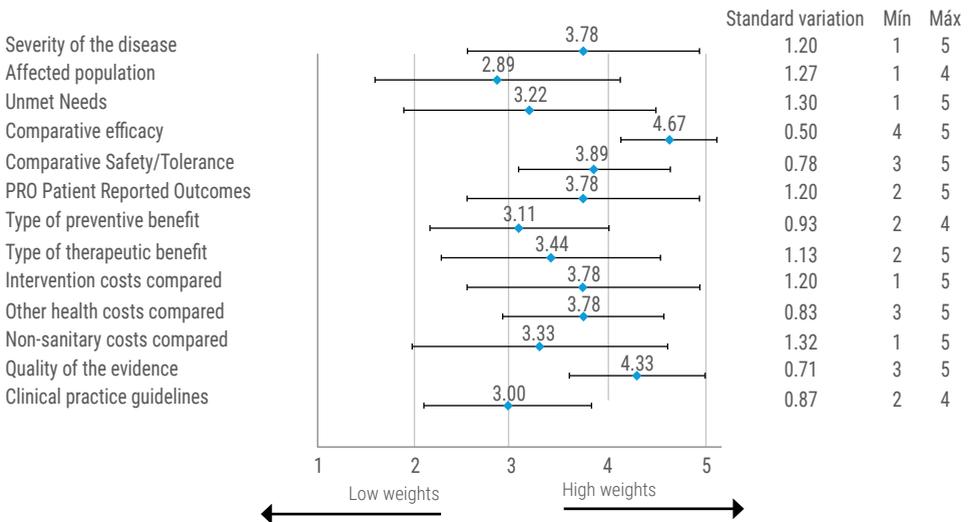
10.3. RESULTS

10.3.1. Weighting of the criteria

The nine members of the Committee weighted the EVIDEM quantitative criteria using two methods: the 5-point direct weight elicitation technique (scale 1-5) and the 100 weighting points allocation method (distribution of 100 points). With the method of scale 1-5, the experts considered that all the criteria had an importance of above 3 (out of a maximum of 5) (Figure 10.2), except for the criterion of the size of the population to which the intervention is directed, which obtained an average weighting of 2.89 ± 1.3 . Based on the results obtained, the most relevant criteria to be taken into account in the evaluation of a healthcare intervention should be the comparative efficacy of the intervention (4.67 ± 0.50) and the quality of the evidence (4.33 ± 0.71), followed by comparative safety/tolerability (3.89 ± 0.78) and, with 3.78 points, the severity of the disease (SD: 1.20), the results reported by the patients (SD: 1.20), the intervention costs (SD: 1.20) and other healthcare costs (SD: 0.83). By contrast, the least relevant criteria would be the size of the affected population (2.89 ± 1.27), the alignment with clinical practice guidelines (3.00 ± 0.87) and the type of preventive benefit (3.11 ± 0.93). The greatest variability in the range of responses occurred in the comparative non-healthcare costs (SD: 1.32), followed by the unmet needs (SD: 1.30) and the size of the affected population (SD: 1.27).

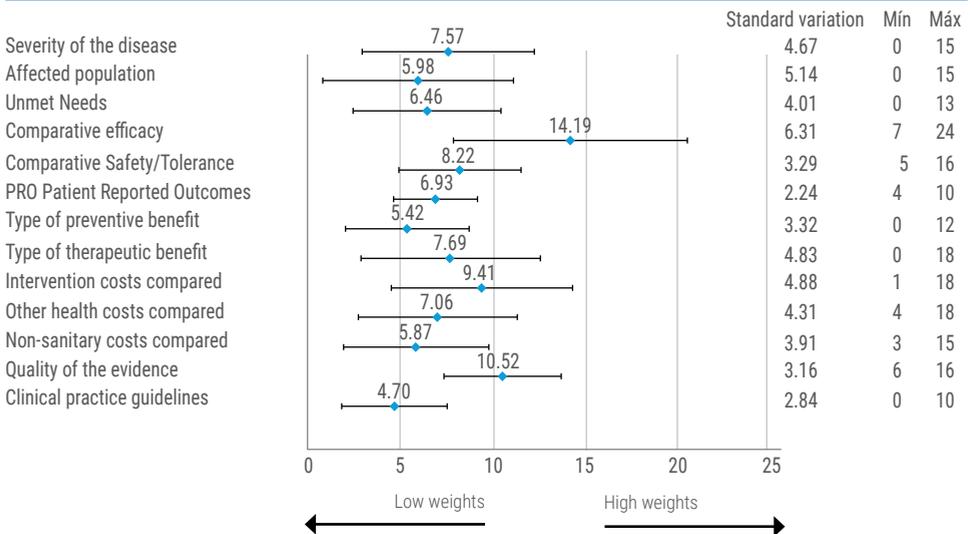
Regarding the weights obtained through the 100-point distribution method, it is worth remembering that the weighting is done in two phases: in the first, the relevance of the 5 domains or groups of criteria is considered, distributing the 100 points among them; in a second phase, the distribution of points is done by individual criterion, within each of the domains.

FIGURE 10.2. AVERAGE WEIGHTS USING THE SCALE 1-5 METHOD



Source: author's preparation

FIGURE 10.3. AVERAGE WEIGHTS USING THE 100-POINTS DISTRIBUTION METHOD (NORMALISED X 100)



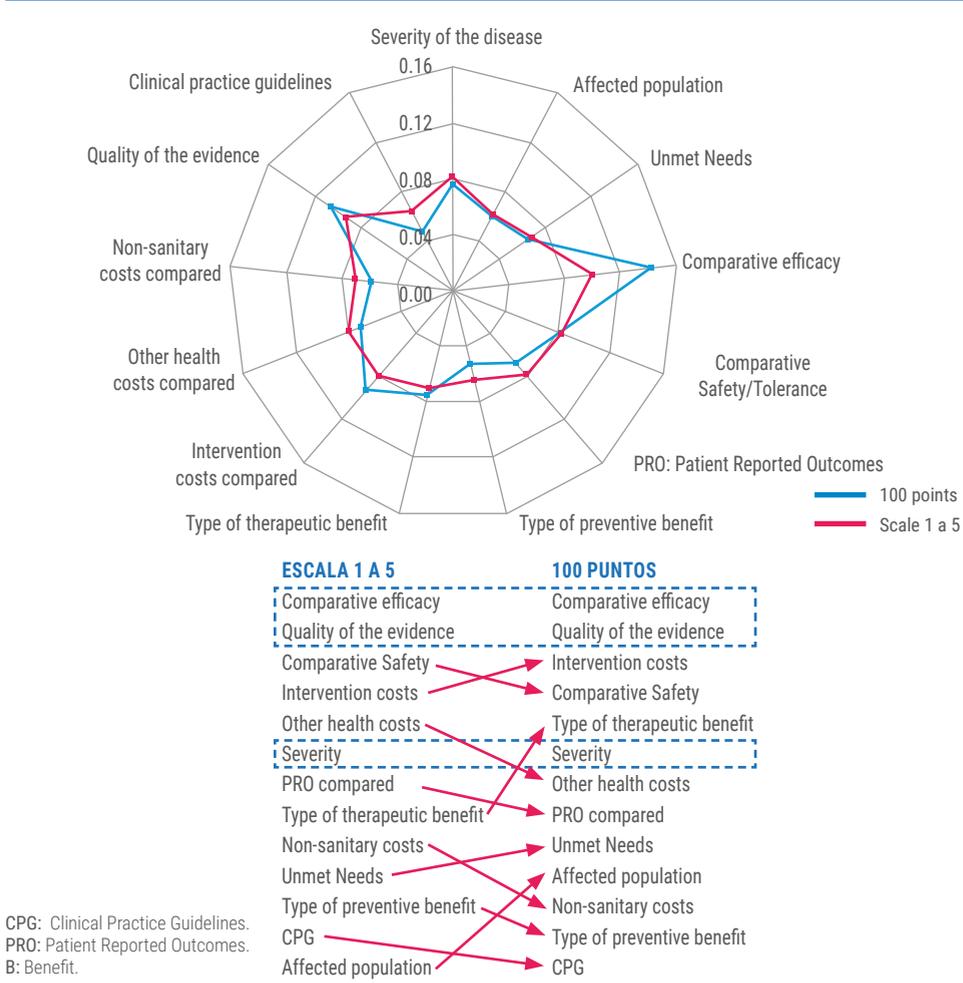
Source: author's preparation

Under the 100-points method, the domain of greatest relative importance for the Committee was the domain that groups the criteria of comparative results of the intervention (29.3 points), followed by the domain of economic consequences (22.3 points), while that of lesser relative importance was that of the type of benefit of the intervention (13.1 points). Regarding the weighting of the criteria that make up each domain, under this method, the Committee considered that the three most relevant criteria should be the comparative efficacy of the intervention (14.19±6.3), the quality of the evidence (10.52±3.2) and

the comparative costs of the intervention (9.41 ± 4.88) (Figure 10.3), coinciding in order of importance with the scale 1-5 method.

In Figure 10.4 we compare the weights obtained through both methods, considering separately the order and the magnitudes reached. Regarding the order or hierarchy, the comparative efficacy and the quality of the evidence were in both cases the criteria with greater relative weight. The rest of the attributes did not keep the same order, although they remained at similar levels, with the exception of the size of the affected population and the type of therapeutic benefit (better positioned in the 100-point method). Regarding the magnitude of the weights (normalised to 1), the greatest discrepancy is observed in the comparative effectiveness criterion (0.10 on scale 1-5 vs. 0.14 in 100 points distribution), followed by alignment with the clinical practice guidelines and the comparative costs of the intervention. These differences can help to better understand the discrepancies obtained in the estimated total value.

FIGURE 10.4. COMPARISON OF RESULTS BETWEEN THE TWO WEIGHTING METHODS (NORMALISED AVERAGE WEIGHTS 0-1)



Source: author's preparation.

10.3.2 Scoring of the interventions

During the face-to-face meeting, the members of the Committee scored, individually, the three fictitious drugs proposed for the three pathologies. The scores were based on the scientific evidence provided about the current situation of the diseases (Table 10.2), the assumptions considered for each fictitious drug in the base case (Table 10.6) and the discussion that took place during the meeting.

TABLE 10.6. BASE CASES CONSIDERED FOR THE FICTITIOUS DRUGS

| CRITERION | BASE CASE OF COPD | BASE CASE OF MPS-I | BASE CASE OF MPS-VII |
|--|--|---|---|
| Severity disease | MR: 121 X 1,000 High risk of exacerbation | They can live to adulthood | Survival for a few months or until adolescence |
| Size of affected population | Very severe: 38 x 10,000 inhabitants Severe: 132 x 10,000 inhabitants | 0.2 x 100,000 inhabitants (93 cases) | <0.1 x 100,000 inhabitants (100 cases in the world) |
| Unmet needs | Several Tx with room for improvement | Substitute enzyme therapy | Nothing (bone marrow transplant) |
| Comparative Effectiveness/Efficacy | FEV: 0.13 L Reduction of exacerbations: +5% | FVC: +7% 6MWT: +45% | Survival: +5 years FVC: +2% |
| Comparative Safety/Tolerability | Similar | Similar | Moderate and severe AEs |
| Comparative patient-perceived health/PRO | Greater comfort | 20 % Improvement | 20 % Improvement |
| Type of preventive benefit | No | No | No |
| Type of therapeutic benefit | It does not cure it. Reduction of symptoms and risk of future events | It does not cure it. SET stops the progression of the disease | Improvement |
| Comparative cost consequences cost of intervention | €1,050/year (+€200) | €100,000/year | €350,000/year |
| Comparative cost consequences other medical costs | Savings of €150 | Potential savings | Similar |
| Comparative cost consequences non-medical costs | Similar or slightly lower | Similar | Similar |
| Quality of the evidence | Numerous CTs with good quality evidence | 2 CTs(n=45) | There is no evidence |
| Expert consensus/clinical practice guidelines | GesEPOC in Spain; GOLD 2017 | Management Guidelines for MPS | Management Guidelines for MPS |

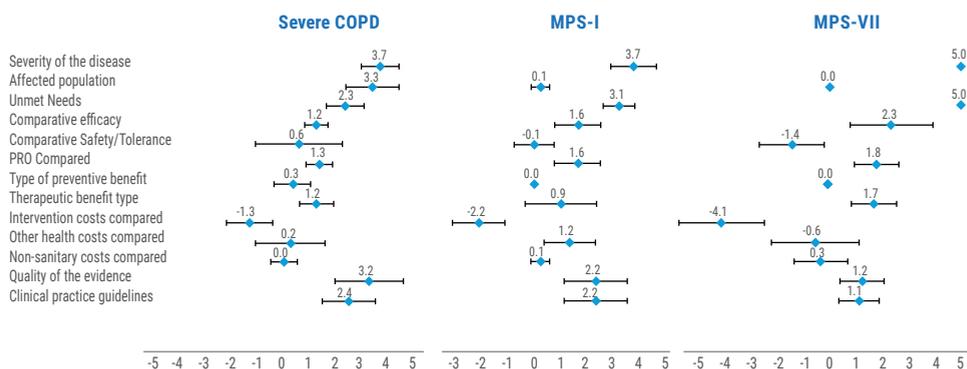
FEV: Forced Expiratory Volume. FVC: Forced Vital Capacity. 6MWT: 6-Minutes Walk Test. MR: Mortality Rate. PROs: Patient-Reported Outcomes. Tx: Treatments. SET: Substitute Enzyme Therapy. CTs: Clinical Trials. AEs: Adverse Effects. GOLD: global initiative for chronic obstructive lung disease. MPS: Mucopolysaccharidosis.

Source: author's preparation based on the available evidence and performed cases.

It should be noted that, even though the scores for each drug were performed independently, an indirect comparison between them was inevitably made, as shown by the results obtained. The average scores awarded by Committee members to the three interventions evaluated are presented in Figure 10.5.

In the case of MPS VII, the Committee gave the highest score (5 out of 5) and the one with the highest consensus (SD: 0.00) to the severity of the disease and to the unmet needs of this rare disease that lacks therapeutic alternatives. The severity of the disease also proved to be the most relevant criterion for the Committee, both for severe COPD and for MPS I, with very similar scores (3.67 ± 0.71 and 3.67 ± 0.87 respectively), lower than for MPS VII. The scores relating to unmet needs reveal that, for the experts, severe COPD (2.3) has more satisfactory treatment options than MPS I (3.1). Consistent with the prevalence of the evaluated pathology, the population size obtained an average score close to zero in MPS I and MPS VII, while for COPD it was 3.3.

FIGURE 10.5. AVERAGE SCORES OF THE EVALUATED INTERVENTIONS (BASE CASES)



Source: author's preparation.

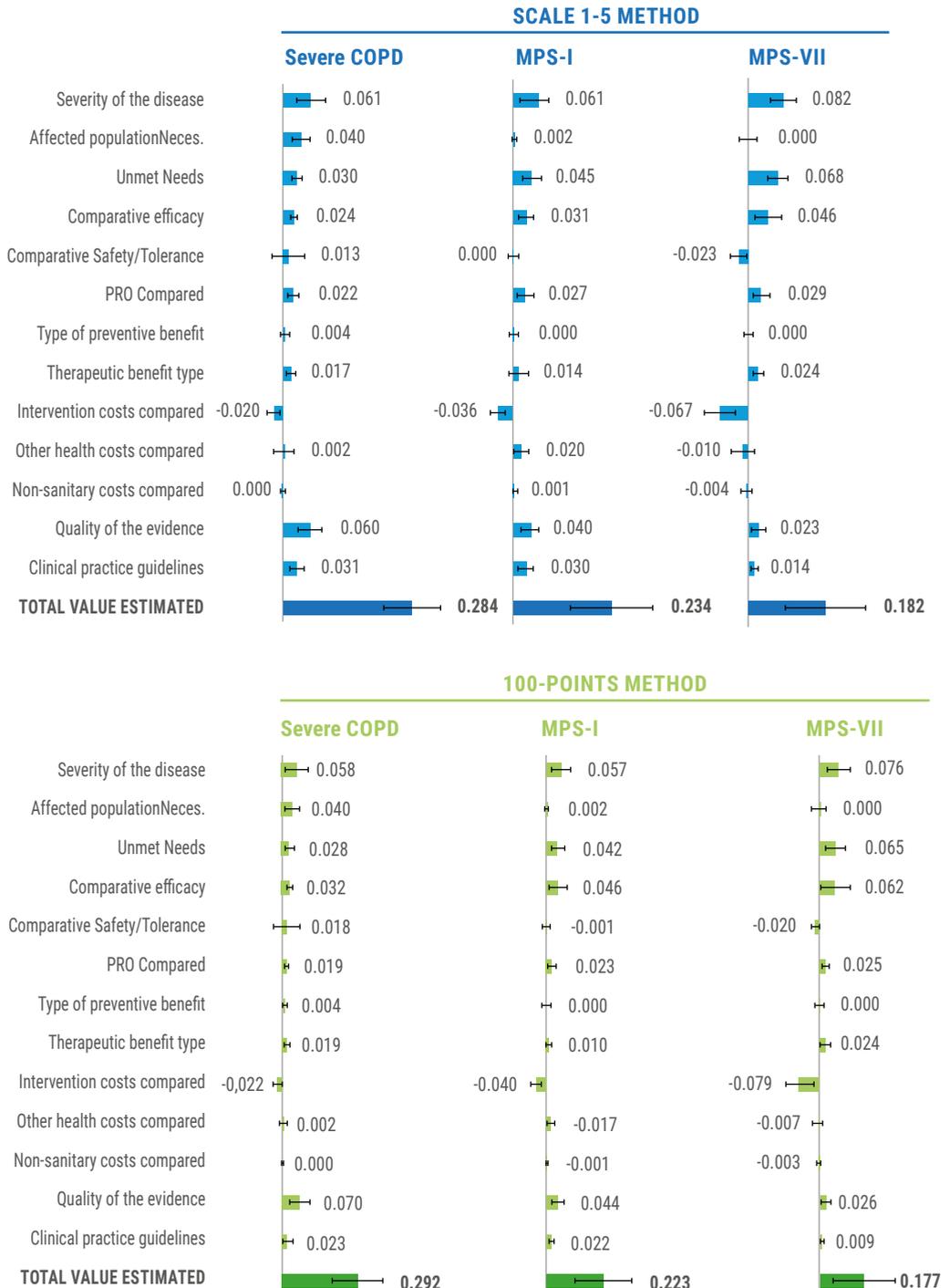
Regarding the clinical efficacy of the drugs considered, the experts scored the fictitious drugs at 1.22 (SD: 0.44) for severe COPD, 1.56 (SD: 0.88) for MPS I, and 2.33 (SD: 1.58) for MPS VII. The score awarded to the criterion patient-reported outcomes (PROs) followed the same gradient. Nevertheless, the Committee scored the tolerability of the drug better for COPD than for the two rare diseases (0.6 versus -0.1 and -1.4). The type of preventive benefit of the two drugs aimed at mucopolysaccharidosis reflects identical results (0.00), while in the case of severe COPD, this criterion obtained a slightly higher score (0.33). The type of therapeutic benefit for the drug for MPS VII was scored at 1.7 versus the 1.2 received by the treatment for severe COPD and 0.9 for the drug for MPS I. And the experts scored the quality of the evidence in COPD (3.2) better than in the MPSs, giving the idea that it is an area with more clinical trials, which provide more relevant and better-reported evidence.

Finally, in the three interventions evaluated, the scores for the compared cost of the intervention were negative, showing that they would represent an increase in costs compared with the current situation. The increase in costs would be greater in the case of MPS VII (score of -4.1), followed by MPS I (-2.2) and COPD (-1.3).

10.3.3 Estimated total value of the interventions

The estimated total value incorporates the weights and scores of each Committee member for each drug evaluated on a scale between 0 and 1. The three drugs evaluated in this

FIGURE 10.6. ESTIMATED TOTAL VALUE USING THE TWO WEIGHTING METHODS



Source: author's preparation.

MCDA were considered value-added drugs, since they obtained positive values compared with their comparators (Figure 10.6).

The total value estimated by the whole Committee for the fictitious drug for severe COPD, compared with the current alternative, was 0.2836 [95%CI: 0.235-0.333] with the scale method between 1 and 5, and slightly higher with the 100-point method, with an estimated value of 0.2916 [95%CI: 0.215-0.368]. This value is higher than that obtained for mucopolysaccharidoses, both in its most attenuated form and in its strongest forms: for MPS I, a value of 0.2341 [95%CI: 0.158-0.310] was obtained with the scale method 1-5 and 0.2231 [95%CI: 0.137-0.309] with the 100-point method while for MPS VII a value of 0.1817 [IC95%: 0,108-0,255] was obtained with the scale 1-5 method, and 0,1773 [IC95%: 0,083-0,271] with the 100-point method (Figure 10.6). This would imply that, using this methodology and given the assumptions made for fictitious drugs, the Committee evaluated the drug more positively for treating COPD than the drugs used to treat mucopolysaccharidoses.

The main difference in total value between the two methods of weighting is given, in the case of COPD, by the contribution of size of the population and quality of the evidence; in the case of MPS I, by the contribution of quality of evidence and costs of the intervention; and in the case of MPS VII, by the contribution of costs of the intervention and unmet needs.

In the case of the drug for COPD, the main elements of value were the severity of the disease, the quality of the available scientific evidence and the size of the affected population, while in MPS, the criteria that contributed the most value were the severity of the disease, the unmet needs and the quality of the evidence. The additional cost that the acquisition would represent was the only element that detracted value from the final estimates, being of greater absolute magnitude for MPS VII, followed by MPS I and COPD.

When comparing the drugs with each other, it is observed that a differential element in the case of severe COPD is its high prevalence (the drug gains 4 hundredths against the MPSs) and a higher quality of the evidence (it gains between 2 and 4 hundredths). For MPS VII, the severity of the pathology adds 2 additional hundredths and the unmet needs up to 4 hundredths. Nevertheless, the value provided by the efficacy of the fictitious drug (between 2 and 3 hundredths more) is more than offset by the value that detracts its lower tolerability (between 2 and 4 hundredths) and higher cost (up to 5 hundredths).

The orphan drug for the most severe disease without therapeutic alternatives, associated with a higher price, obtained a lower final value than the other orphan drug considered. This was because the relative gains of MPS VII versus MPS I in terms of severity, unmet needs, efficacy and type of therapeutic benefit were more than offset by the relative reductions of value resulting from the criteria of safety, costs, quality of the evidence and alignment with the CPGs.

10.3.4. Managing uncertainty

In this section we try to quantify the impact of uncertainty (structural, stochastic, about parameters, etc.) on the results, in order to have a measure of the validity of the results. Thus, on the one hand, we performed several deterministic sensitivity analyses on some of the most relevant parameters, to observe the influence of these changes on the results. On the other hand, to try to manage the uncertainty in the weighting and scoring, we performed a re-evaluation (or re-test) at another time.

10.3.4.1. Univariate deterministic sensitivity analysis

A deterministic sensitivity analysis consists of varying only one parameter, leaving the rest unaltered, and observing the influence of this change on the result and on the conclusions reached. Although it does not allow one to observe the interaction between several parameters at the same time, it does allow one to check the sensitivity of the model for parameters considered to be crucial, such as the efficacy and cost of the treatments evaluated. Specifically, changes in the variables were made, for the three pathologies, in the following four scenarios with respect to the base case, *ceteris paribus* the other criteria (Table 10.7):

- **Scenario A:** Increase in the effectiveness of the evaluated intervention
- **Scenario B:** Decrease in the effectiveness of the evaluated intervention
- **Scenario C:** Increase in the intervention costs
- **Scenario D:** Decrease in the intervention costs

As expected, the results obtained in the sensitivity analyses show that, with an improvement in efficiency or a reduction in the relative price, the estimated total value of the drugs would increase, while it would decrease compared with the base case in the event of a worse efficacy or a higher cost of the drug evaluatedⁱⁱ.

TABLE 10.7. ESTIMATED TOTAL VALUE OF THE INTERVENTIONS IN THE FOUR SCENARIOS OF THE SENSITIVITY ANALYSIS, USING THE TWO WEIGHTING METHODS

| | SEVERE COPD | | | MPS-I | | | MPS-VII | | |
|--|---------------------------|-------------|------------|-----------------------------|-------------|------------|-----------------------------|-------------|------------|
| BASE CASE | BASE CASE | 0.2836 | 0.2916 | BASE CASE | 0.2341 | 0.2231 | BASE CASE | 0.1817 | 0.1773 |
| VARIABLE CHANGE | VARIABLE CHANGE | Scale 1 - 5 | 100 points | VARIABLE CHANGE | Scale 1 - 5 | 100 points | VARIABLE CHANGE | Scale 1 - 5 | 100 points |
| SENSITIVITY ANALYSIS 1 COMPARATIVE EFFICIENCY | Scenario A FEV: 0.15 | 0.3038 | 0.3209 | Scenario A FVC: +8.4% | 0.2409 | 0.2344 | Scenario A +7 years | 0.2003 | 0.2057 |
| | Scenario B FEV: 0.11 | 0.2506 | 0.2455 | Scenario B FVC: +4.5% | 0.1992 | 0.1829 | Scenario B +3 years | 0.1552 | 0.1435 |
| SENSITIVITY ANALYSIS 2 INTERVENTION COST COMPARED | Scenario C €1,200 year | 0.2764 | 0.2851 | Scenario C €120,000 year | 0.2264 | 0.2164 | Scenario C €420,000 year | 0.1858 | 0.1862 |
| | Scenario D €680 year | 0.3162 | 0.3290 | Scenario D €62,000 year | 0.2752 | 0.2693 | Scenario D €280,000 year | 0.2074 | 0.2128 |

FEV: Forced Expiratory Volume. FVC: Forced Vital Capacity.

Source: author's preparation.

ⁱⁱ With the exception of scenario C (increase in the cost of intervention) of MPS VII, in which the total value increases compared with the base case. The result is due to the inconsistent scoring of one of the experts, which was corrected in the re-test.

Table 10.8 shows the results of the sensitivity analyses in percentage terms compared with the base case. It should be noted that sensitivity over total value is greater with a worsening of the effectiveness than in the case of an improvement. Similarly, the results vary more with a rise in relative cost than with a decrease (compared with the base case). Additionally, the model is generally more sensitive to the changes proposed when using the weighting method of 100 points than when using the scale 1 to 5.

Moreover, in the case of MPS VII, there is evidence of greater sensitivity of the results to changes in the parameters than in the other two pathologies. This may be due to the assumptions made for the parameters, but also to a greater sensitivity of the evaluators when faced with these very serious diseases, and without alternatives. For example, with an improvement in comparative efficacy, the final estimated value in MPS VII would increase between 10.2% and 16% compared with the base case, while these ranges are lower in the cases of severe COPD (7.1% - 10%) and MPS I (2.9% - 5.1%) (Table 10.8).

TABLE 10.8. PERCENTAGE VARIATIONS IN THE TOTAL ESTIMATED VALUE COMPARED WITH THE BASE CASE, FOR THE FOUR SCENARIOS OF THE SENSITIVITY ANALYSIS, UNDER THE TWO WEIGHTING METHODS

| VARIATIONS % vs. BASE CASE | SEVERE COPD | | MPS-I | | MPS-VII | |
|-------------------------------|-------------|------------|-------------|------------|-------------|------------|
| | SCALE 1 - 5 | 100 POINTS | SCALE 1 - 5 | 100 POINTS | SCALE 1 - 5 | 100 POINTS |
| SCENARIO A ▲ Efficacy | 7.1% | 10.0% | 2.9% | 5.1% | 10.2% | 16.0% |
| SCENARIO B ▼ Efficacy | -11.7% | -15.8% | -14.9% | -18.0% | -14.5% | -19.0% |
| SCENARIO C ▲ Cost | -2.6% | -2.2% | -3.3% | -3.0% | 2.2% | 5.0% |
| SCENARIO D ▼ Cost | 11.5% | 12.8% | 17.6% | 20.7% | 14.2% | 20.0% |

Source: author's preparation.

10.3.4.2. Re-test

In order to verify the replicability of the analysis, three weeks after the face-to-face meeting the Committee of Experts were asked to weight the criteria again (with the two methods) and to score the three fictitious drugs, in their base cases and in the four scenarios of the sensitivity analysis.

Between the test and the re-test, 63% of the weights were identical using the scale 1-5 method, 35% differed by 1 point and 2% differed by 2 points. When using the 100-point distribution method, 38% of the weights differed by 1 point or less (out of 100), 41% differed by between 1 and 3 points, 14% by between 3-5 points and 7% by more than 5 points. The individual correlation between the weights reported by the experts on the two occasions was 0.824 using the 'Scale 1 to 5' method and 0.786 using the '100 points' method.

Regarding the scores, the highest individual correlation between the two occasions was for MPS VII (0.914), followed by COPD (0.808) and MPS I (0.667). In MPS VII, 68% of the scores were identical between the test and the re-test, 23% varied by 1 point, 5% by 2 points and 5% by 3 or 4 points. In COPD, 56% of the scores were identical, 30% varied by 1 point, 9% by 2 points and 4% by 3 or 4 points. And in MPS I, 54% of the scores were identical, 27% varied by 1 point, 10% by 2 points and 9% by 3 points or more (both for and against).

In general, the total value obtained in the re-test for the three drugs was lower than that obtained in the base evaluation (test), with the exception of the exercise performed in COPD with the scale 1-5, in which there was an increase (Table 10.9). The reduction occurring in the re-test was greater with the 100-point method than with the method of scale 1-5. Nevertheless, the value of the drug for COPD is still higher than that of MPS I, which in turn remains slightly higher than that of MPS VII.

The correlation analysis indicates a fair to good reproducibility of the exercise. According to the individual intraclass correlations of the total estimated values, the degree of reliability and consistency of the results was higher for MPS than for COPD (individual ICC of 0.607-0.679 for MPS I; 0.696-0.862 for MPS VII; 0.527-0.790 for COPD).

TABLE 10.9. COMPARISON OF THE TOTAL ESTIMATED VALUE IN THE TEST AND RE-TEST, AND INTRACLAS CORRELATIONS

| | WEIGHTING METHOD | FINAL VALUE | | | INTRACLAS CORRELATIONS | |
|-------------|------------------|-------------|---------------|-----------|------------------------|-------------|
| | | Test (n=9) | Re-test (n=9) | Variation | Individual ICC | Average ICC |
| SEVERE COPD | SCALE 1 - 5 | 0.2836 | 0.3105 | 9.5% | 0.527 | 0.690 |
| | 100 POINTS | 0.2916 | 0.2819 | -3.3% | 0.790 | 0.883 |
| MPS-I | SCALE 1 - 5 | 0.2341 | 0.1980 | -15.4% | 0.679 | 0.809 |
| | 100 POINTS | 0.2231 | 0.1717 | -23.0% | 0.607 | 0.755 |
| MPS-VII | SCALE 1 - 5 | 0.1817 | 0.1669 | -8.1% | 0.696 | 0.821 |
| | 100 POINTS | 0.1773 | 0.1414 | -20.2% | 0.862 | 0.926 |

Source: author's preparation.

Finally, the re-test performed on the univariate sensitivity analyses reveals the same tendency in the results as in the test: in all cases, a higher value is obtained when the efficiency increases or the price decreases and a lower value when the opposite occurs. The model is still more sensitive with the 100-points distribution method. Likewise, it is more sensitive to a worsening of the effectiveness with respect to the base case than in the case of an improvement, and also in the face of a price reduction rather than an increase. There is no clear pattern between the test and the re-test, the sensitivity being in some cases greater and in other cases lower (Figure 10.10).

TABLE 10.10. TOTAL ESTIMATED VALUE OBTAINED IN THE RE-TEST FOR SENSITIVITY ANALYSES

| | SEVERE COPD | | | MPS-I | | | MPS-VII | | |
|--|---------------------------|-------------|------------|-----------------------------|-------------|------------|-----------------------------|-------------|------------|
| BASE CASE | BASE CASE | 0.3105 | 0.2819 | BASE CASE | 0.1980 | 0.1717 | BASE CASE | 0.1669 | 0.1414 |
| | VARIABLE CHANGE | Scale 1 - 5 | 100 points | VARIABLE CHANGE | Scale 1 - 5 | 100 points | VARIABLE CHANGE | Scale 1 - 5 | 100 points |
| SENSITIVITY ANALYSIS 1 COMPARATIVE EFFICIENCY | Scenario A FEV: 0.15 | 0.3200 | 0.2940 | Scenario A FVC: +8.4% | 0.2117 | 0.1881 | Scenario A +7 years | 0.1826 | 0.1600 |
| | Scenario B FEV: 0.11 | 0.2922 | 0.2580 | Scenario B FVC: +4.5% | 0.1729 | 0.1350 | Scenario B +3 years | 0.1367 | 0.0963 |
| SENSITIVITY ANALYSIS 2 INTERVENTION COST COMPARED | Scenario C €1,200 year | 0.2973 | 0.2637 | Scenario C €120,000 year | 0.1803 | 0.1480 | Scenario C €420,000 year | 0.1574 | 0.1303 |
| | Scenario D €680 year | 0.3423 | 0.3256 | Scenario D €62,000 year | 0.2240 | 0.2077 | Scenario D €280,000 year | 0.1810 | 0.1627 |

FEV: Forced Expiratory Volume. FVC: Forced Vital Capacity.

Source: author's preparation.

10.4. DISCUSSION

This case study has allowed us to approach a numerical value for three interventions aimed at a prevalent disease such as severe COPD and at two subtypes of a rare disease such as mucopolysaccharidosis. Although these are different pathologies, in a way the study allows us to compare the results obtained for the three drugs, considering the order, rather than the magnitude, of the results and analysing their sensitivity to certain changes in the parameters. The results indicate that, given the assumptions made, the hypothetical drug for the prevalent disease has a higher final value than the drugs for the rare disease. We must not forget that the drugs evaluated are fictitious, with fictitious characteristics considered ad hoc for the exercise. From the exercise it is possible to derive the following reflections:

- This type of analysis allows us to consider the appraisal of interventions from a broader perspective than usual, taking into account additional elements that may be relevant for agents, such as the severity of the disease, the existence of therapeutic alternatives and the quality of clinical evidence.
- The drug for the prevalent disease obtained a higher final value than the interventions for rare diseases. In turn, the orphan drug for the most serious disease without therapeutic alternatives, associated with a higher price, obtained a lower final value than the other orphan considered, since the disadvantages offered in terms of cost, quality of evidence and safety weighted more than the value added by the severity of the disease, the absence of alternatives and the greater comparative efficacy.

- The weighting method used affects the final results. The way in which experts rank the importance of the criteria has some effect, although not substantial, on the final result. The different ways of weighting implicitly entail a different prioritisation of the criteria, which in turn can lead to results of different magnitude. A common pattern cannot be identified: the results with the 1-5 scale were higher for the rare diseases, but not for the prevalent disease.
- The analysis involve some cognitive complexity, especially when several drugs are evaluated simultaneously, in different scenarios. The model requires concentration by the evaluator and a very clear understanding of how the tool works. Errors may occur in scoring that give rise to inconsistent responses. So it is important that the scoring should be done during the face-to-face meeting, preceded by a thorough explanation of the MCDA methodology and the assumptions made, and making sure that the evaluators have enough time to perform the assessments.
- The consistency of the analysis must be explored through a re-test of the weights and scores. In line with other studies, the reproducibility of scores was higher than that of the weights (Wagner et al., 2017; Goetghebeur et al., 2010). On both occasions the results were consistent, but in some cases they differed by 20%. The greatest differences occurred for the rare diseases, when the 100-points method was used, which in our example gave more importance to the relative value of the efficacy and the comparative costs.
- The various deterministic sensitivity analyses proposed showed that the results are sensitive to changes in the variables of efficacy and cost, especially when the parameters worsen compared with the base case.
- The price is a substantial component of value, although its weight differs in each pathology. For MPS without alternatives, reducing the price by 20% compared with the base case was associated with an increase in the total value of between 14% and 20%, while an increase of 20% reduced the value by less than 5%. In the case of MPS with alternatives, reducing the price by 35% raised the total value by between 18% and 21%, figures that in the prevalent disease were approximately 12%.
- The way in which evidence about drugs and pathologies is presented can affect the final results. In this case, presenting the Committee information about pathologies and drugs together could have influenced the results, since they could indirectly make comparisons between the three cases. On the other hand, in the meeting they commented that providing graphic information (photos) about the effects of pathologies could influence the scores.
- The composition of the evaluating Committee will have an effect on the results. Beyond the effect that the particular view of each member of the Committee may have, the profile of each of them will also affect the final evaluation. In this case, it was a Committee formed by health economists, so the weight given to the economic criteria may be greater than in other committees and in society in general. In fact, the group gave the three cost criteria a greater relative importance than did the multi-

disciplinary committees of other MCDA (Gilabert-Perramon et al., 2017). This fact contributed to the better relative result of severe COPD compared with mucopolysaccharidoses, which are rare diseases associated with high costs.

This study has certain limitations, some specific to the MCDA and others specific to this case. As it was the EVIDEM approach that was adopted a fixed set of criteria is being taken into account, so other relevant criteria could have been excluded. The results of the exercise are subjective, and depend to a large extent on the value judgements, experience and training of the members of the evaluation Committee. In addition, the evaluation obtained is not generalisable to other comparators, nor lasting over time. In this case, the evaluation Committee, composed mainly of health economists, may be collecting only a partial view of the value of the interventions. Moreover, the assumptions raised have not taken into account the psychometric capacities of the scales.

Nevertheless, the exercise has allowed us to complement the theoretical perspective of the book's authors with practical experience, which has allowed a deeper reflection about the scope of MCDA and about its advantages and limitations. Using the EVIDEM methodology (a homogeneous, validated and widely used methodology) allowed us to focus in this case on the practical applicability of the tool in different types of diseases, leaving aside the debate about the attributes to be included. The exchange of opinions between the experts contributes to enrich the analysis and the individual evaluations. The MCDA allows us to analyse the interventions from an expanded perspective, explicitly considering different value attributes, and determining to what extent each of them affects the final value. Thus, it can serve as a complementary tool for economic evaluation in healthcare decision-making.

10.5. CONCLUSIONS

The results obtained in this practical application of an MCDA help us to better understand where the value of healthcare interventions lies for a specific group of agents, in this case, health economists. The exercise has allowed us to explore the additional value of the proposed interventions beyond the economic perspective, proving the importance of other attributes. Based on the assumptions made, the fictitious drug aimed at the prevalent disease has resulted in greater added value than those directed at rare diseases. This result supports the fact that the cost of treatment, the quality of the scientific evidence and the size of the affected population may have sufficient weight to counteract the value contributed by an effective treatment aimed at a serious, less prevalent pathology and with important uncovered needs, but associated with a high cost. Nevertheless, the greatest usefulness of the exercise is probably in helping to reflect on the robustness and sensitivity of the methodology and the process.

REFERENCES

AECOM, y SEPEAP. 2015. «Guía para el manejo de las MPS». Guías de práctica clínica en el SNS. Madrid. Available from: https://www.fesemi.org/sites/default/files/documentos/publicaciones/guia_manejo_mps.pdf.

Barrecheguren M, Monteagudo M, Ferrer J, Borrell E, Llor C, Esquinas C, Miravittles M. Treatment Patterns in COPD Patients Newly Diagnosed in Primary Care. A Population-Based Study. *Respir Med*. 2016; 111: 47-53. <https://doi.org/10.1016/j.rmed.2015.12.004>.

BotPlus. 2017a. «PVL Adurazyme». 2017. <https://botplusweb.portalfarma.com/botplus.aspx>.

BotPlus. 2017b. «PVL Elapraxe». 2017. <https://botplusweb.portalfarma.com/botplus.aspx>.

Cazzola M, Matera MG. Triple Combinations in Chronic Obstructive Pulmonary Disease - Is Three Better than Two? *Expert Opin Pharmacother*. 2014; 15(17): 2475-78. <https://doi.org/10.1517/14656566.2014.972367>.

Clarke LA, Wraith JE, Beck M, Kolodny EH, Pastores GM, Muenzer J, et al. Long-term Efficacy and Safety of Laronidase in the Treatment of Mucopolysaccharidosis I. *Pediatrics*. 2009; 123(1): 229-40. doi.org/10.1542/peds.2007-3847.

EVIDEM Collaboration. EVIDEM Framework 10th edition- Overview. 2017 Available from: <https://www.evidem.org/evidem-framework/>.

Gilbert-Perramon A, Torrent-Farnell J, Catalan A, Prat A, Fontanet M, Puig-Peiró R, Merino-Montero S, Khoury H, Goetghebeur MM, Badia X. Drug evaluation and decision making in Catalonia: development and validation of a methodological framework based on multi-criteria decision analysis (MCDA) for orphan drugs. *Int J Technol Assess Health Care*. 2017: 1–10.

Global Initiative for Chronic Obstructive Lung Disease. Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease 2017. 2017. Available from: <http://goldcopd.org/download/326/>.

Goetghebeur MM, Wagner M, Khoury H, Levitt RJ, Erickson LJ, Rindress D. Evidence and Value: Impact on DEcisionMaking – the EVIDEM framework and potential applications. *BMC Health Serv Res*. 2008; 8(1).

Granda-Orive D, Ignacio J, Solano-Reina S. Mortalidad en la enfermedad pulmonar obstructiva crónica. Año SEPAR EPOC/Tabaco. *Arch Bronconeumol*. 2016; 52(8): 407-8.

Gruffydd-Jones K, Jones MM. NICE guidelines for chronic obstructive pulmonary disease: implications for primary care. *Br J Gen Pract*. 2011; 61(583): 91-2.

Iskrov G, Miteva-Katrandzhieva T, Stefanov R. Multi-Criteria Decision Analysis for Assessment and Appraisal of Orphan Drugs. *Front Public Health*. 2016; 4.

Izquierdo JL. The burden of COPD in Spain: results from the confronting COPD survey. *Respiratory Medicine*. 2003;97:S61-9. [doi.org/10.1016/S0954-6111\(03\)80026-4](https://doi.org/10.1016/S0954-6111(03)80026-4).

Khakban A, Sin DD, FitzGerald JM, McManus BM, Ng R, Hollander Z, et al. The Projected Epidemic of Chronic Obstructive Pulmonary Disease Hospitalizations over the Next 15 Years. A Population-based Perspective. *Am J Respir Crit Care Med*. 2016; 195(3): 287-91.

Khan SA, Peracha H, Ballhausen D, Wiesbauer A, Rohrbach M, Gautschi M, et al. Epidemiology of mucopolysaccharidoses. *Mol Genet Metab*. 2017; 121(3): 227-40.

Lampe C, Jain M, Olaye A, Meesen B, Decker C, Mengel E. Relationship Between Patient-Reported Outcomes and Clinical Outcomes in Patients With Morquio A Syndrome. *J Inborn Errors Metab Screen*. 2015.

Mannino DM, Kiri VA. Changing the burden of COPD mortality. *Int J Chron Obstruct Pulmon Dis*. 2006; 1(3): 219-33.

Martín Escudero JC. Calidad de vida y EPOC. *Rev Clin Esp.* 2009; 209(2): 59-60. doi.org/10.1016/S0014-2565(09)70626-5.

Miravittles M, Soler-Cataluña JJ, Calle M, Molina J, Almagro P, Quintano JA, et al. Guía española de la enfermedad pulmonar obstructiva crónica (GesEPOC) 2017. Tratamiento farmacológico en fase estable. *Arch Bronconeumol.* 2017; 53(6): 324-35.

Molina París J, Palacios Gómez L, Piñera Salmerón P, Pleguezuelos Cobo E, Simonet P, Soriano Ortiz JB, et al. Guía de Práctica Clínica para el Diagnóstico y Tratamiento de Pacientes con Enfermedad Pulmonar Obstructiva Crónica (EPOC) – Guía Española de la EPOC (GesEPOC). Versión 2017. Available from: <https://www.sogapar.info/wp-content/uploads/2016/12/Guia-GESEPOC-2017.pdf>.

MSSSI, BOE-A-2006-18375. Coste medio Trasplante de médula ósea alogénico. 2006. Available from: <https://www.boe.es/buscar/act.php?id=BOE-A-2006-18375>.

Muenzer J. Overview of the mucopolysaccharidoses. *Rheumatology (Oxford).* 2011; 50 (suppl 5): v4-12.

Muenzer J. Early initiation of enzyme replacement therapy for the mucopolysaccharidoses. *Mol Genet Metab.* 2014; 111(2): 63-72.

Müllerova H, Maselli DJ, Locantore N, Vestbo J, Hurst JR, Wedzicha JA, et al. Hospitalized exacerbations of COPD: risk factors and outcomes in the ECLIPSE cohort. *Chest.* 2015; 147(4): 999-1007.

Noh H, Lee JI. Current and potential therapeutic strategies for mucopolysaccharidoses. *J Clin Pharm Ther.* 2014;39(3):215-24.

Péntek M, Gulácsi L, Brodsky V, Baji P, Boncz I, Pogány G, et al. Social/economic costs and health-related quality of life of mucopolysaccharidosis patients and their caregivers in Europe. *Eur J Health Econ.* 2016; 17(1): 89-98. doi.org/10.1007/s10198-016-0787-0.

Pérez-López J, Morales-Conejo M, López-Rodríguez M, Hermida-Ameijeiras Á, Moltó-Abad M. Efficacy of laronidase therapy in patients with mucopolysaccharidosis type I who initiated enzyme replacement therapy in adult age. A systematic review and meta-analysis. *Mol Genet Metab.* 2017; 121(2): 138-49.

Pinto R, Caseiro C, Lemos M, Lopes L, Fontes A, Ribeiro H, et al. Prevalence of lysosomal storage diseases in Portugal. *Eur J Hum Genet.* 2003; 12(2): 87-92.

Santibáñez M, Garrastazu R, Ruiz-Nuñez M, Helguera JM, Arenal S, Bonnardeux C, et al. Predictors of Hospitalized Exacerbations and Mortality in Chronic Obstructive Pulmonary Disease. *PLOS ONE.* 2016; 11(6): e0158727.

Soler-Cataluña JJ, Martínez-García M, Roman S, Salcedo E, Navarro M, Ochoa R. Severe acute exacerbations and mortality in patients with chronic obstructive pulmonary disease. *Thorax.* 2005;60(11):925-31. doi.org/10.1136/thx.2005.040527.

Soler-Cataluña JJ. Severe acute exacerbations and mortality in patients with chronic obstructive pulmonary disease. *Thorax.* 2005;60(11):925-31. doi.org/10.1136/thx.2005.040527.

Soler-Cataluña JJ, Martínez-García MÁ, Serra PC. Impacto multidimensional de las exacerbaciones de la EPOC. *Arch Bronconeumol.* 2010; 46(Supplement 11):12-9. doi.org/10.1016/S0300-2896(10)70057-0.

Soriano JB, Calle M, Montemayor T, Álvarez-Sala JL, Ruiz-Manzano J, Miravittles M. Conocimientos de la población general sobre la enfermedad pulmonar obstructiva crónica y sus determinantes: situación actual y cambios recientes. *Arch Bronconeumol.* 2012;48(9):308-15. doi.org/10.1016/j.arbres.2012.04.008.

Soriano JB, Lamprecht B, Ramírez AS, Martínez-Cambor P, Kaiser B, Alfageme I, et al. Mortality prediction in chronic obstructive pulmonary disease comparing the GOLD 2007 and 2011 staging systems: a pooled analysis of individual patient data. *Lancet Respir Med.* 2015; 3(6): 443-50. doi.org/10.1016/S2213-2600(15)00157-5.

Thokala P, Devlin N, Marsh K, Baltussen R, Boysen M, Kalo Z, Longrenn T, Mussen F, Peacock S, Watkins J, Ijzerman M. Multiple Criteria Decision Analysis for Health Care Decision Making--An Introduction: Report 1 of the ISPOR MCDA Emerging Good Practices Task Force. *Value Health.* 2016; 19(1): 1–13.

Tomatsu S, Alméciga-Díaz CJ, Montañó AM, Yabe H, Tanaka A, Dung VC, et al. Therapies for the bone in mucopolysaccharidoses. *Mol Genet Metab.* 2015; 114(2): 94-109. doi.org/10.1016/j.ymgme.2014.12.001.

Tylki-Szymanska A, Marucha J, Jurecka A, Syczewska M, Czartoryska B. Efficacy of recombinant human α -L-iduronidase (aronidase) on restricted range of motion of upper extremities in mucopolysaccharidosis type I patients. *J Inher Metab Dis*. 2010; 33(2): 151-7. doi.org/10.1007/s10545-010-9059-9.

Valayannopoulos V, Wijburg FA. Therapy for the mucopolysaccharidoses. *Rheumatology*. 2011; 50(5): v49-59. doi.org/10.1093/rheumatology/ker396.

Verma J, Thomas DC, Kasper DC, Sharma S, Puri RD, Bijarnia-Mahay S, et al. Inherited Metabolic Disorders: Efficacy of Enzyme Assays on Dried Blood Spots for the Diagnosis of Lysosomal Storage Disorders. *JIMD Rep*. 2017; 31: 15-27.

World Health Organization (WHO). Enfermedad pulmonar obstructiva crónica (EPOC). 2004 [Accessed 7 december 2017]. Available from: <http://www.who.int/respiratory/copd/es/>.

Wraith JE, Clarke LA, Beck M, Kolodny EH, Pastores GM, Muenzer J, et al. Enzyme replacement therapy for mucopolysaccharidosis I: a randomized, double-blinded, placebo-controlled, multinational study of recombinant human α -L-iduronidase (aronidase). *J Pediatr*. 2004; 144(5): 581-8. doi.org/10.1016/j.jpeds.2004.01.046.

Wyatt K, Henley W, Anderson L, Anderson R, Nikolaou V, Stein K, et al. The effectiveness and cost-effectiveness of enzyme and substrate replacement therapies: a longitudinal cohort study of people with lysosomal storage disorders. *Health Technol Assess*. 2012; 16(39): 1-543. doi.org/10.3310/hta16390.



SUMMARY AND FINAL CONCLUSIONS

Néboa Zozaya González

Juan Oliva Moreno

Álvaro Hidalgo Vega

Bleric Alcalá Revilla

Xavier Badía Llach

Carlos Campillo Artero

Jaime Espín Balbino

Nuria García-Agua Soler

Antonio J. García Ruiz

Antoni Gilabert Perramon

Javier Mar Medina

Carlos Martín Saborido

Jaume Puig Junoy

Marta Traperó Bertran

CHAPTER

11

Decision-making is inherent in the daily work of healthcare professionals, whether in the field of clinical management or in micro, meso or macro management. Deciding involves having to choose between different options, usually in a context with a certain degree of uncertainty, depending on the context. It is not always easy, given the opportunity cost of each case. In democratic societies, there is a growing tendency to expect decisions to be informed by evidence, and that decision-making processes should be as participatory as possible or, at least, that the social preferences delegated to the decision-makers should be reflected in processes with transparent rules which provide for accountability.

This does not imply that the decision derives directly from the decision-making technique. On the contrary, the tools chosen should facilitate decision-making, but this will always be a political issue in which value judgements will intervene, in addition to the evidence. In this sense, and as happens in other sectors (Thokala et al., 2016), MCDA can be useful as an instrument to help decision-making in the healthcare field, under conditions of correct choice and application of the technique. Nevertheless, it must be borne in mind that, despite its growing interest, MCDA is still beginning its journey in the healthcare field, and is not always exempt from criticism.

This book has focused on explaining this method, from a theoretical and practical point of view, so that it can help healthcare professionals (clinicians, managers, researchers, professionals in the pharmaceutical industry, etc.) to understand it, interpret it and apply it correctly. We have seen that MCDA, under good practice conditions, is a structured and explicit approach which could serve to better inform decision-making. It can be applied to all types of healthcare interventions, not only to drugs, but also to devices, surgical techniques, etc. Its use is not limited to a certain area of the decision; it can be used for approval decisions, prioritisation of coverage, financing or access to benefits, classification of diseases, allocation of resources for R&D, etc. MCDA can be used to organise interventions or to evaluate them in isolation or in a comparative manner. In addition, it can be used retrospectively, evaluating alternatives to those to which resources have already been allocated, or prospectively, to evaluate alternatives on which a decision must be made.

Generally, developing an MCDA involves some key steps, which entail the definition of the decision problem, the selection and structuring of the decision criteria, the gathering of evidence, the measurement of the alternatives performance, the weighting and scoring of the criteria, the calculation of the aggregated values, the handling of uncertainty and the interpretation of the results (Thokala et al., 2016). Throughout the book we have seen that there are different methods of weighting, scoring and aggregation, and that the criteria included may vary depending on the objective set.

11.1. MCDA IN RARE VERSUS PREVALENT DISEASES

In the book we have tried to analyse the use of MCDA in rare diseases separately from its use in prevalent diseases, in order to determine whether there are differences between the two approaches. In the case of rare diseases, in which there are often important unmet needs and the intervention is directed at small population groups, there are authors who argue that the appraisal of interventions should consider other criteria in addition to the cost and effectiveness of the interventions (Iskrov et al., 2016; Schlander et al., 2016), and who advocate the use of MCDA for that reason. We have verified that different published MCDA have been concerned

with this type of pathologies, using either fictitious or real drugs. Most of the studies reviewed conclude that MCDA constitutes a methodological framework which incorporates wider decision criteria, and from different perspectives, necessary in decision-making for this type of pathologies. In the field of prevalent diseases, there are currently few bibliographical references which discuss the application of an MCDA, apart from those relating to its application to cancer.

The practical exercise carried out allowed us to appreciate the additional value of the proposed interventions from beyond the economic perspective, and revealed the importance of other attributes, in this case, for a group of experts whose professional profile is focused on health economics. Based on the assumptions made, the drug aimed at the prevalent disease was of greater added value than those aimed at rare diseases. This result highlights the fact that, for the evaluators of this exercise, the cost of treatment, the quality of the scientific evidence and the size of the affected population could have enough weight to counterbalance the value contributed by an effective treatment aimed at a serious pathology which is not very prevalent and has important unmet needs, but also a high cost. Nevertheless, the greatest usefulness of the exercise lies in helping us to reflect on the process of developing MCDA and the robustness and sensitivity of the methodology.

11.2. STRENGTHS AND LIMITATIONS OF MCDA

Once the practical case has been made, and after reviewing the literature, it is possible to identify some of the main strengths and limitations of MCDA, which we refer to below and summarise in Figure 11.1 and Figure 11.2.

Noteworthy among its advantages is the fact that MCDA allows us to analyse the interventions from a broad perspective, explicitly considering different attributes of value. MCDA thus allows us to explicitly take into account a wide range of criteria (in addition to the traditional criteria of efficacy, safety and price) which may be relevant to society when making a decision about an intervention. This may be particularly relevant in the case of some diseases in which the standard criteria may not be sufficient. In addition, among the attributes to be considered, qualitative elements related to ethics, solidarity, etc., can also be included. These are explicitly taken into account in countries such as Sweden and Australia (Paris and Belloni, 2003; Whitty and Littlejohns, 2015), and are also often implicitly considered in economic appraisals.

Also noteworthy is the fact that MCDA helps to formalise the decision process and harmonises the joint work of decision-makers and other stakeholders, allowing the parties involved to organise their ideas and systematise and organise, in an explicit and numerical manner, the healthcare interventions evaluated.

The MCDA framework is postulated by its advocates as a potential help in deciding in a consensual and explicit manner which criteria will be used in the decision, allowing all parties to take into account the same criteria, in a standardised and transparent manner. Another of the most valued aspects of MCDA is that it requires each committee member to explain systematically the reasoning behind the scores which they assign to the evaluated intervention (Goetghebeur, 2012). It should be noted that there are many regulators, including Spanish ones (RD-Law 16/2012, Article 89) which in practice take into account different criteria or elements of value to decide about the inclusion of drugs in public financing. However, the weight given to each criterion is unknown. In this sense, MCDA would contribute to providing greater

transparency and accountability to the decision-making process, by specifying the extent to which each of the pre-defined criteria affects the final value.

The framework also allows a reflective, reasoned and interdisciplinary debate. It can help us to share and understand different perspectives, to clarify the points on which the stakeholders agree and disagree and to find a common language among the parties. From another point of view, if it gives a voice to the different types of agents of the system, it could help in a certain way to democratise the decision-making process. Even when the type of agents involved is restricted, it would serve as a good means of communication between members (of, for example, pharmaco-therapeutic committees). In addition, it can facilitate a thorough understanding of the situation being considered.

FIGURE 11.1. STRENGTHS OF MULTI-CRITERIA DECISION ANALYSIS



Source: authors' preparation.

MCDA is aligned with making decisions based on scientific evidence. The scoring of each criterion must be based on the published scientific evidence, and the quality of that evidence can also be assessed, either directly (as one more explicit criterion in the framework), or indirectly (weighting lower the criterion if it is based on a low methodological quality). In addition, it allows one to consider the non-scientific evidence which may be relevant for the decision. In this way, the framework could facilitate the efficient synthesis of information, make it possible to identify uncertainties, and facilitate the consideration of complex information when making decisions, such as those involving different criteria which may conflict with each other.

MCDA divides the problem into more manageable parts, allowing the data and value judgments to be applied separately to each part, and all or some of them to be aggregated later into a single final value. At each step of the process (weights, scores, aggregate measures), visual representations of the quantitative results which reflect the reasoning of the parties can be made, finally allowing the impacts of the evaluated options to be compared. The ease with which the MCDA is performed will depend, among other things, on the chosen technique.

In short, MCDA would provide an explicit planning framework which can contribute to increasing the transparency, soundness and consistency of decisions, thus improving the quality of decision-making (Baltussen and Niessen, 2006; Angelis and Kanavos, 2016). It would provide contextualisation and better traceability, facilitating greater accountability to citizens and greater understanding by pharmaceutical companies about the rules of the game, and the most valued aspects of innovations. The framework could also help to standardise the criteria on which healthcare decisions are made, and would help agents to be more aligned with a value-based planning of interventions.

However, MCDA also has limitations. One of its main drawbacks is that it does not solve the problem of subjectivity, inherent in all decision-making. The weights and specific scores will depend to a great extent on the chosen evaluators: on their profile, training, experience, individual perception, value judgments, biases and even conflicts of interest, without exhausting the list. For this reason, the external validity of the results will not be evident either, and often the final assessment obtained will not be directly generalisable to other comparators, nor will it last over time (Marsh et al., 2014).

Also, in some cases, there may be some inconsistency or arbitrariness in the assessment of the scoring scales used, if their psychometric capacities are not being adequately considered. To this are added the limitations deriving from the uncertainty about the evidence, as well as the cognitive demands that the analysis can entail, especially when several interventions are evaluated at the same time, with different sensitivity analyses, and in which the excess of complexity can lead to errors or inconsistencies in the appraisal process.

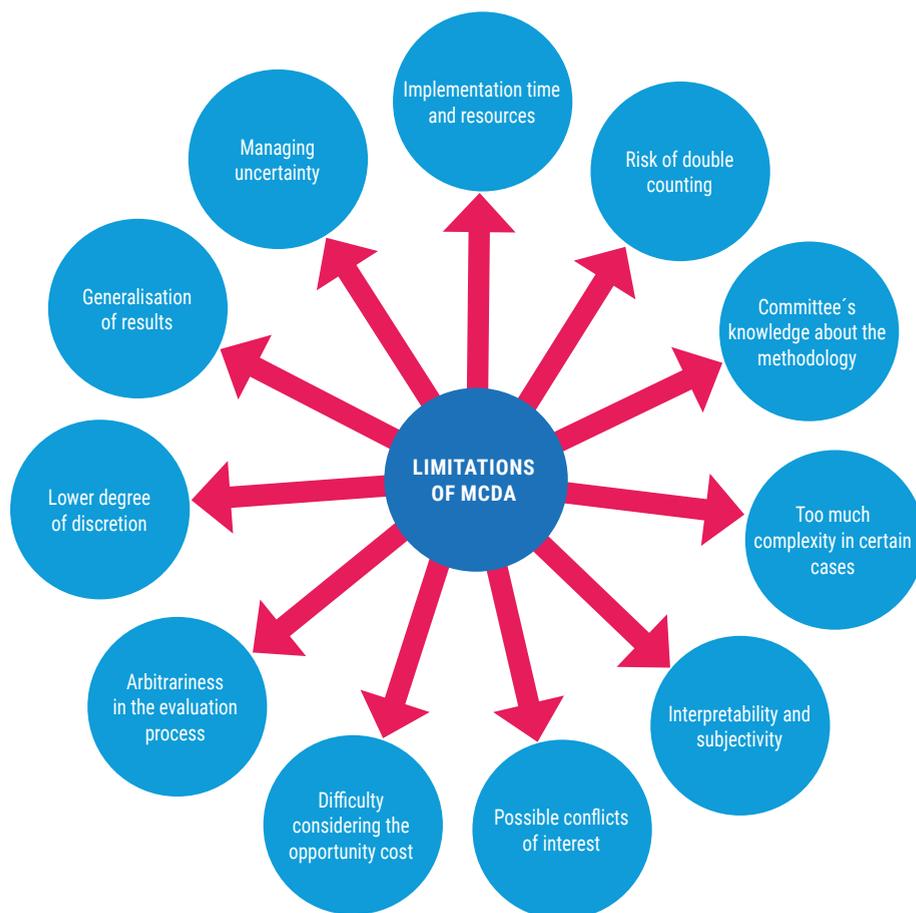
Because it is so specific, the methodology of MCDA should be adequately explained to the members of the evaluation committee, paying special attention when dealing with people who are unfamiliar with the appraisal, such as patients or representatives of the general population. For researchers, the implementation of the analysis will involve time and resources, which should be evaluated when choosing one method or another (Marsh et al., 2016). In addition, the exercise will require the adequate management of the uncertainty associated with it, which can represent a technical challenge for researchers (Marsh et al., 2017).

Another possible criticism of MCDA is its risk of double counting for the same criterion, by incorporating related attributes, such as the cost-effectiveness ratio, when costs and effectiveness have already been incorporated separately. Also, an element to highlight is that some authors consider that this technique is not sufficient when considering the opportunity cost of decisions, as it does not clearly differentiate this concept from accounting cost or healthcare expenditure (Campillo-Artero et al., 2018).

The disclosure of the preferences of the stakeholders involved does not necessarily lead to a better outcome in terms of efficiency and equity, given that healthcare services are affected

by high doses of uncertainty (from the supply and demand side), asymmetry and imperfections in information, positive and negative externalities in production and consumption, and strong socio-economic gradients which affect health. From the point of view of public managers, by making their preferences explicit, the MCDA could reduce their degree of discretion and the degree of freedom in their decisions. On the other hand, there could be an inappropriate use of the results, without forgetting the potential influence that certain pressure groups can have in the choice of the participants and of the criteria to be considered.

FIGURE 11.2. LIMITATIONS OF MULTI-CRITERIA DECISION ANALYSIS



Source: authors' preparation.

Finally, as has been discussed in several sections of this book, it is worth mentioning the existing debate about the complementarity or substitutability of the MCDA and the economic evaluation in HTA. There are authors who consider MCDA as a new evaluation paradigm which could even come to replace the traditional economic appraisal (Kanavos and Angelis, 2013). At the other extreme, there are those who argue that the strengths of MCDA are already included in the economic evaluation (like cost-effectiveness analysis), and that the deficiencies of MCDA (risk of double counting, confusion between elements of value, lack

of transparency, etc.) prevent it from becoming a satisfactory instrument for optimal decision-making (Campillo-Artero et al., 2018). Another school of thought considers that the two methods are complementary, and that MCDA provides certain advantages which can be useful for decision-makers (Garau et al., 2017; Iskov 2016; Goetghebeur et al., 2010).

11.3. CONCLUSIONS AND FINAL RECOMMENDATIONS

1. Although it is still early to know the scope of MCDA in real decision-making, this reflective method is designed to support the culture of decision-making. MCDA can help to structure and reflect about decisions, but in no case can it replace decision-making. Nor is it a substitute for economic evaluation, but it is one more (complementary) tool which decision-makers in healthcare can use to organise their ideas, understand other approaches and support their final decisions.
2. We are at an incipient stage of its development and application in the healthcare field, so there is a small number and a remarkable variability of analysis models, with some methodological limitations. The methodological principles governing the design, conduct and analysis of MCDA still do not enjoy general consensus among the researchers, and there are no consolidated guidelines nor any which are commonly accepted as valid. Thus, one of the pending subjects is to expand the development and use of consolidated guidelines for the design, method, conduct, analysis, interpretation and dissemination of the results of MCDA, preferably on an international scale. One could try to perfect the only existing checklist (Marsh et al., 2016), improving its objectivity and generating more closed and concrete response options to the different questions posed.
3. The selection of the evaluators is one of the bases of the analysis, and will be decisive for the internal and external validity of the results of the MCDA. As such, it must be done with extreme care, paying attention both to the size of the group (not too small to be representative of the real stakeholders, not too big to be manageable, unless it is online) and to the profile of the evaluators. There should be detail of how and why the evaluators were chosen, and an explanation of the potential conflicts of interest, so as to be able to decide to what extent and in what circumstances one should rely on the opinion of those. Note that we are not suggesting that parties who have conflicts of interest should be excluded, because that would exclude, de facto, the opinions of patients and other stakeholders. Also, it will be important to ensure that the profile of the evaluation/debate group is multidisciplinary, so as to give everyone a voice, and that the exchange of opinions among the experts contributes to enrich the analysis and the individual assessments.
4. When tackling an MCDA, researchers should set out what criteria should be considered for that particular problem, without losing sight of the required characteristics: integrity, non-redundancy, non-overlapping and independence of preferences. The number of criteria is not fixed, and will normally be between 3 and 19 (Marsh et al., 2016). The selection of criteria can be done ad hoc, or by using an already existing framework, such as the EVIDEM. Using a pre-established approach such as the EVIDEM has advantages and disadvantages. On the one hand, it provides a set of validated and proven criteria, including those of a qualitative nature, it saves time and facilitates comparisons between studies. On the other hand, the EVIDEM framework can be less flexible and appropriate than one created ad hoc for the decision problem.

5. The MCDA encompasses different methods, and there is no consensus about which one is preferable, since all have advantages and limitations. The method to be applied in each case will depend on the objective of the study, the time and resources available, as well as on the training of the researchers. It would be desirable to deepen the methodological development of MCDA, to allow more solid validation exercises and comparisons between methods which help researchers to choose the most appropriate option.
6. The identification and management of uncertainty are key elements which all MCDA must deal with and manage in the best possible way. Any decision is subject to a certain degree of uncertainty, and the greater the uncertainty, the more difficult it will be to face the situation and make the right decision. The uncertainty extends to different phases of the MCDA, through different elements. Although it is not always possible, or even desirable, to reduce uncertainty, it must be made explicit, and one should try to quantify, through deterministic and/or probabilistic sensitivity analysis, how it may be affecting the results of the analysis.
7. To counteract the cognitive difficulty which MCDA can entail methodologically, especially when dealing with several interventions simultaneously, with their respective sensitivity analyses, it is recommended to have an understanding of the tool and the logical order of the scores in the different scenarios of sensitivity. For this, exhaustive methodological explanations prior to the meeting and during it will be fundamental components, especially for the patients, who are less accustomed to analysing the results of clinical trials and to understanding clinical and economic concepts (providing a glossary of terms could help, as could asking the evaluators for their feedback about the process, to improve in the future).
8. Beyond further improving the methodological aspects of MCDA, it would also be desirable to deepen its applicability in actual practice, either through retrospective exercises which compare real decisions with those resulting from the MCDA, or by means of multi-criteria evaluations which help to make decisions in real time, giving public dissemination to these exercises.

There is no perfect tool capable of answering all the questions that arise in a decision-making process in the field of healthcare. Nevertheless, well applied, MCDA can promote greater transparency, consistency and coherence in decision-making in healthcare. If one can overcome the specific limitations in the conception and initial applications of this method when applied to healthcare, MCDA could offer an additional methodological tool which serves as a useful complement in the evaluation of healthcare interventions. We encourage readers, and especially decision-makers, to engage in practical exercises which enable them to experiment with the method, to better understand its strengths and limitations, and to consider incorporating it as a more supportive tool in decision-making.

REFERENCES

- Angelis A, Kanavos P. Value-Based Assessment of New Medical Technologies: Towards a Robust Methodological Framework for the Application of Multiple Criteria Decision Analysis in the Context of Health Technology Assessment. *PharmacoEconomics*. 2016; 34(5):435–46.
- Baltussen R, Niessen L. Priority setting of health interventions: the need for multi-criteria decision analysis. *Cost Eff Resour Alloc*. 2006; 4(1):14.
- Campillo-Artero C, Puig-Junoy J, Culyer AJ. Does MCDA Trump CEA? *Appl Health Econ Health Policy*. 2018; 16(2): 147-151.
- Garau M, Devlin NJ. Using MCDA as a Decision Aid in Health Technology Appraisal for Coverage Decisions: Opportunities, Challenges and Unresolved Questions. In: *Multi-Criteria Decision Analysis to Support Healthcare Decisions*. Springer, Cham. 2017; 277–98. Disponible en: https://link.springer.com/chapter/10.1007/978-3-319-47540-0_14.
- Goetghebeur MM, Wagner M, Khoury H, Rindress D, Grégoire J-P, Deal C. Combining multi-criteria decision analysis, ethics and health technology assessment: applying the EVIDEM decisionmaking framework to growth hormone for Turner syndrome patients. *Cost Eff Resour Alloc*. 2010; 8(1):4.
- Goetghebeur MM, Wagner M, Khoury H, Levitt RJ, Erickson LJ, Rindress D. Bridging health technology assessment (HTA) and efficient health care decision making with multi-criteria decision analysis (MCDA): applying the EVIDEM framework to medicines appraisal. *Med Decis Making*. 2012 Mar-Apr; 32(2): 376-88. doi: 10.1177/0272989X11416870. Epub 2011 Oct 10.
- Iskrov G, Miteva-Katrandzhieva T, y Stefanov R. 2016. Multi-Criteria Decision Analysis for Assessment and Appraisal of Orphan Drugs. *Frontiers in Public Health* 4. doi.org/10.3389/fpubh.2016.00214.
- Kanavos P, Angelis A. Multiple Criteria Decision Analysis for Value Based Assessment of New Medical Technologies: A Conceptual Framework. *LSE Work Pap*. 2013;33.
- Marsh K, Goetghebeur M, Thokala P, Baltussen R. *Multi-Criteria Decision Analysis to Support Healthcare Decisions*. Springer; 2017.
- Marsh K, Ijzerman M, Thokala P, Baltussen R, Boysen M, Kaló Z, Lönngren T, Mussen F, Peacock S, Watkins J, Devlin N. Multiple Criteria Decision Analysis for Health Care Decision Making—Emerging Good Practices: Report 2 of the ISPOR MCDA Emerging Good Practices Task Force. *Value Health*. 2016; 19(2): 125-137.
- Marsh K, Lanitis T, Neasham D, Orfanos P, Caro J. Assessing the Value of Healthcare Interventions Using Multi-Criteria Decision Analysis: A Review of the Literature. *PharmacoEconomics*. 2014; 32(4): 345–65.
- Paris, V, Belloni, A. Value in Pharmaceutical Pricing. 2013. Report No.: 63. Disponible en: http://www.oecd-ilibrary.org/social-issues-migration-health/value-in-pharmaceutical-pricing_5k43jc9v6knx-en.
- Real Decreto-ley 16/2012, de 20 de abril, de medidas urgentes para garantizar la sostenibilidad del Sistema Nacional de Salud y mejorar la calidad y seguridad de sus prestaciones. *BOE Num*. 98.
- Schlander, M., Garattini, S., Kolominsky-Rabas, P., Nord, E., Persson, U., Postma, M., Richardson, J., Simoons, S., de Solà Morales, O., Tolley, K., and Toumi, M. Determining the value of medical technologies to treat ultra-rare disorders: a consensus statement. *J Mark Access Health Policy*. 2016; Oct 27; 4: 10.3402.
- Thokala P, Devlin N, Marsh K, Baltussen R, Boysen M, Kalo Z, Longrenn T, Mussen F, Peacock S, Watkins J, Ijzerman M. Multiple Criteria Decision Analysis for Health Care Decision Making—An Introduction: Report 1 of the ISPOR MCDA Emerging Good Practices Task Force. *Value Health*. 2016; 19(1): 1–13.
- Whitty JA, Littlejohns P. Social values and health priority setting in Australia: An analysis applied to the context of health technology assessment. *Health Policy*. 2015 Feb 1; 119(2): 127–36.
-

Sponsored by:

