DEVELOPMENT OF A VALUE ASSESSMENT FRAMEWORK FOR HEALTH TECHNOLOGY ASSESSMENT AND COVERAGE DECISION MAKING IN CHINA

DEVELOPMENT OF A VALUE ASSESSMENT FRAMEWORK FOR HEALTH TECHNOLOGY ASSESSMENT AND COVERAGE DECISION MAKING IN CHINA

By MENGMENG ZHANG, BSc., MSc.

A Thesis Submitted to the School of Graduate Studies in Partial Fulfilment of

the Requirements for the Degree

Doctor of Philosophy

DOCTOR OF PHILOSOPHY (2023)	McMaster University
Health Research Methodology	Hamilton, Ontario
Title	Development of a value assessment framework for health technology assessment and coverage decision making in China
Author	Mengmeng Zhang B.Sc. Peking University, 2012 M.Sc. Peking University, 2014
Supervisor	Dr. Feng Xie
Committee	Dr. Melissa Kimber Dr. Mitchell Levine
Number of pages	xii, 192

LAY ABSTRACT

Efficient allocation of limited healthcare resources has been a challenge for many healthcare systems around the world. In recent years, a decision-support tool called value assessment framework (VAF) has become promising in assessing health technologies' (e.g., drugs, medical devices) value and supporting value-based coverage decision-making. We developed a VAF for China by 1) identifying dimensions that are important for value assessment through a review of existing VAFs 2) conducting open-ended interviews with 34 Chinese stakeholders, and 3) developing scoring methods for the VAF. Using the developed VAF, decision makers can estimate the value of drugs to be assessed and whether the drug is covered by insurance based on the drug's performance. Thus, decision makers can make more transparent and consistent coverage decisions to promote the use of health technologies with high value in China.

ABSTRACT

Value assessment framework (VAF) has become a promising tool for assessing the value of health technologies and informing coverage decision making. Most VAFs have been developed for high-income countries and are insufficient for various contexts given. There were limited patient and public engagement in the framework development process and the uncertainty in coverage decision making was not accommodated. This doctoral thesis aimed to develop a VAF that involved multiple stakeholders to support transparent and consistent coverage decision making in China.

This thesis begins with an overview of coverage decision making and health technology assessment (HTA), and the emergence and application of VAF in this field in recent years. This thesis subsequently presents a systematic review of existing VAFs that investigated how value is defined and measured in healthcare and summarized the methods of framework development in existing VAFs. Then, this thesis presents a qualitative description study informed by the systematic review and the principles of qualitative description (QD). Through open-ended semi-structured interviews with 34 Chinese stakeholders, as well as a review and analysis of 16 publicly available government documents related to HTA and coverage policies in China, 12 value attributes were identified for the development of a VAF in China. Then, this thesis includes an online factorial survey among 365 Chinese stakeholders to generate value scoring algorithms. With the developed VAF, the value of a health technology under assessment and its probabilities of entering negotiation or being covered by the national medical insurance in China for diseases with different levels of severity, can be estimated. This thesis ends with a discussion of the key findings, limitations, and implications of this program of research and presents our perspectives on challenges and future directions in the field of VAF.

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my supervisor Dr. Feng Xie. The complete trust and extreme freedom he has given me allowed me to choose projects that interested me and to explore many possibilities for my career. His continuous support encouraged me to push myself past my limits and embrace my weaknesses. His extraordinary vison and expertise and passionate commitment to research and teaching have inspired and will always inspire me.

I want to thank my supervisory committee members, Dr. Melissa Kimber, and Dr. Mitchell Levine for their valuable advice, professional guidance, and extreme support. They have always been available throughout my PhD study, from preparing applications to obtain ethics approval, collecting and analyzing data and refining the reports, to sourcing the appropriate software for my project. It has been such a pleasure to work with them and learn from them.

I want to thank Dr. Lehana Thabane for reminding me to be proud of myself and Dr. Jinhui Ma for all the valuable advice on my independent study project. I also want to thank all the researchers I have worked with and all the research participants in this thesis. I want to thank Li and Peter at the research data center for giving me so much kind help and support.

I would also like to thank my fellow graduate students for the warm friendship and happy collaborations for courses and projects.

Finally, I want to thank my husband, Zhikang for his continuous love, support and understanding. I am so lucky to have him in my life. I also want to thank my boys Guanzhong and Guanzhang for giving me so many wonderful distractions and making me the happiest mother in the world.

This thesis is dedicated to Zhikang, Guanzhong and Guanzhang.

TABLE OF CONTENTS

LIST OF FIGURES

Chapter 2

Chapter 4

LIST OF TABLES

Chapter 2

Chapter 3

Chapter 4

LIST OF ABBREVIATIONS

Abbreviation	Definition
AHP	Analytic hierarchy process
AIC	Akaike information criterion
ASCO	American Society of Clinical Oncology
BIA	Budget impact analysis
BWS	Best worst scaling
CBA	Cost-benefit analysis
CEA	Cost-effectiveness analysis
CI	Confidence interval
COVID-19	Coronavirus disease 2019
CUA	Cost-utility analysis
DCE	Discrete choice experiment
ESMO	European Society of Medical Oncology
EVIDEM	The Evidence and Value: Impact on DEcisionMaking framework
GDP	Gross domestic product
GRADE	Grading of Recommendations Assessment, Development, and Evaluation
GRA-QALY	Generalized risk-adjusted quality-adjusted life years
HiREB	Hamilton Integrated Research Ethics Board
HTA	Health technology assessment
ICER	Incremental cost-effectiveness ratio
IDA	Initial data analysis
IOM	Institute of Medicine
IQRs	Interquartile ranges
ISPOR	International Society for Pharmacoeconomics and Outcomes Research
MCDA	Multi-criteria decision analysis
MCDM	Multi-criteria decision making
MSKCC	Memorial Sloan Kettering Cancer Center
NCCN	National Comprehensive Cancer Network
NHSA	National Healthcare Security Administration
PAPRIKA	Potentially All Pairwise RanKings of all possible Alternatives
PROs	Patient-reported outcomes
QALY	Quality-adjusted life-year
QD	Qualitative description
QoL	Quality of life
SD	Standard deviation
SMART	Simple Multi-Attribute Rating Technique
SRQR	Standards for Reporting Qualitative Research
UEBMI	Urban employee basic medical insurance
URRBMI	Urban-rural resident basic medical insurance
VAF	Value assessment framework
WHO	World Health Organization
WTP	Willingness to pay

DECLARATION OF ACADEMIC ACHIEVEMENT

This sandwich thesis consists of three manuscripts. At the time of writing the thesis, one manuscript has been published, one has been accepted for publication, and one under preparation for publication. MZ is the first author of all three manuscripts. Chapters 1&5 are unpublished and MZ is the sole author.

Chapter 2 presents a systematic review of existing VAFs and has been published in *Value in Health* on February 1, 2022. MZ led the research and contributed to the conception and design of the study, preparation of data collection forms, literature screening, data extraction and interpretation, drafting the manuscript, coordinating submissions to journals for peer review and incorporating feedback into the final manuscript. FX contributed to the conceptualization and design of the work. YB, YL, SF, MK, ML and FX contributed to the acquisition, analysis and interpretation of data, critical revision of the manuscript and statistical analysis.

Chapter 3 presents a qualitative description study for attribute identification and has been accepted for publication in *Pharmacoeconomics* on December 22, 2022. MZ was the primary researcher responsible for conceptualizing and designing the studies; preparing the protocol and obtaining approval from the Hamilton Integrated Research Ethics Board (HiREB, project No. 12993); preparing study materials including forms, and questionnaires; recruiting and screening research participants; collecting, analysing and interpreting data; drafting the manuscript; coordinating submissions to journals for peer review; and preparing responses to editors' and reviewers' comments and feedback. FX and MK contributed to the conceptualization and design of study. YB, YY, MK, ML and FX contributed to participant recruitment, data collection, analysis and interpretation and critical revision of the manuscript.

Chapter 4 presents a factorial survey experiment for developing scoring functions of the VAF which is under preparation for publication. MZ led the research and was the primary researcher responsible for conceptualizing and designing the studies; preparing the protocols and obtaining approvals from the HiREB (Project No. 14710); preparing study materials including forms, questionnaires, and surveys; recruiting research participants; collecting, analysing and interpreting data; drafting the manuscript; and coordinating submissions to journals for peer review. FX contributed to the conceptualization and design of the work. YY, YB, MK, ML and FX contributed to the recruitment of participants, acquisition, analysis and interpretation of data, and critical revision of the manuscript.

CHAPTER 1. INTRODUCTION

HTA, value in healthcare and coverage decision making

Health technologies are interventions developed for the prevention, diagnosis, treatment and management of medical conditions.[1] The interventions include drugs, devices, medical procedures, and programs.[1] They have been a mainstay of healthcare, significantly contributing to improved population health outcomes.[2] For instance, the use of antibacterial agents have saved millions of lives around the world.[2] In addition, treatments and secondary preventive therapies for coronary heart disease led to approximately half of the mortality reduction of this disease in the US from 1980 through 2000.[3] However, there is substantial variation among health technologies in terms of the degree of their benefits relative to the risks they could pose to patients.[2,4] Some technologies are ineffective and their risks outweigh the benefits; some examples include testing for C-reactive protein and liver function, spinal fusion for non-specific low back pain, and vertebroplasty for osteoporotic fractures, among others.[2,5] On the other hand, the progress and diffusion of health technologies has become a main driver of the rapid growth of healthcare expenditures.[2] Global health expenditures have more than doubled over the past 20 years, reaching US$ 8.5 trillion in 2019.[6] Its growth has outpaced economic growth, increasing from 8.5 to 9.8% of global gross domestic product (GDP) over the same period of time.[6] Previous studies have estimated that technological progress accounts for 35% of total healthcare expenditure growth on average.[7] The mechanisms for this impact include, but are not limited to, rising costs of new technologies, expanding volume of old and new services, extended life expectancy, and increased disease diagnosis.[2,7]

With scarce resources, variable degree of benefits and risks delivered by health technologies and escalating healthcare expenditures, it is neither feasible nor rational to pay for everything entering into healthcare markets.[8,9] Choices must be made through coverage decision making which regulates the resource allocation process and determines the reimbursement or coverage of technologies through medical insurance.[8,9] For a given health technology, manufactures seek to achieve fast market entry, patients value access to the technology and its health benefits, while payers manage the budget impact.[10,11] With the purpose of balancing multiple objectives, coverage decision making is inherently complex, an instrument to assess the value of technologies which can advise health technology-related coverage decision making is needed.

Health technology assessment (HTA), first introduced in the 1970s, has served as such a tool to couple evidence with decision making and has been increasingly adopted to inform coverage decision making in organizations and jurisdictions worldwide.[2,12–14] HTA is a multidisciplinary process that uses explicit methods to rigorously review and synthesize evidence related to the health technology.[13,15] The evidence assessed in HTA was described as the "short- and long-term consequences of the application of a technology" in its early years, then as "the properties, effects or impact of health technologies" in the 1990s, the "direct and intended effects and indirect and unintended consequences" later, and most recently the "value of a health technology".[12,15,16]

Over the past decade, value in healthcare has been increasingly discussed by stakeholders for coverage decision making. Terms such as "paying for value" and "value-based payment" have been often used in the context of HTA and coverage decision making.[17–19] However,

there is no global consensus on the definition of value in healthcare which highly depends on how value is measured and which perspectives are adopted (e.g., patients, the healthcare system or the society).[20,21] From the economic perspective, the value in healthcare can be measured in different ways. Under welfarism, the aim of healthcare is to generate improvement in global welfare (well-being) which occurs when "one individual can be made better off without any other being made worse off".[22] Based on this compensation principle, cost-benefit analysis (CBA) compares a health technology's costs and benefits, which are measured by the amount of money individuals are willing to pay for. However, this approach is limited due to the difficulty of valuing health benefits in monetary terms.[22] Under extra-welfarism, value is measured by the health improvements obtained with constrained budget.[22] Cost-utility analysis (CUA) is the most commonly used economic analysis technique under this approach.[22] Health improvements or health gains of health technology are measured using quality-adjusted life-years (QALY), an aggregate metric combining length of life and quality of life.[23] A QALY is one year in perfect health.[23] The cost-effectiveness or the value of the health technology is estimated using the incremental cost-effectiveness ratio (ICER), i.e. the ratio of incremental costs to incremental health gains (e.g., measured in QALYs in CUA) of the new technology compared to the reference technology.[24] By comparing the estimated ICERs to a predefined threshold, decision makers are able to compare and make choices among health technologies across different diseases.[24] Despite that the extra-welfarist approach has been widely adopted by many HTA agencies, it is criticized for its sole goal of health maximization and failure of capturing societal aspects, and population preferences concerning the distribution of the

costs and benefits associated with the use of the technology across populations.[25] In the most recent definition of HTA, value is multi-dimensional, referring to the technology's "clinical effectiveness, safety, costs and economic implications, ethical, social, cultural and legal issues, organizational and environmental aspects, as well as wider implications for the patient, relatives, caregivers, and the population".[26] In spite of using societal perspectives in CUA to incorporate non-health costs and outcomes, broader value dimensions (e.g., severity of disease) are still more likely left excluded in current methods.[24,25,27] This lack of inclusiveness has left coverage decision making based on HTA results prone to ambiguity, inconsistency, and subjectivity. To incorporate relevant value dimensions into HTA and facilitate transparent, consistent, and objective coverage decision making, a broader, comprehensive "value assessment framework" has been proposed.[28,29]

VAFs for HTA and coverage decision making

Value assessment frameworks (VAFs) have emerged as a response to the rising healthcare expenditures and the shift of our health system from "volume-based" to "value-based".[30,31] It provides a relatively structured approach that includes all relevant value aspects and aggregates them in a relatively explicit way so that value assessment and coverage decision-making can be completed simultaneously.[28,32]

VAFs are considered promising tools to measure and communicate the value of health technologies.[32] In recent years, a few institutions and organizations, including the American Society of Clinical Oncology (ASCO), the Memorial Sloan Kettering Cancer Center (MSKCC), the National Comprehensive Cancer Network (NCCN), and the European Society of Medical Oncology (ESMO), and HTA bodies such as the Institute for

Clinical and Economic Review have developed VAFs to support health technology value assessment and inform decision making.[33–37] The oncology-oriented frameworks developed by ASCO, ESMO and NCCN are for the shared decision making between healthcare providers and patients.[33,35,36] The other two frameworks (i.e., the MSKCC and the Institute for Clinical and Economic Review frameworks) are for coverage and reimbursement decisions.[34,37] These VAFs represent the considerable effort that has been devoted to investigating value dimensions important to different stakeholders affected by healthcare decision making and innovative methods to aggregate the value dimensions such as estimation of net benefit, determination of a target price, grading and deliberation.[33–37] Nevertheless, these frameworks are mostly disease-specific with limited patients or public involvement in framework development.[20] The methods for aggregation of value dimensions also remain unvalidated.[38]

Patient and public engagement is vital for value assessment.[32] Although the end user of VAFs is the payers, decisions made through the VAF will ultimately affect patients and members of the public (hereafter collectively referred to as the public), the recipients of healthcare services.[32,39] Besides, compared to payers who tend to define value from a population perspective, the public's interpretation of value is individualized, reflecting personal perceptions and experiences.[39] Outcomes important to payers could include budget impact, cost-effectiveness and societal impact.[10,11] In contrast, the public values outcomes such as accessibility and quality of life.[11,40] Thus, it is critical to explore different stakeholders' perspectives on value and develop a VAF with multiple stakeholder engagement.[32,39] A few institutions have developed patient-centered VAFs or principles for

partnering with the public to include patient perspectives into value assessment.[41,42] The National Health Council in the US has released the Patient-Centered Value Model Rubric to promote patient-centeredness in the development and dissemination processes of VAF.[39] Despite these efforts, engagement of public is insufficient in current VAFs which calls for methods to meaningfully work with the public.[32]

For VAFs comprised of multiple value attributes, an important feature is the approaches to aggregate the attributes. Multi-criteria decision analysis (MCDA) has been used as an alternative approach to traditional CUA/CEA and deliberation.[20,25,43] This approach originated in the discipline of operational research is concerned with decision making situations where multiple dimensions are to be combined or aggregated.[43] MCDA has been increasingly explored in healthcare decision making and adopted or tested by various HTA agencies.[20,44–49] MCDA involves various methods that differ in how the attributes are combined. The most commonly used methods are those adopting the weighted-sum models where each health technology is assigned a numerical value based on the technology's performance on the attributes and the weights of the attributes through a value function or scoring algorithm.[20,43] In spite of its wide use in VAFs, the weighted-sum model is limited in 1) its inability to account for the possible dependence between value attributes, 2) the lack of appropriate thresholds to advise decision-making and 3) the fixed mechanism to accommodate uncertainty in coverage decision making.[25,50,51] The additive nature of the weighted-sum model allows compensation between attributes and assumes independence and non-overlapping characteristics between attributes, which is often violated in real world applications.[25,50] For example, the value of a new intervention with substantial health

gain but fair safety profiles could be similar to the value of a control with small health gain but good safety profiles. Also, the value of the same health gain varies with the severity of the disease, which is not incorporated in an additive model in which severity of disease and health gain are two independent attributes.[52] The single index generated through the weighted-sum model requires a combination with acceptable thresholds to facilitate decision-making.[25] However, this approach remains a subject of ongoing debate.[25] Due to the fact that there are a variety of perspectives and needs to incorporate a large volume of information with uncertainty to process in coverage decision making, the decisions made may not always be a simple "yes" or "no".[53]

China's healthcare system and coverage decision making

China is the most populous country in the world with 1.44 billion residents and 12% (~172 million) over 65 years old.[54] China's healthcare system is facing unprecedented challenges in meeting the healthcare needs for its population. It was estimated that the healthcare spending in China accounted for 6.6% of gross domestic product (GDP) in 2019.[55,56] With the current healthcare delivery model, primarily volume-based, the expenditure is expected to exceed 9% by 2035.[55,56] Reforms on the healthcare delivery model is needed to reduce the use of unnecessary or non-beneficial procedures and treatments and ease the pressure from increasing healthcare expenditures.

As part of China's latest healthcare reforms, the National Healthcare Security Administration (NHSA) was established to adopt a centralized approach to drug pricing and coverage decision making.[57,58] The NHSA updates the National Medical Insurance drug list annually. Drugs that are newly approved or which were not included in the drug

list previously, may be covered, enter the negotiation, or be rejected by the NHSA based on deliberation of the drugs' performance on multiple dimensions by a committee comprised of physicians, health economists, and policymakers.[59] There are a large number of drugs to be assessed each year and the value of drugs is considered, albeit informally, in the decision-making process.

Given the lack of a formal VAF in China to account for the population preferences and resources available in the local setting, a VAF that is suitable for China's health system and society is needed to guide the value assessment of new technologies and to support consistent and efficient coverage policy making. In recent years, some Chinese researchers have adapted the Evidence and Value: Impact on DEcisionMaking (EVIDEM) framework to support HTA in China.[60,61] The EVIDEM was developed through a literature review; some of the resulting attributes were noted as context sensitive.[62] Moreover, when EVIDEM was adapted to China's contexts, the stakeholder engagement was limited, more specifically, the input from the general public was omitted.[60,61] Without the engagement of patients and the general public in VAF development, healthcare decision making could be compromised as it would not incorporate the patients and the general public's perspectives about the value of health technologies.[32] Consequently, the uptake and efficiency of health technologies in real-world practice could be impacted. For example, from 2009 to 2017, China's government subsidies to primary care programs accounted for an increase from 12.3 to 32.5% of primary care revenue.[63] However, from 2012 to 2017, the proportion of services provided by primary healthcare facilities decreased by 7%.[63] Except for some

historic and institutional factors, lack of understanding of the public's preferences plays an important role in this misalignment between government decisions and the public's choices.

The goal of this program of research was to develop a VAF for HTA and coverage decision making in China. There are three specific objectives: 1) identification and selection of candidate value attributes; 2) aggregation and scoring of selected attributes and 3) testing and validation of a VAF in practice to support healthcare decision making. In this thesis, we have accomplished the first two stages of the study and developed a VAF.

Outline of thesis

This is a sandwich thesis of three papers presented in Chapters 2-4 covering the attribute identification and aggregation processes to develop a VAF in China.

Chapter 2 systematically reviews and summarizes the existing VAFs to investigate how value is defined and measured in healthcare. The findings of this systematic review have provided insights into how the value attributes have been identified and aggregated in existing VAFs and informed the design of the subsequent studies.

Chapter 3 presents the process of attribute identification using the methods of qualitative description. Through 34 open-ended semi-structured interviews and a review and analysis of 16 government documents, we have identified 12 value attributes deemed important to Chinese stakeholders. These attributes were used for the development of a VAF incorporating perspectives of multiple stakeholders.

Chapter 4 describes the process of developing scoring algorithms for the VAF using the methods of factorial survey and MCDA. The relative weights of attributes and the scoring

algorithm for value assessment were developed. Using the VAF, the value of a health technology and its probabilities of negotiation or coverage for diseases of different levels of severity can be predicted.

Chapter 5 discusses the main findings and limitations of the thesis and its policy implications. Challenges and opportunities related to VAF development and application are also discussed.

References

1. HTA Glossary. HtaGlossary.net | health technology. HtaGlossary.net. Accessed June 24, 2022. http://htaglossary.net/health-technology

2. The Organisation for Economic Co-operation and Development (. New Health Technologies: Managing Access, Value and Sustainability | en | OECD. Published 2017. Accessed June 24, 2022. https://www.oecd.org/health/managing-new-technologies-in-health-care-9789264266438-en.htm

3. Ford ES, Ajani UA, Croft JB, et al. Explaining the decrease in U.S. deaths from coronary disease, 1980-2000. *N Engl J Med*. 2007;356(23):2388-2398. doi:10.1056/NEJMsa053935

4. Taylor RS, Drummond MF, Salkeld G, Sullivan SD. Inclusion of cost effectiveness in licensing requirements of new drugs: the fourth hurdle. *BMJ*. 2004;329(7472):972-975. doi:10.1136/bmj.329.7472.972

5. Elshaug AG. Over 150 potentially low-value health care practices: an Australian study. *Med J Aust*. 2013;198(2):85. doi:10.5694/mja12.11770

6. World Health Organization. Global Health Expenditure Database. Published 2022. Accessed June 29, 2022. https://apps.who.int/nha/database

7. Marino A, Lorenzoni L. *The Impact of Technological Advancements on Health Spending: A Literature Review*. OECD; 2019. doi:10.1787/fa3bab05-en

8. Sood N, de Vries H, Gutierrez I, Lakdawalla DN, Goldman DP. The effect of regulation on pharmaceutical revenues: experience in nineteen countries. *Health Aff (Millwood)*. 2009;28(1):w125-137. doi:10.1377/hlthaff.28.1.w125

9. World Health Organization. Regional Office for Europe, Policies EO on HS and, Velasco Garrido M, Kristensen FB, Nielsen CP, Busse R. *Health Technology Assessment and Health Policy-Making in Europe: Current Status, Challenges and Potential*. World Health Organization. Regional Office for Europe; 2008. Accessed September 30, 2022. https://apps.who.int/iris/handle/10665/107911

10. Fischer KE. A systematic review of coverage decision-making on health technologies-evidence from the real world. *Health Policy*. 2012;107(2-3):218-230. doi:10.1016/j.healthpol.2012.07.005

11. Institute of Medicine (US) Roundtable on Value & Science-Driven Health Care. *Value in Health Care: Accounting for Cost, Quality, Safety, Outcomes, and Innovation*. (Yong PL, Olsen L, McGinnis JM, eds.). National Academies Press (US); 2010. Accessed May 20, 2020. http://www.ncbi.nlm.nih.gov/books/NBK50929/

12. O'Rourke B, Oortwijn W, Schuller T. Announcing the New Definition of Health Technology Assessment. *Value in Health*. 2020;23(6):824-825. doi:10.1016/j.jval.2020.05.001

13. Banta D. The development of health technology assessment. *Health Policy*. 2003;63(2):121-132. doi:10.1016/S0168-8510(02)00059-3

14. Kristensen FB, Husereau D, Huić M, et al. Identifying the Need for Good Practices in Health Technology Assessment: Summary of the ISPOR HTA Council Working Group Report on Good Practices in HTA. *Value Health*. 2019;22(1):13-20. doi:10.1016/j.jval.2018.08.010

15. Banta D. What is technology assessment? *International Journal of Technology Assessment in Health Care*. 2009;25(S1):7-9. doi:10.1017/S0266462309090333

16. Draborg E, Gyrd-Hansen D, Poulsen PB, Horder M. International comparison of the definition and the practical application of health technology assessment. *Int J Technol Assess Health Care*. 2005;21(1):89-95. doi:10.1017/s0266462305050117

17. Porter ME. What is value in health care? *N Engl J Med*. 2010;363(26):2477-2481. doi:10.1056/NEJMp1011024

18. Landon S n, Padikkala J, Horwitz L i. Defining value in health care: a scoping review of the literature. *International Journal for Quality in Health Care*. 2021;33(4):mzab140. doi:10.1093/intqhc/mzab140

19. Teisberg E, Wallace S, O'Hara S. Defining and Implementing Value-Based Health Care: A Strategic Framework. *Acad Med*. 2020;95(5):682-685. doi:10.1097/ACM.0000000000003122

20. Zhang M, Bao Y, Lang Y, et al. What Is Value in Health and Healthcare? A Systematic Literature Review of Value Assessment Frameworks. *Value Health*. 2022;25(2):302-317. doi:10.1016/j.jval.2021.07.005

21. Garrison LP, Pauly MV, Willke RJ, Neumann PJ. An Overview of Value, Perspective, and Decision Context-A Health Economics Approach: An ISPOR Special Task Force Report [2]. *Value Health*. 2018;21(2):124-130. doi:10.1016/j.jval.2017.12.006

22. Coast J, Smith RD, Lorgelly P. Welfarism, extra-welfarism and capability: The spread of ideas in health economics. *Social Science & Medicine*. 2008;67(7):1190-1198. doi:10.1016/j.socscimed.2008.06.027

23. Weinstein MC, Torrance G, McGuire A. QALYs: the basics. *Value Health*. 2009;12 Suppl 1:S5-9. doi:10.1111/j.1524-4733.2009.00515.x

24. Neumann PJ, Sanders GD, Basu A, Brock DW, Feeny D, Krahn M. *Cost-Effectiveness in Health and Medicine*. Oxford University Press; 2016.

25. Phelps CE, Lakdawalla DN, Basu A, Drummond MF, Towse A, Danzon PM. Approaches to Aggregation and Decision Making-A Health Economics Approach: An ISPOR Special Task Force Report [5]. *Value Health*. 2018;21(2):146-154. doi:10.1016/j.jval.2017.12.010

26. O'Rourke B, Oortwijn W, Schuller T, International Joint Task Group. The new definition of health technology assessment: A milestone in international collaboration. *Int J Technol Assess Health Care*. 2020;36(3):187-190. doi:10.1017/S0266462320000215

27. Drummond M. *Methods for the Economic Evaluation of Health Care Programmes*. Fourth edition. Oxford University Press; 2015.

28. Neumann PJ, Willke RJ, Garrison LP. A Health Economics Approach to US Value Assessment Frameworks-Introduction: An ISPOR Special Task Force Report [1]. *Value Health*. 2018;21(2):119-123. doi:10.1016/j.jval.2017.12.012

29. Riva S, Pravettoni G. Value-Based Model: A New Perspective in Medical Decision-making. *Front Public Health*. 2016;4. doi:10.3389/fpubh.2016.00118

30. Centers for Medicare and Medicaid Services. CMS' Value-Based Programs | CMS. Published January 12, 2021. Accessed January 16, 2022. https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/Value-Based-Programs/Value-Based-Programs

31. Miller HD. From Volume To Value: Better Ways To Pay For Health Care. *Health Affairs*. 2009;28(5):1418-1428. doi:10.1377/hlthaff.28.5.1418

32. Dubois RW, Westrich K. As Value Assessment Frameworks Evolve, Are They Finally Ready for Prime Time? *Value Health*. 2019;22(9):977-980. doi:10.1016/j.jval.2019.06.002

33. Schnipper LE, Davidson NE, Wollins DS, et al. Updating the American Society of Clinical Oncology Value Framework: Revisions and Reflections in Response to Comments Received. *J Clin Oncol*. 2016;34(24):2925-2934. doi:10.1200/JCO.2016.68.2518

34. Memorial Sloan Kettering Cancer Center. DrugAbacus. Drug Pricing Lab. Accessed February 11, 2020. https://drugpricinglab.org/tools/drug-abacus/

35. National Comprehensive Cancer Network. NCCN Clinical Practice Guidelines in Oncology (NCCN Guidelines) with NCCN Evidence Blocks. Accessed February 11, 2020. https://www.nccn.org/evidenceblocks/

36. Cherny NI, Dafni U, Bogaerts J, et al. ESMO-Magnitude of Clinical Benefit Scale version 1.1. *Ann Oncol*. 2017;28(10):2340-2366. doi:10.1093/annonc/mdx310

37. Institute for Clinical and Economic Review. 2020 Value Assessment Framework | Topic. ICER. Published 2020. Accessed August 20, 2020. https://icer-review.org/topic/2020-value-assessment-framework/

38. Willke RJ, Neumann PJ, Garrison LP, Ramsey SD. Review of Recent US Value Frameworks-A Health Economics Approach: An ISPOR Special Task Force Report [6]. *Value Health*. 2018;21(2):155-160. doi:10.1016/j.jval.2017.12.011

39. Perfetto EM, Oehrlein EM, Boutin M, Reid S, Gascho E. Value to Whom? The Patient Voice in the Value Discussion. *Value in Health*. 2017;20(2):286-291. doi:10.1016/j.jval.2016.11.014

40. Havrilesky LJ, Alvarez Secord A, Ehrisman JA, et al. Patient preferences in advanced or recurrent ovarian cancer. *Cancer*. 2014;120(23):3651-3659. doi:10.1002/cncr.28940

41. Avalere Health. Avalere and FasterCures Release Patient-Perspective Value Framework to Incorporate Patient Preferences into Healthcare Treatment Decisions. Published 2017. Accessed March 19, 2021. https://avalere.com/press-releases/avalere-and-fastercures-release-patient-perspective-value-framework-to-incorporate-patient-preferences-into-healthcare-treatment-decisions

42. Innovation and Value. Partnering with Patients: principles and commitments. Innovation and Value Initiative. Published 2020. Accessed October 2, 2022. https://www.thevalueinitiative.org/partnering-with-patients/

43. Hansen PH, Devlin ND. Multi-Criteria Decision Analysis (MCDA) in Healthcare Decision-Making. In: *The Oxford Encyclopedia of Health Economics*. Oxford University Press; 2020. Accessed March 6, 2022. http://www.oxfordreference.com/view/10.1093/acref/9780190902179.001.0001/acref-9780190902179-e-98

44. Diaby V, Campbell K, Goeree R. Multi-criteria decision analysis (MCDA) in health care: A bibliometric analysis. *Operations Research for Health Care*. 2013;2(1):20-24. doi:10.1016/j.orhc.2013.03.001

45. Angelis A, Linch M, Montibeller G, et al. Multiple Criteria Decision Analysis for HTA across four EU Member States: Piloting the Advance Value Framework. *Soc Sci Med*. 2020;246:112595. doi:10.1016/j.socscimed.2019.112595

46. Institute for Quality and Efficiency in Health Care (IQWiG). General Methods Version 6.0. Published 2020. https://www.iqwig.de/en/about-us/methods/methods-paper/

47. 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):105-112. doi:10.1017/S0266462313000639

48. Jehu-Appiah C, Baltussen R, Acquah C, et al. Balancing equity and efficiency in health priorities in Ghana: the use of multicriteria decision analysis. *Value Health*. 2008;11(7):1081-1087. doi:10.1111/j.1524-4733.2008.00392.x

49. 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

50. Marsh KD, Sculpher M, Caro JJ, Tervonen T. The Use of MCDA in HTA: Great Potential, but More Effort Needed. *Value in Health*. 2018;21(4):394-397. doi:10.1016/j.jval.2017.10.001

51. Krahn M, Miller F, Bayoumi A, et al. DEVELOPMENT OF THE ONTARIO DECISION FRAMEWORK: A VALUES BASED FRAMEWORK FOR HEALTH TECHNOLOGY ASSESSMENT. *Int J Technol Assess Health Care*. 2018;34(3):290-299. doi:10.1017/S0266462318000235

52. Nord E, Daniels N, Kamlet M. QALYs: Some Challenges. *Value in Health*. 2009;12:S10-S15. doi:10.1111/j.1524-4733.2009.00516.x

53. Sanford Schwartz J. Chapter 7 - Health Services Research: Translating Discovery and Research Into Practice and Policy. In: Robertson D, Williams GH, eds. *Clinical and Translational Science (Second Edition)*. Academic Press; 2017:111-133. doi:10.1016/B978-0-12-802101-9.00007-7

54. United Nations Population Division. World Population Prospects: 2019 Revision. Published 2020. Accessed September 25, 2020. https://population.un.org/wpp/

55. National Health Commission of the People's Republic of China. 2019 Statistical Bulletin of China's Health Development. Published 2020. Accessed October 17, 2020. http://www.nhc.gov.cn/guihuaxxs/s10748/202006/ebfe31f24cc145b198dd730603ec4442.shtml

56. The World Bank. Healthy China : Deepening Health Reform in China : Building High-Quality and Value-Based Service Delivery (Chinese). World Bank. Accessed September 28, 2020. https://documents.worldbank.org/en/publication/documents-reports/documentdetail

57. General Office of the State Council of the People's Republic of China. The pilot of National Centralized Drug Procurement Policy on Drug Utilization and Drug Expenditures. Published 2019. http://www.gov.cn/zhengce/content/2019-01/17/content_5358604.htm

58. General Office of the State Council of the People's Republic of China. National Healthcare Security Administration and other eight Departments' Report on Expanding the Pilot of The National Centralized Drug Procurement Policy on Drug Utilization and Drug Expenditures. Published 2019. http://www.gov.cn/zhengce/zhengceku/2019-09/30/content_5456439.htm

59. National Healthcare Security Administration (NHSA). Work Plan for the Adjustment of the National Medical Insurance Drug List in 2022. Published 2022. Accessed January 20, 2023. http://www.nhsa.gov.cn/art/2021/6/30/art_62_5386.html

60. Geng J, Chen X, Yu X, Huang J, He D, Dong J. Study on the evidence-based decision-making framework for reimbursement technologies in view of EVIDEM. *Chinese Journal of Health Policy*. 2018;11(04):50-54.

61. Bao Y, Gao B, Meng M, et al. Impact on decision making framework for medicine purchasing in Chinese public hospital decision-making: determining the value of five dipeptidyl peptidase 4 (DPP-4) inhibitors. *BMC Health Serv Res*. 2021;21(1):1-7. doi:10.1186/s12913-021-06827-0

62. EVIDEM. Evidence and Value: Impact on DEcision Making (EVIDEM) 10th Edition. Published 2017. Accessed February 11, 2020. https://www.evidem.org/

63. Jiang M, Fu Q, Xiong J yang, et al. Preferences heterogeneity of health care utilization of community residents in China: a stated preference discrete choice experiment. *BMC Health Serv Res*. 2020;20(1):1-11. doi:10.1186/s12913-020-05134-4

CHAPTER 2. WHAT IS VALUE IN HEALTH AND HEALTH CARE? A SYSTEMATIC LITERATURE REVIEW OF VALUE ASSESSMENT FRAMEWORKS

Status: This manuscript was published in Value in Health on February 1, 2022. It is reproduced here with permission.

Original publication:

Zhang M, Bao Y, Lang Y, Fu S, Kimber M, Levine M, Xie F. What Is Value in Health and Healthcare? A Systematic Literature Review of Value Assessment Frameworks. Value Health. 2022 Feb;25(2):302-317. Doi: 10.1016/j.jval.2021.07.005.

What is value in health and healthcare? A systematic literature review of value assessment frameworks

Highlights

- The goal of value assessment is to promote an efficient and equitable healthcare system. However, there is no global consensus on how to define and measure value in healthcare.
- Fifty-seven value assessment frameworks (VAFs) were included in this review. The attributes of value can be broadly grouped into nine categories, namely, health benefits, affordability, societal impact, burden of disease, quality of evidence, cost-effectiveness, ethics and equity, unmet needs, and innovation. Literature review has been the primary method to define value, while weighting is commonly used to derive value scores.
- There are substantial variations in defining and measuring value. A noticeable weakness of existing VAFs is that patient/public engagement was generally very limited or missing in framework development process. Existing VAFs tend to aggregate multiple value attributes into a single index for decision making.

Abstract

Objectives: This study aimed to investigate how value is defined and measured in existing value assessment frameworks (VAFs) in healthcare.

Methods: We searched PubMed, Embase, the Cochrane Library and Centre for Review and Dissemination from 2008 to 2019. We also performed backward citation chaining of included studies and previously published systematic reviews. Studies reporting the development of a VAF in healthcare were included. For each included framework, we extracted and compared the context, target users, intended use, methods used to identify value attributes, description of the attributes, and attribute scoring approaches.

Results: Of the 8,151 articles screened, 57 VAFs were included. The value attributes included in 55 VAFs were grouped into nine categories: health benefits (n=53, 96%), affordability (n=45, 82%), societal impact (n=42, 76%), the burden of disease (n=36, 65%), quality of evidence (n=32, 58%), cost-effectiveness (n=31, 56%), ethics and equity (n=27, 49%), unmet needs (n=21, 38%), and innovation (n=15, 27%). The remaining two VAFs used broad attributes or user-defined attributes. Literature review was the main approach to identify value attributes in 36 VAFs. Patient or public was engaged through the development of only 11 VAFs. Weighting has been used to score 29 VAFs, of which 19 used the methods of multicriteria decision analysis.

Conclusions: There are substantial variations in defining and measuring value. A noticeable weakness of existing VAFs is that patient/public engagement was generally very

limited or missing in framework development process. Existing VAFs tend to aggregate multiple value attributes into a single index for decision making.

Keywords: healthcare decision making, health technology assessment, value assessment frameworks.

Introduction

The World Health Organization (WHO) estimated that 10% of global Gross Domestic Product (GDP) was spent on health in 2016.[1] Healthcare spending is projected to exceed 20% of GDP in many countries by 2050.[2] Facing rising healthcare demand with limited resources, healthcare systems are shifting from "fee-for-service" reimbursement to value-based models.[3] Health technology assessment (HTA) that provides a multidisciplinary and dynamic assessment on the adoption and diffusion of a new technology (e.g. drug, device, test kit, etc.) has been widely used to inform healthcare decisions and coverage policymaking under the value-based models.[4,5] Considerable efforts have been made to develop frameworks to facilitate the value assessment process in HTA over the past decade.[6–10] Recently, a few institutions and organizations, including the American Society of Clinical Oncology (ASCO), the Memorial Sloan Kettering Cancer Center (MSKCC), the National Comprehensive Cancer Network (NCCN), and the European Society of Medical Oncology (ESMO), as well as HTA bodies such as the Institute for Clinical and Economic Review (ICER) have developed VAF to support value assessment and optimize the decision making process about the use or coverage of new technologies.[11–15]

Despite the growing interest in developing and using VAFs, there is no global consensus on how to define a health technology's value, from whom the value was derived, to how

value is to be used to inform health technology decision making. The value associated with a health technology has conventionally been assessed by comparing cost with health outcomes (e.g. incremental cost per quality adjusted life years (QALYs) gained)[16,17]. However, others have proposed a definition of value beyond cost-effectiveness, adding that equity and ethical considerations are essential attributes for health technology decision making.[10,18] According to a report by the Institute of Medicine (IOM), the meaning of "value" varies substantially.[19] Healthcare providers tend to define value based on the appropriateness and effectiveness of interventions and the evidence supporting their use. In contrast, patients tend to value a health technology based on its ability to increase the accessibility, equity, and quality of healthcare.[19]

VAFs were designed to measure and communicate the value of health technologies to inform healthcare decision making.[10] Examining their development and characteristics such as constituent value attributes (also known as value elements, criteria, or domains), scoring methods, and target users, can provide insight into understanding the impacts and influence of value assessment for health technology decision making. Previous literature reviews have attempted to summarize existing VAFs.[20–23] For example, Seixas et al. identified and classified the value assessment approaches and strategies adopted by VAFs.[20] In an International Society for Pharmacoeconomics and Outcomes Research (ISPOR) special task force report, Willke et al. purposively summarized the limitations and critiques of 4 recent US-oriented VAFs.[22] González-Lorenzo et al. and Morgan et al. specifically focused on the attributes included in frameworks supporting decision making on vaccines and high-cost but effective and desirable technologies.[21,23] Although these

reviews provide some important information regarding the breadth and characteristics of VAFs, their scope was narrow, focusing on, for example, the US only, a specific type of health technology (e.g. vaccines), or only a specific component of the framework (e.g. attributes or attribute aggregation strategies).[21–23] There has been no comprehensive review of existing VAFs with respect to what attributes are included, how attributes are aggregated and the extent to which various stakeholders have been engaged in framework development. The purpose of this systematic review was to investigate how value is defined and measured in healthcare by systematically synthesizing the available literature on VAFs. This review was motivated by a growing interest in expanding the use of cost-effectiveness analysis (CEA) (with or without using the QALY metric) in assessing the value of health technologies. Therefore, its scope focuses on VAFs beyond cost-effectiveness.

Methods

We reported this review according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines.[24]

Search strategy

We conducted a systematic literature search in PubMed, Embase, Cochrane Library and Centre for Review and Dissemination to identify published VAFs. Given that VAF is a relatively recent concept and has evolved rapidly over time, we restricted our search from January 2008 to October 2019.[3,25] Search terms included "value," "value-based," "framework," "decision making," "decision support," "health technology assessment," and "reimbursement". The search strategy was tested against a set of frameworks that were

determined by the research team, a priori, as relevant for inclusion in the present review. The final search was executed by a team member (MZ). Detailed search strategies are presented in Appendix 1 in Supplemental Materials found at https://doi.org/10.1016/j.jval.2021.07.005. The studies included in full-text review and previously published systematic reviews underwent backward citation chaining, which identifies further relevant literature to screen by looking backwards in time and checking the references of these studies.[26] The ICER 2020 framework was released during the review and thus was included to replace the 2019 version.[11,27]

Study selection

We conducted a two-stage screening, which consisted of the initial title and abstract screening and subsequent full-text review. In both stages, two pairs of reviewers (MZ and YB, YL and SF) used predesigned screening forms to review and select eligible publications independently. The full texts of potentially eligible articles after the title and abstract screening were retrieved and reviewed. Any disagreement between reviewers during the screening were discussed and resolved by consulting the senior investigator (FX).

It Is important to note that "framework" is a loosely defined term in the literature and there are noticeable variations in using or not using this term. For this review, articles satisfying the following two criteria were included: 1) the study output is a VAF, which, for the purpose of this review, is defined as a structure that explicitly describes and defines value attributes and their relationship; and 2) the intended use of the framework is to support HTA and inform decision making about the use or coverage of the health technologies that

may include clinical decision making, formulary listing, pricing, coverage, and reimbursement policymaking. Articles were excluded if they met one of the following criteria: 1) developing frameworks for other purposes (e.g., setting general health policy priorities); 2) identifying or measuring specific value attributes (e.g., cost-effectiveness), but not for the development of a framework; 3) evaluating or testing an existing value framework; and 4) protocols, clinical practice guidelines, comments or opinions about VAFs. We restricted our review to VAFs published in English. For VAFs with updates, we only included the most recent version of the VAF. In cases where a framework was presented by more than one publication, we included all relevant publications.

Data extraction and analyses

The data extraction was conducted independently by two pairs of reviewers (MZ and YB, YL and SF), and any discrepancy was resolved through full group discussions. For each VAF included in the review, we extracted target jurisdictions, type of health technologies, target users, intended use, methods used to identify and select value attributes, description of attributes, and attribute scoring methods.

Two reviewers (MZ and YB) examined and compared the original descriptions of attributes in each framework and proposed initial categories. The principle of nonoverlapping was followed through the categorization process by cross-checking the attributes within each category to avoid double counting.[28] After consulting the senior investigator, nine categories are used to summarize the attributes, including 1) health benefits of technology which assess efficacy, effectiveness, safety, or impact on patient-reported outcomes; 2)

quality of evidence which is related to the credibility and certainty of evidence; 3) cost-effectiveness; 4) innovation which assesses the level of novelty and advancement in terms of the mechanism of action or the approach/technique used to improve the properties, usage and performance of health technology; 5) burden of disease which describes mortality, morbidity, and economic impact of the target disease; 6) unmet needs which describe availability or limitations of alternative health technologies; 7) affordability which assesses costs and budget impact associated with the use of health technology; 8) ethics and equity; 9) societal impact which covers socio-cultural, organizational, legal and political implications associated with the use of health technology. The detailed descriptions of these nine categories are provided in Appendix 2 in Supplemental Materials found at https://doi.org/10.1016/j.jval.2021.07.005. We did not assess the quality of VAFs since no tool is available for assessing the quality of this type of research to our knowledge.

Results

Search results

As shown in Figure 1, a total of 8,151 records were identified in the database searches. After removing duplicates and title/abstract screening, 339 were included in the full-text review. Through backward citation chaining and reviewer's suggestion, another 47 articles were added to the full-text screening. Out of the 386 articles, 57 frameworks (described in 62 articles) are included in the review (Figure 1).[11–15,23,29–79] Of these 57 frameworks, 33 were identified via the database searches,[12,23,31–59,77,78] 23 were identified via backward citation chaining,[11,13–15,30,60–76,79] and one published in 2020 included following a

reviewer's suggestion during the peer review.[29] The characteristics of these VAFs are shown in Table 1 and summarized below.

The context and target jurisdiction

There are 48 VAFs that reported their target jurisdictions. Eleven (23%) were developed for global use[12,14,15,23,31,33,40,63,64,67,79] and five (10%) for regional use (one each for Latin America,[56] Asia-Pacific[41] and Europe[34] and two for low/middle-income countries[37,75]). The remaining 32 (67%) VAFs targeted specific countries, with 26 for high-income countries (nine for the US,[11,13,50,51,53,54,66,70,74] eight for Canada,[35,42,48,49,55,58,62,76] eight for European countries[32,44,52,61,65,68,69,73] and one for Israel[43]) and six for middle- or low-income countries (three for countries in Asia,[45,59,71] two for countries in Africa[38,47] and one for Bulgaria[46]).

Target health technology

There are 19 VAFs targeting health technologies in general[11,23,29–32,43,47,54,56,58,59,61,63,72,73,77–79] and one of them focused on highly specialized technologies that are high cost/low volume medication or procedures (e.g. heart transplant).[32] Of the 38 frameworks for a specified type of health technology, 27 are for drugs (10 for cancer drugs,[12–15,34,35,48,62,69,75] six for drugs in general,[33,37–41] four for orphan drugs,[44,46,55,76] four for vaccines,[57,65,67,71] one each for biotechnology,[74] anti-diabetic[36] and off-patent drugs[45]) and 11 for non-drug health technologies including preventative health interventions targeting issues such as substance use and mental disorders (n=4),[51–53,66] health service programs targeting other healthcare services (e.g. prenatal care) (n=2),[42,70] diagnostics or

genetic tests (e.g. blood gas analysis)(n=4)[50,60,64,68] and non-drug health technologies in general (e.g. cardiac pacemaker)(n=1).[49]

Target users, intended use/decision context, and perspectives

Fifty-six VAFs target specific users; among these 36 (64%) are for HTA policymakers supporting reimbursement or coverage policy making about health technologies.[32,33,35,37–39,41–49,51,52,55–59,61,62,64–66,68–71,73,75–78] Among the 17 (30%) frameworks with multiple target users, three are for shared decision making between healthcare professionals and patients,[12,14,36] and 14 for multiple stakeholders to make recommendations or prioritization decisions regarding health technologies.[11,13,15,23,29–31,34,50,53,60,63,67,79] One framework was developed to inform coverage decisions regarding health technology at the hospital level.[74] Among the remaining two frameworks, one is for pharmaceutical companies to guide new drug development[40] and the other for educators or residency programs to teach the trainees about health technology value assessment.[54] Out of 32 frameworks reporting perspectives, 13 (41%) use societal perspective,[33,43,46,47,51,53,61,63–67,71] nine (28%) the health system perspective,[37,38,40,44,48,55,60,69,76] three (9%) both the societal and the health system perspectives,[11,41,79] six (19%) the patient's perspective,[12,14,30,34,36,78] and one (3%) the health system or patient's perspective.[31]

Identification and selection of attributes

Forty-four frameworks reported their methods for the identification and selection of value attributes; eleven (25%) frameworks used literature review,[13,23,39,43,53,57,63–65,73,74] eight (18%) involved stakeholder engagement[11,12,15,30,47,66,70,76] and 25 (57%) included a

combination of literature review and engaging stakeholders.[31–34,36,38,40,42,44–46,48,49,51,52,55,56,58,59,67–69,71,75,79] When stakeholders were involved, the consensus on the selection of attributes was achieved by means of consultation, interview, focus group, and/or survey.

Description of value attributes

The number of attributes included in the VAFs ranges from 3 to 35. Based on the descriptions of these attributes, they are grouped into 9 categories, except for the two frameworks by Lee et al and Lakdawalla and Phelps.[29,50] Lee's framework consists of three value attributes for diagnostics: medical value (impact on treatment decisions), planning value (impact on patients' health, work and life plans) and psychic value (impact on patients' sense of satisfaction)[50] and these three attributes were deemed too broad to be included in our defined categories and thus are listed separately. The augmented CEA framework derives a generalized single index, the Generalized Risk-Adjusted QALY (GRA-QALY), to incorporate risk aversion in quality of life (QoL) and uncertainty in treatment outcomes.[29] The distributions of the attribute categories included in the remaining 55 VAFs are shown in Figure 2. Health benefits of technology (53, 96%), affordability (45, 82%), and societal impact (42, 76%) are the three most frequently included. In contrast, attributes about ethics and equity (27, 49%), unmet needs (21, 38%) and innovation (15, 27%) are least frequently included. The remaining three attributes are the burden of disease (36, 65%), quality of evidence (32, 58%) and cost-effectiveness (31, 56%). Three VAFs cover all nine attribute categories,[48,55,73] and seven VAFs eight categories,[11,23,40,49,51,56,57] and seven VAFs seven categories.[33,42,44,46,63,67,79] Thirty-six

VAFs span three to six categories, two VAFs cover two categories.[15,72] Twenty-four VAFs use QALY to measure health benefits or cost-effectiveness of health technologies,[11,29,38,40,41,47,49,51,52,55,57,59,61,62,64–67,69,70,76–79] 22 did not use QALY as a value attribute,[12–15,30,33,35–37,39,42–44,50,53,54,60,63,68,71,72,74] and the other 11 are VAFs that include cost-effectiveness as a value attribute without specifying whether QALY is used or not.[23,31,32,34,45,46,48,56,58,73,75]

Scoring methods and decision criterion

Out of the 48 VAFs that have explicitly described their attribute scoring methods, 26 (54%) adopt attribute weighting,[12,13,32–39,42,43,45–47,52,55,59,61,67–71,74,75] 13 (27%) deliberation,[11,14,31,40,48,49,54,58,60,62,65,66,76] three (6%) a combination of weighting and deliberation,[44,63,73] and one (2%) grading.[15] Three (6%) frameworks that fundamentally build on CEA combined the attributes by calculating incremental cost per GRA-QALY or incremental cost per unit of financial risk protection benefits and displaying distributions of health gains across different groups of people.[29,77,78] Two frameworks only discussed different scoring methods (e.g. ordered sequence, weighting system and simultaneous consideration) but did not explicitly indicate which one to use.[51,79] Of the 29 VAFs with weighting, 19 are based on multi-criteria decision analysis (MCDA) methods[32–34,36,38,43–47,52,59,63,67,68,71,73–75] which seek to take account of multiple criteria explicitly for complex decisions.[80] Even among those that used the MCDA approach, a wide range of techniques were used to elicit the relative weights of included attributes, including the direct rating techniques (such as ranking, scale rating or point allocation) or analytic hierarchy process (AHP) (n=11),[32,34,36,44,46,63,67,68,73–75] discrete choice experiments (DCE) (n=3),[38,47,52] best

worst scaling (BWS) (n=1),[71] Potentially All Pairwise RanKings of all possible Alternatives (PAPRIKA) (n=1),[43] a combination of the modified Simple Multi-Attribute Rating Technique (SMART) and swing weighting (n=1),[45] and value measurement model in general (n=1).[33] There is one framework that adopts equal weight across attributes.[59] The remaining 10 VAFs used scoring tools (e.g. estimation of net health benefit) (n=7)[12,35,37,42,55,61,69], user-defined weighting (n=1)[13], or did not provide technique details (n=2).[39,70] Among all 57 frameworks, 13 (23%) have specified or recommended a decision criterion for "good value". These include thresholds for CEA, cut-off points to select the top ranking technologies, and a funding line for the prioritization list in accordance with the available budget.[11,13,37–39,41,46,59,67,69,70,77,79]

Patient/public engagement

Patient/public engagement can take place at two stages of the framework development. The first stage is the identification of important and relevant value attributes and the second developing the scoring methods. Only four (7%) VAFs engaged patients/public in the process of attribute identification[30,56,69,79] and six (11%) engaged patients/public in the process of attribute scoring[14,35,36,39,54,63]. Eleven (19%) VAFs have engaged patients/public in both attribute identification and scoring.[11,12,31–34,42,46,55,59,75] Out of these 11 VAFs, eight included patients/public representatives through consultation, survey, or workshop in attribute identification, of which four involved patient/public engagement prior to the conduct of a literature review [32,42,46,59] and another four after completing a literature review.[33,34,55,75] The remaining three VAFs obtained patient or public input after the draft

framework or recommendation was developed.[11,12,31] None of the frameworks reported the characteristics of patients/public on the panel.

Discussion

VAFs represent an important effort with respect to supporting HTA and addressing the issue of allocating limited resources to meet rising healthcare demand. This systematic review summarizes and describes the key characteristics of existing VAFs developed since 2008. Most of these VAFs were developed in high-income countries and focused on drugs. Engagement of patients and the public in VAF development has been limited or missing. These VAFs differ noticeably in the value attributes included and scoring methods used. As a method with the ability of combining multiple value attributes into a single index, MCDA has gained more attention in recent years.

Patients/public are consumers of healthcare services and will ultimately be affected by decisions made using the VAFs.[25] Although some progress has been made in involving patients/public in the VAF development, it remains inadequate.[11,25,81] This could be due to the interplay between the increased recognition of the importance of patient/public engagement and the challenge of achieving patient-centered decision making in practice.[82,83] We found that the framework development method may affect the patient/public engagement. For example, VAFs using weighting are more likely to engage patients/public in their development compared to those using deliberation, which could be cognitively demanding or requiring knowledge on healthcare.[84,85]

As noted in previous studies[86,87], value assessment has expanded from the traditional territory of health benefit to affordability, societal impact, and disease burden (all included by at least two-thirds of the VAFs reviewed in this study). This reflects the increasing recognition of the multi-dimensional value concept among decision-makers.[88,89] Notably, existing VAFs vary in the way the attributes are described and organized. For example, most frameworks treat "health benefit," "cost," and "cost-effectiveness" as separate attributes, but two use "cost-effectiveness" as one attribute to cover all three.[38,47] When updating their framework, Danko et al. recommend using "cost-effectiveness" with caveats rather than including it as an attribute as in the original framework. This modification was based on the consideration that there are different institutional environments when using different sources of evidence (i.e. localized economic evaluations versus referencing previous evaluations in other countries) to incorporate "cost-effectiveness" in decision making processes.[37,90] This heterogeneity in value attributes among existing VAFs highlights the contextual complexity of VAFs and the importance of using appropriate and the best available evidence to assess a health technology's performance in each attribute. Furthermore, a wide range of methods was used to score value attributes. The methodological variation reflects the diversity in the intended uses (e.g. health technology ranking), stakeholders' needs (e.g. patients'/public's needs), and feasibility (e.g. cognitive burden and resource requirement of DCE) of these frameworks. Guiding principles for VAF development, including transparency, dynamic value assessment, and stakeholder engagement could be useful and needed to improve the quality and implementation of VAFs.[25,91]

Previous health economics studies have developed approaches to measuring the value of health technologies to aid healthcare decision makings. These approaches include but are not limited to opportunity cost, willingness to pay (WTP), experienced utility, and QALY.[92–95] The opportunity cost can be measured in different ways under welfarism versus extra-welfarism.[94] The WTP approach derives from welfarism where the aim is to maximize "social welfare" defined as the aggregate individual well-being and a global improvement occurs for a new health technology when the benefits compensate and outweigh the costs.[92,93] Under this compensation rule, the "social welfare" can be estimated in terms of total net WTP as in cost-benefit analysis (CBA). However, the use of WTP has been limited due to its susceptibility to individuals' ability to pay and the challenges of valuing health benefits in monetary terms.[92] Alternatively, utility measurement provides a way to measure individuals' hedonic welfare or individuals' happiness or satisfaction. This approach's problem is that the judgements depend only on individuals' utility without considering non-utility information.[92] Extra-welfarism argues that there are social values beyond individual utility maximization that can lead to the objective of maximizing the health of the community given a fixed budget.[92] For example, the implementation of this in the UK has been based on QALY as the metric of health gain and using a constant, opportunity-cost-based threshold to determine cost-effectiveness.[96] However, concerns related to the sole goal of health maximization under extra-welfarism and the failure of QALY to capture patient-centered outcomes other than health have driven researchers to seek other options of value measurements to achieve the multiple goals in healthcare decision making.[92–94]

Over the past decades, there is an increasing acceptance that value in healthcare is multi-dimensional. One of the significant moves was the Impact Inventory introduced by the Second Panel on Cost-Effectiveness in Health and Medicine.[16] The impact inventory is a checklist to encourage the identification and enumeration of consequences related to the choice of a health technology in both healthcare and non-healthcare sectors. The impact inventory, with modifications, could potentially be used as a framework to support value assessment in the future.[97] Our review found that existing VAFs tend to aggregate multiple value attributes into a single index for decision making. Among the various methods used by VAFs, MCDA approaches have been adopted by a number of VAFs in recent years. MCDA improves transparency and reproducibility in the value assessments of health technologies.[33,98,99] However, there still remains some challenges when using MCDA. First, it is challenging to choose from many existing methods to elicit or derive attribute weights for attributes such as societal impact, ethics and equity.[49,82] The difficulty of measuring these attributes and the scarcity of evidence related to them compared to those that are more easily measurable (e.g. effectiveness and cost) might result in inaccurate weights for the framework.[49] Another concern of applying the MCDA approach is its "rigidity" in the process of coverage decision making, which would be introduced by its structured and fixed mechanism.[49] This rigidity might limit stakeholder engagement in decision making, generate an inappropriate model algorithm, and eventually lead to suboptimal decisions. Integration of MCDA with stakeholder consultation and discussion could be a promising alternative to address this concern.[100]

Strength and limitations

Compared with previous reviews, the strength of our study is that we set no limit on the target jurisdiction, heath technology type or intended use of VAFs. It presents a full picture of existing VAFs. The findings of our study could help inform further development and application of VAFs.

This review has a few limitations. First, it is worth mentioning that the concept of VAF is relatively new and not well defined. Different terms have been used and some organizations published the framework on their websites only.[11,13,62,69,70] This increases the difficulty in identifying VAFs through bibliographic databases. In our review, approximately 40% of the VAFs were identified via backward citation chaining of references instead of database search strategies. Second, we chose to group value attributes into nine categories to meaningfully summarize and present the large amount of information contained in existing VAFs. The categorization was based on our understanding of content overlap among originally reported value attributes and thus should not be interpreted as an attempt to standardize the value attribute definition. The original VAFs should be consulted as the sole source for value attribute description.

Implications for policy and research

The majority of the VAFs included in our review are intended for assessing the value of health technologies and subsequently supporting policies on the use and coverage of these technologies in healthcare systems. They could profoundly impact access to innovative technologies and the efficient use of limited resources while meeting rising healthcare demands. We focused on examining and comparing methodological aspects among

existing VAFs. By highlighting similarities and differences in the framework development, this review can help further our understanding of value, shed light on the areas for future improvement in the framework development, and consequently enhance their role in informing policy making.

Conclusions

There are substantial variations in defining and measuring value. The heterogeneity in methodology could be due to context specific factors including but not limited to the characteristics of target technology or requirement for local policy making. Regardless of the methods used, most existing VAFs can be applied to assessing the value for a wide range of technologies. A weakness seen across existing VAFs is that patient/public engagement was very limited or missing in framework development process. Aggregating the value attributes to generate a single index is gaining increasing interest among the VAF developers.

References

1 World Health Organization. WHO | Public Spending on Health: A Closer Look at Global Trends. WHO. 2018.http://www.who.int/health_financing/documents/health-expenditure-report-2018/en/ (accessed 8 Feb2020).

2 Dixon-Fyle S, Kowallik T. Engaging consumers to manage healthcaredemand | McKinsey. 2010.https://www.mckinsey.com/industries/healthcare-systems-and-services/our-insights/engaging-consumers-to-manage-health-care-demand (accessed 8 Feb2020).

3 Centers for Medicare and Medicaid Services. CMS' Value-Based Programs | CMS. https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/Value-Based-Programs/Value-Based-Programs (accessed 3 Sep2020).

4 National Library of Medicine. HTA 101: II. FUNDAMENTAL CONCEPTS. https://www.nlm.nih.gov/nichsr/hta101/ta10104.html#Heading1 (accessed 29 Jun2020).

5 Health Technology assessment international. About HTAi | Health Technology Assessment International (HTAi). HTAi. https://htai.org/about-htai/ (accessed 19 Nov2020).

6 Evidence-Based Medicine Working Group. Evidence-based medicine. A new approach to teaching the practice of medicine. *JAMA* 1992; **268**: 2420–2425.

7 Dobrow MJ, Goel V, Upshur REG. Evidence-based health policy: context and utilisation. *Soc Sci Med* 2004; **58**: 207–217.

8 Hoffmann C, Graf von der Schulenburg JM. The influence of economic evaluation studies on decision making. A European survey. The EUROMET group. *Health Policy* 2000; **52**: 179–192.

9 Trueman P, Drummond M, Hutton J. Developing guidance for budget impact analysis. *Pharmacoeconomics* 2001; **19**: 609–621.

10 Neumann PJ, Willke RJ, Garrison LP. A Health Economics Approach to US Value Assessment Frameworks-Introduction: An ISPOR Special Task Force Report [1]. *Value Health* 2018; **21**: 119–123.

11 Institute for Clincial and Economic Review. 2020 Value Assessment Framework | Topic. ICER. 2020.https://icer-review.org/topic/2020-value-assessment-framework/ (accessed 20 Aug2020).

12 Schnipper LE, Davidson NE, Wollins DS, Blayney DW, Dicker AP, Ganz PA *et al.* Updating the American Society of Clinical Oncology Value Framework: Revisions and Reflections in Response to Comments Received. *J Clin Oncol* 2016; **34**: 2925–2934.

13 Memorial Sloan Kettering Cancer Center. DrugAbacus. Drug Pricing Lab. https://drugpricinglab.org/tools/drug-abacus/ (accessed 11 Feb2020).

14 National Comprehensive Cancer Network. NCCN Clinical Practice Guidelines in Oncology (NCCN Guidelines) with NCCN Evidence Blocks. https://www.nccn.org/evidenceblocks/ (accessed 11 Feb2020).

15 Cherny NI, Dafni U, Bogaerts J, Latino NJ, Pentheroudakis G, Douillard J-Y *et al.* ESMO-Magnitude of Clinical Benefit Scale version 1.1. *Ann Oncol* 2017; **28**: 2340–2366.

16 Neumann PJ, Sanders GD, Basu A, Brock DW, Feeny D, Krahn M. *Cost-effectiveness in Health and Medicine*. Oxford University Press: New York, 2016.

17 Drummond M. *Methods for the economic evaluation of healthcareprogrammes*. Fourth edition. Oxford University Press: Oxford, United Kingdom ; New York, NY, USA, 2015.

18 Boztepe S. User Value: Competing Theories and Models. *International Journal of Design* 2007; **1**: 55–63.

19 Institute of Medicine (US) Roundtable on Value & Science-Driven Health Care. *Value in Health Care: Accounting for Cost, Quality, Safety, Outcomes, and Innovation*. National Academies Press (US): Washington (DC), 2010http://www.ncbi.nlm.nih.gov/books/NBK50929/ (accessed 20 May2020).

20 Seixas BV, Dionne F, Conte T, Mitton C. Assessing value in health care: using an interpretive classification system to understand existing practices based on a systematic review. *BMC Health Serv Res* 2019; **19**: 560.

21 González-Lorenzo M, Piatti A, Coppola L, Gramegna M, Demicheli V, Melegaro A *et al.* Conceptual frameworks and key dimensions to support coverage decisions for vaccines. *Vaccine* 2015; **33**: 1206–1217.

22 Willke RJ, Neumann PJ, Garrison LP, Ramsey SD. Review of Recent US Value Frameworks-A Health Economics Approach: An ISPOR Special Task Force Report [6]. *Value Health* 2018; **21**: 155–160.

23 Morgan RL, Kelley L, Guyatt GH, Johnson A, Lavis JN. Decision-making frameworks and considerations for informing coverage decisions for healthcare interventions: a critical interpretive synthesis. *J Clin Epidemiol* 2018; **94**: 143–150.

24 Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Ann Intern Med* 2009; **151**: 264–269, W64.

25 Dubois RW, Westrich K. As Value Assessment Frameworks Evolve, Are They Finally Ready for Prime Time? *Value Health* 2019; **22**: 977–980.

26 Berry F, Choi C. UNIVERSITY OF SYDNEY LIBRARY: CITATION CHAINING: HOW TO FIND MORE INFORMATION FROM ONE SOURCE. 2012.

27 Institute for Clincial and Economic Review. Final Value Assessment Framework: Updates for 2017-2019. ICER. 2017.https://icer-review.org/material/final-vaf-2017-2019/ (accessed 11 Feb2020).

28 Marsh K, IJzerman M, Thokala P, Baltussen R, Boysen M, Kaló Z *et al.* Multiple Criteria Decision Analysis for HealthcareDecision Making--Emerging Good Practices: Report 2 of the ISPOR MCDA Emerging Good Practices Task Force. *Value Health* 2016; **19**: 125–137.

29 Lakdawalla DN, Phelps CE. Health technology assessment with risk aversion in health. *J Health Econ* 2020; **72**: 102346.

30 Avalere Health. Avalere and FasterCures Release Patient-Perspective Value Framework to Incorporate Patient Preferences into Healthcare Treatment Decisions. 2017.https://avalere.com/press-releases/avalere-and-fastercures-release-patient-perspective-value-framework-to-incorporate-patient-preferences-into-healthcare-treatment-decisions (accessed 19 Mar2021).

31 Alonso-Coello P, Schünemann HJ, Moberg J, Brignardello-Petersen R, Akl EA, Davoli M *et al.* GRADE Evidence to Decision (EtD) frameworks: a systematic and transparent approach to making well informed healthcare choices. 1: Introduction. *BMJ* 2016; **353**: i2016.

32 Anderson P, Webb P, Groves S. Prioritisation of specialist healthcareservices; not NICE, not easy but it can be done. *Health Policy* 2017; **121**: 978–985.

33 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.

34 Badia X, Aguarón A, Fernández A, Gimón A, Nafria B, Gaspar B *et al.* Patient involvement in reflective multicriteria decision analysis to assist decision making in oncology. *Int J Technol Assess Healthcare*2019; **35**: 56–63.

35 Browman GP, Manns B, Hagen N, Chambers CR, Simon A, Sinclair S. 6-STEPPPs: A Modular Tool to Facilitate Clinician Participation in Fair Decisions for Funding New Cancer Drugs. *J Oncol Pract* 2008; **4**: 2–7.

36 Choi SE, Berkowitz SA, Yudkin JS, Naci H, Basu S. Personalizing Second-Line Type 2 Diabetes Treatment Selection: Combining Network Meta-analysis, Individualized Risk, and Patient Preferences for Unified Decision Support. *Med Decis Making* 2019; **39**: 239–252.

37 Dankó D, Molnár MP. Balanced assessment systems revisited. *J Mark Access Health Policy* 2017; **5**: 1355190.

38 Diaby V, Lachaine J. An application of a proposed framework for formulary listing in low-income countries: the case of Côte d'Ivoire. *Appl Health Econ Health Policy* 2011; **9**: 389–402.

39 Doyle JJ, Hawryluk E, Niemira J, Wood B. Evidence-Based Valuation in Oncology:: Lessons Learned from a Case Study. *Ther Innov Regul Sci* 2019; **53**: 403–411.

40 Dunlop WCN, Mullins CD, Pirk O, Goeree R, Postma MJ, Enstone A *et al.* BEACON: A Summary Framework to Overcome Potential Reimbursement Hurdles. *Pharmacoeconomics* 2016; **34**: 1051–1065.

41 Finkelstein EA, Krishnan A, Doble B. Beyond cost-effectiveness: A five-step framework for appraising the value of health technologies in Asia-Pacific. *Int J Health Plann Manage* 2019; **35**: 397–408.

42 Gibson J, Mitton C, DuBois-Wing G. Priority setting in Ontario's LHINs: ethics and economics in action. *Healthc Q* 2011; **14**: 35–43.

43 Golan O, Hansen P. Which health technologies should be funded? A prioritization framework based explicitly on value for money. *Isr J Health Policy Res* 2012; **1**: 44.

44 Guarga L, Badia X, Obach M, Fontanet M, Prat A, Vallano A *et al.* Implementing reflective multicriteria decision analysis (MCDA) to assess orphan drugs value in the Catalan Health Service (CatSalut). *Orphanet J Rare Dis* 2019; **14**: 157.

45 Inotai A, Brixner D, Maniadakis N, Dwiprahasto I, Kristin E, Prabowo A *et al.* Development of multi-criteria decision analysis (MCDA) framework for off-patent pharmaceuticals - an application on improving tender decision making in Indonesia. *BMC Health Serv Res* 2018; **18**: 1003.

46 Iskrov G, Miteva-Katrandzhieva T, Stefanov R. Multi-Criteria Decision Analysis for Assessment and Appraisal of Orphan Drugs. *Front Public Health* 2016; **4**: 214.

47 Jehu-Appiah C, Baltussen R, Acquah C, Aikins M, d'Almeida SA, Bosu WK *et al.* Balancing equity and efficiency in health priorities in Ghana: the use of multicriteria decision analysis. *Value Health* 2008; **11**: 1081–1087.

48 Keech J, Beca J, Eisen A, Kennedy E, Kim J, Kouroukis CT *et al.* Impact of a novel prioritization framework on clinician-led oncology drug submissions. *Curr Oncol* 2019; **26**: e155–e161.

49 Krahn M, Miller F, Bayoumi A, Brooker A-S, Wagner F, Winsor S *et al.* DEVELOPMENT OF THE ONTARIO DECISION FRAMEWORK: A VALUES BASED FRAMEWORK FOR HEALTH TECHNOLOGY ASSESSMENT. *Int J Technol Assess Healthcare*2018; **34**: 290–299.

50 Lee DW, Neumann PJ, Rizzo JA. Understanding the medical and nonmedical value of diagnostic testing. *Value Health* 2010; **13**: 310–314.

51 Maciosek MV, Coffield AB, Edwards NM, Flottemesch TJ, Solberg LI. Prioritizing clinical preventive services: a review and framework with implications for community preventive services. *Annu Rev Public Health* 2009; **30**: 341–355.

52 Marsh K, Dolan P, Kempster J, Lugon M. Prioritizing investments in public health: a multi-criteria decision analysis. *J Public Health (Oxf)* 2013; **35**: 460–466.

53 Miller G, Roehrig C, Russo P. A Framework for Assessing the Value of Investments in Nonclinical Prevention. *Prev Chronic Dis* 2015; **12**: E216.

54 Patel MS, Davis MM, Lypson ML. The VALUE Framework: training residents to provide value-based care for their patients. *J Gen Intern Med* 2012; **27**: 1210–1214.

55 Paulden M, Stafinski T, Menon D, McCabe C. Value-based reimbursement decisions for orphan drugs: a scoping review and decision framework. *Pharmacoeconomics* 2015; **33**: 255–269.

56 Pichon-Riviere A, Garcia-Marti S, Oortwijn W, Augustovski F, Sampietro-Colom L. Defining the Value of Health Technologies in Latin America: Developments in Value Frameworks to Inform the Allocation of Healthcare Resources. *Int J Technol Assess Healthcare*2019; **35**: 64–68.

57 Piso B, Wild C. Decision support in vaccination policies. *Vaccine* 2009; **27**: 5923–5928.

58 Stafinski T, Menon D, McCabe C, Philippon DJ. To fund or not to fund: development of a decision-making framework for the coverage of new health technologies. *Pharmacoeconomics* 2011; **29**: 771–780.

59 Youngkong S, Baltussen R, Tantivess S, Mohara A, Teerawattananon Y. Multicriteria decision analysis for including health interventions in the universal health coverage benefit package in Thailand. *Value Health* 2012; **15**: 961–970.

60 Anonychuk A, Beastall G, Shorter S, Kloss-Wolf R, Neumann P. A Framework for Assessing the Value of Laboratory Diagnostics. *Healthc Manage Forum* 2012; **25**: S4–S11.

61 Broqvist M, Branting Elgstrand M, Carlsson P, Eklund K, Jakobsson A. *National Model for Transparent Prioritisation in Swedish Health Care : Reviderad version*. Linköping University Electronic Press, 2011http://urn.kb.se/resolve?urn=urn:nbn:se:liu:diva-111767 (accessed 29 Jun2020).

62 CADTH pan-Canadian Oncology Drug Review Committee. The pCODR Expert Review Committee (pERC). CADTH.ca. 2016.https://www.cadth.ca/collaboration-and-outreach/advisory-bodies/pcodr-expert-review-committee-perc (accessed 11 Feb2020).

63 EVIDEM. Evidence and Value: Impact on DEcision Making (EVIDEM) 1^{0t}h Edition. 2017.https://www.evidem.org/ (accessed 11 Feb2020).

64 Garrison L, Mestre-Ferrandiz J, Zamora B. The Value of Knowing and Knowing the Value: Improving the Health Technology Assessment of Complementary Diagnostics. 2016.https://www.ohe.org/news/value-knowing-and-knowing-value-improving-health-technology-assessment-complementary (accessed 11 Feb2020).

65 Houweling H, Verweij M, Ruitenberg EJ, National Immunisation Programme Review Committee of the Health Council of the Netherlands. Criteria for inclusion of vaccinations in public programmes. *Vaccine* 2010; **28**: 2924–2931.

66 Institute of Medicine. *An Integrated Framework for Assessing the Value of Community-Based Prevention*. 2012 doi:10.17226/13487.

67 Committee on Identifying and Prioritizing New Preventive Vaccines for Development, Phase III, Board on Population Health and Public Health Practice, Board on Global Health, Institute of Medicine, National Academy of Engineering. *Ranking Vaccines: Applications of a Prioritization Software Tool: Phase III: Use Case Studies and Data Framework*. National Academies Press (US): Washington (DC), 2015http://www.ncbi.nlm.nih.gov/books/NBK284622/ (accessed 11 Feb2020).

68 Kroese M, Burton H, Whittaker J, Lakshman R, Alberg C. A framework for the prioritization of investment in the provision of genetic tests. *Public Health Genomics* 2010; **13**: 538–543.

69 NHS England Chemotherapy Clinical Reference Group for NHS. Standard Operating Procedures: The Cancer Drugs Fund (CDF) Guidance to support operation of the CDF in 2015-16. 2015.https://www.england.nhs.uk/cancer/cdf/ (accessed 11 Feb2020).

70 Oregon Health Services Commission. Prioritization of health services: A Report to the Governor and the 7^{5t}h Oregon Legislative Assembly. 2009.

71 Pooripussarakul S, Riewpaiboon A, Bishai D, Muangchana C, Tantivess S. What criteria do decision makers in Thailand use to set priorities for vaccine introduction? *BMC Public Health* 2016; **16**: 684.

72 Porter ME. What is value in health care? *N Engl J Med* 2010; **363**: 2477–2481.

73 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 Healthcare*2014; **30**: 105–112.

74 Seigfried RJ, Corbo T, Saltzberg MT, Reitz J, Bennett DA. Deciding which drugs get onto the formulary: a value-based approach. *Value Health* 2013; **16**: 901–906.

75 Venhorst K, Zelle SG, Tromp N, Lauer JA. Multi-criteria decision analysis of breast cancer control in low- and middle- income countries: development of a rating tool for policy makers. *Cost Eff Resour Alloc* 2014; **12**: 13.

76 Winquist E, Bell CM, Clarke JTR, Evans G, Martin J, Sabharwal M *et al.* An evaluation framework for funding drugs for rare diseases. *Value Health* 2012; **15**: 982–986.

77 Asaria M, Griffin S, Cookson R. Distributional Cost-Effectiveness Analysis: A Tutorial. *Med Decis Making* 2016; **36**: 8–19.

78 Verguet S, Kim JJ, Jamison DT. Extended Cost-Effectiveness Analysis for Health Policy Assessment: A Tutorial. *Pharmacoeconomics* 2016; **34**: 913–923.

79 Garrison LP, Neumann PJ, Willke RJ, Basu A, Danzon PM, Doshi JA *et al.* A Health Economics Approach to US Value Assessment Frameworks-Summary and Recommendations of the ISPOR Special Task Force Report [7]. *Value Health* 2018; **21**: 161–165.

80 Thokala P, Devlin N, Marsh K, Baltussen R, Boysen M, Kalo Z *et al.* Multiple Criteria Decision Analysis for HealthcareDecision Making--An Introduction: Report 1 of the ISPOR MCDA Emerging Good Practices Task Force. *Value Health* 2016; **19**: 1–13.

81 Innovation and value initiative. Partnering with Patients. Innovation and Value Initiative. https://www.thevalueinitiative.org/partnering-with-patients/ (accessed 28 Oct2020).

82 Coyle D, Durand-Zaleski I, Farrington J, Garrison L, Graf von der Schulenburg J-M, Greiner W *et al.* HTA methodology and value frameworks for evaluation and policy making for cell and gene therapies. *Eur J Health Econ* 2020. doi:10.1007/s10198-020-01212-w.

83 Vahdat S, Hamzehgardeshi L, Hessam S, Hamzehgardeshi Z. Patient Involvement in HealthcareDecision Making: A Review. *Iran Red Crescent Med J* 2014; **16**. doi:10.5812/ircmj.12454.

84 Mccarron T, Moffat K, Wilkinson G, Zelinsky S, Boyd J, White D *et al.* Understanding patient engagement in health system decision-making: A co-designed scoping review. *Systematic Reviews* 2019; **8**. doi:10.1186/s13643-019-0994-8.

85 Perfetto EM, Oehrlein EM, Boutin M, Reid S, Gascho E. Value to Whom? The Patient Voice in the Value Discussion. *Value Health* 2017; **20**: 286–291.

86 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**: 126–135.

87 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 decisionmaking. *Cost Eff Resour Alloc* 2012; **10**: 9.

88 Reed SD, Dubois RW, Johnson FR, Caro JJ, Phelps CE. Novel Approaches to Value Assessment Beyond the Cost-Effectiveness Framework. *Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research* 2019; **22**: S18–S23.

89 Lakdawalla DN, Doshi JA, Garrison LP, Phelps CE, Basu A, Danzon PM. Defining Elements of Value in Health Care-A Health Economics Approach: An ISPOR Special Task Force Report [3]. *Value Health* 2018; **21**: 131–139.

90 Dankó D. Health technology assessment in middle-income countries: recommendations for a balanced assessment system. *J Mark Access Health Policy* 2014; **2**. doi:10.3402/jmahp.v2.23181.

91 Neumann PJ. Yes, Improve the US Value Frameworks, But Recognize They Are Already in Prime Time. *Value in Health* 2019; **22**: 975–976.

92 Coast J, Smith RD, Lorgelly P. Welfarism, extra-welfarism and capability: The spread of ideas in health economics. *Social Science & Medicine* 2008; **67**: 1190–1198.

93 Brouwer WBF, Culyer AJ, van Exel NJA, Rutten FFH. Welfarism vs. extra-welfarism. *J Health Econ* 2008; **27**: 325–338.

94 Booth N. On value frameworks and opportunity costs in health technology assessment. *Int J Technol Assess Healthcare*2019; **35**: 367–372.

95 Birch S, Donaldson C. Valuing the benefits and costs of healthcareprogrammes: where's the 'extra' in extra-welfarism? *Soc Sci Med* 2003; **56**: 1121–1133.

96 McCabe C, Claxton K, Culyer AJ. The NICE cost-effectiveness threshold: what it is and what that means. *Pharmacoeconomics* 2008; **26**: 733–744.

97 McQueen RB, Slejko JF. Toward Modified Impact Inventory Tables to Facilitate Patient-Centered Value Assessment. *Pharmacoeconomics* 2021; **39**: 379–382.

98 Roldán ÚB, Badia X, Marcos-Rodríguez JA, de la Cruz-Merino L, Gómez-González J, Melcón-de Dios A *et al.* MULTI-CRITERIA DECISION ANALYSIS AS A DECISION-SUPPORT TOOL FOR DRUG EVALUATION: A PILOT STUDY IN A PHARMACY AND THERAPEUTICS COMMITTEE SETTING. *Int J Technol Assess Healthcare*2018; **34**: 519–526.

99 Phelps CE, Lakdawalla DN, Basu A, Drummond MF, Towse A, Danzon PM. Approaches to Aggregation and Decision Making-A Health Economics Approach: An ISPOR Special Task Force Report [5]. *Value in Health* 2018; **21**: 146–154.

100 Angelis A, Kanavos P, Phillips LD. ICER Value Framework 2020 Update: Recommendations on the Aggregation of Benefits and Contextual Considerations. *Value Health* 2020; **23**: 1040–1048.

101 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 Services Research* 2008; **8**: 270.

102 Treweek S, Oxman AD, Alderson P, Bossuyt PM, Brandt L, Brożek J *et al.* Developing and evaluating communication strategies to support informed decisions and practice based on evidence (DECIDE): protocol and preliminary results. *Implement Sci* 2013; **8**: 6.

103 Iskrov G, Stefanov R. Criteria for Drug Reimbursement Decision-Making: An Emerging Public Health Challenge in Bulgaria. *Balkan Med J* 2016; **33**: 27–35.

104 Committee on Identifying and Prioritizing New Preventive Vaccines for Development, Board on Population Health and Public Health Practice, Board on Global Health, Institute of Medicine. *Ranking Vaccines: A Prioritization Framework: Phase I: Demonstration of Concept and a Software Blueprint*. National Academies Press (US): Washington (DC), 2012http://www.ncbi.nlm.nih.gov/books/NBK201608/ (accessed 29 Oct2020).

List of figures and tables

Notes: The 9 colored tracks and 55 sectors (labelled using author name year) represent the 9 attribute categories and 55 VAFs, respectively. The outer a track, the more frequent the corresponding attribute category is included in existing VAFs. The intersection of a sector and a track indicates the presence (colored) or absence (blank) of the corresponding attribute category in that VAF.

Abbreviations: VAF: value assessment framework.

Figure 1 Flow diagram of literature screening

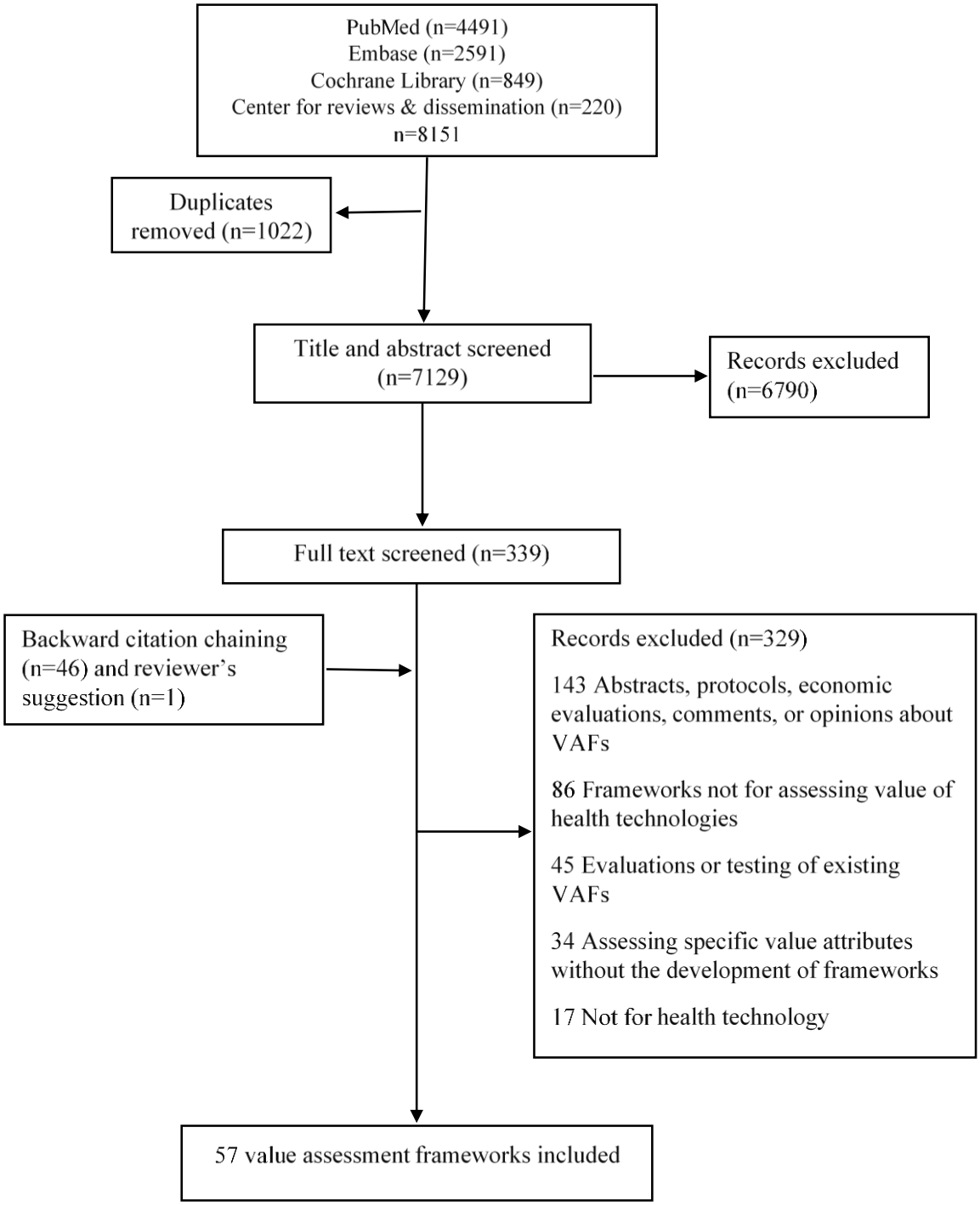


Figure 2 Distribution of attribute categories in VAFs

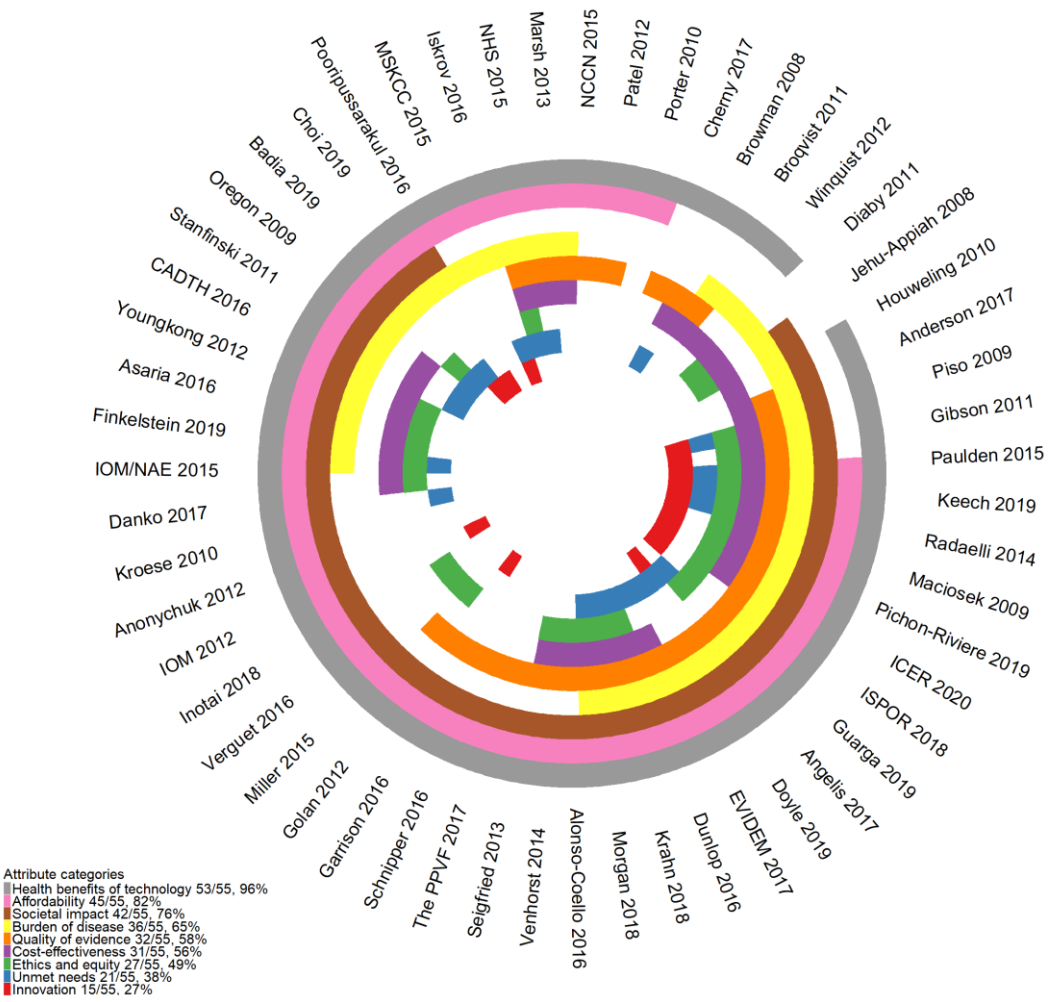


Table 1. Characteristics of included value assessment frameworks

Author year (Framework name)	Target jurisdiction	Health technology	Target users	Intended use /decision context	Perspective	Identification and selection of attributes*	Number of attributes	Use of QALY	Scoring methods*	Decision criterion
ICER value assessment framework[11] 2020	US	Health technologies	HTA policymakers Healthcare professionals Industry	Deliberation support on medical policies related to health technologies at the population level	Both the health system perspective and societal perspective	Stakeholder panel consultation (including patients and public)	5 attributes	Yes	Stakeholder panel judgements (including patients and the public)	$100,000-$150,000 per QALY and evLYG for CEA. Annual budget threshold of $819 million for an individual intervention.
Lakdawalla et al.[29] 2020 (Augmented CEA)	NR	Health technologies	HTA policymakers Researchers Industry	Healthcare resource allocation decisions	NR	NR	User-defined	Yes**	Generalized ICER using the GRA-QALY	NR
Badia et al.[34] 2019	Europe	Cancer drugs	HTA policymakers Healthcare professionals Industry	Cancer drug development and funding decisions	Patient's perspective	Literature review; patient workshop	8 attributes	NR	Weighting by a non-hierarchical simple 1-5 scale from patients	NR
Choi et al.[36] 2019	NR	Anti-diabetic drugs	Healthcare professionals Patients	Shared decision making about the selection of second-line glycemic therapy for type 2 diabetes patients	Patient's perspective	Literature review; national patient survey data	12 attributes	No	Weighting by a 0 to 1 scale using patient survey data	NR
Doyle et al.[39] 2019 (EBV framework)	NR	Drugs	HTA policymakers	Estimation of the value-based pricing range for the new drugs	NR	Literature review	4 attributes	No	Weighting by quantitative studies performed within stakeholder panel (including patients)	Payer evidence thresholds based on the value offered by selected products at various prices
Finkelstein et al.[41] 2019	The Asia-Pacific area	Drugs	HTA policymakers	Drug coverage and reimbursement decisions	Both the health system perspective and societal perspective	NR	5 attributes	Yes	NR	Implicit willingness-to-pay threshold is recommended for CEA

Guarga et al.[44] 2019	Spain	Orphan drugs	HTA policymakers	Orphan drug appraisal and reimbursement decisions	The health system perspective	Literature review; healthcare professional consensus	10 quantitative attributes in the Core Model and 4 qualitative attributes in the Contextual Tool	No	Weighting by a non-hierarchical simple 1-5 scale for quantitative attributes and judgements for qualitative attributes among healthcare professionals	NR
Keech et al. [48] 2019 (CCO Prioritization Framework)	Canada	Cancer drugs	HTA policymakers	Cancer drug coverage and reimbursement decisions	The health system perspective	Literature review; healthcare professional consensus	7 attributes	NR	Policymaker and healthcare professional judgements	NR
Pichon-Riviere et al. [56] 2019	Latin America	Health technologies	HTA policymakers	Healthcare resource allocation decisions	NR	Literature review; stakeholder panel consensus (including patients)	15 attributes	NR	NR	NR
Inotai et al. [45] 2018	Indonesia	Off-patent drugs	HTA policymakers	National off-patent drug procurement	NR	Literature review; stakeholder panel consensus (without patients/public)	7 attributes	NR	The modified Simple Multi-Attribute Rating Technique (SMART) and swing weighting from stakeholder panel (without patients/public)	NR
Krahn et al. [49] 2018 (OHTAC framework)	Canada	Non-drug health technologies	HTA policymakers	Non-drug health technology coverage decisions	NR	Literature review; expert consensus;	16 attributes	Yes	Stakeholder panel judgements (without patients/public)	NR
ISPOR[79] 2018	Global	Health technologies	HTA policymakers Healthcare professionals Industry Researchers	Health technologies value assessment	Both the health system perspective and societal perspective	Literature review, stakeholder panel consensus (including patients and public)	12 attributes	Yes	Different criteria combination approaches were discussed[99]	Value thresholds are recommended for coverage and reimbursement decisions
Morgan et al. [23] 2018	Global	Health technologies	HTA policymakers Guideline	Health technology coverage decisions	NR	Literature review	7 attributes	NR	NR	NR

			developers HTA producers							
Anderson et al. [32] 2017 (WHSSC framework for HST prioritization)	Welsh	Highly specialized medical technologies	HTA policymakers	HST funding decisions	NR	Literature review; Stakeholder panel consensus (including patients and public)	5 attributes	NR	Weighting by the Portsmouth Scorecard from stakeholder panel (including patients and public)	NR
Angelis and Kanavos[33] 2017 (The AVF framework)	Global	Drugs	HTA policymakers	Drug coverage and reimbursement decisions	Societal perspective	Literature review; Stakeholder panel consensus (including patients)	11 attributes	No	Weighting by value measurement methods from stakeholder panel (including patients)	NR
Cherny et al. [15] 2017 The ESMO-MCBS	Global	Cancer treatments	HTA policymakers Guideline developers HTA producers	Clinical benefit assessment of cancer treatments	NR	Expert consensus	5 attributes	No	Grading of clinical benefit by experts	NR
Dankó & Molnár[37] 2017 (BAS framework)	Middle-income countries	Drugs	HTA policymakers	Drug coverage and reimbursement decisions	The health system perspective	NR	6 attributes	No	Use of a scoring system and cut-off values for evaluation	Selecting top X drugs using cut-off scores, or selecting top X medicines to be reimbursed until the available budget is exhausted
EVIDEM[63,101] 2017	Global	Health technologies	HTA policymakers; Healthcare professionals; Patients; Public	Decision making on health technologies	Societal perspective	Literature review	20 attributes	No	Weighting by direct rating scale or hierarchical point allocation for quantitative attributes and judgements for qualitative attributes from stakeholder panel	NR

									(including patients/public)	
The Patient-Perspective Value Framework[30] 2017	NR	Health technologies	HTA policymakers Healthcare professionals Researchers Industry Patients	Health technology value assessment and shared decision making between clinicians and patients	Patient's perspective	stakeholder panel consensus (including patients)	4 attributes	No	NR	NR
Alonso-Coello et al. [31,102] 2016 (GRADE EtD framework)	Global	Health technologies	HTA policymakers; Healthcare professionals Guideline developers and users	Clinical recommendations, coverage decisions, and health system or public health recommendations and decisions regarding health technologies	Patient's or health system perspective	Literature review; consultation, testing and survey among stakeholder panel (including patients and the public)	10 attributes	NR	Stakeholder panel judgement (including patients and the public)	NR
Asaria et al.[77] 2016 (Distributional CEA)	NR	Health technologies	HTA policymakers	Health technology prioritization decisions	NR	NR	7 attributes	Yes	Social distributions of health gains associated with different health technologies	Use of dominance rules and social welfare indices
CADTH[62] 2016 (pCODR Framework)	Canada	Cancer drugs	HTA policymakers	Public funding decisions on cancer drugs	NR	NR	8 attributes	Yes	Policymaker judgements	NR
Dunlop et al. [40] 2016 (BEACON framework)	Global	Drugs	Pharmaceutical companies	Development of drugs	The health system perspective	Literature review; stakeholder panel consensus (without patients)	6 attributes	Yes	Pharmaceutical company judgement using a color-coding system	NR
Garrison et al[64] 2016 (OHE/EPEMED framework for diagnostics)	Global	Diagnostics	HTA policymakers	Policy recommendations about complementary diagnostics	Societal perspective	Literature review	10 attributes	Yes	NR	NR
Iskrov et al[46,103] 2016	Bulgaria	Orphan drugs	HTA policymakers	Orphan drug appraisal and	Societal perspective	Literature review; stakeholder panel	11 attributes	NR	Weighting by a two-step 0-100 scale elicitation	50 and 70 as the two cut-off points for reimbursement

				reimbursement decisions		survey (including patients)			technique from stakeholder panel (including patients)	and conditional reimbursement
Pooripussarakul et al. [71] 2016	Thailand	Vaccines	HTA policymakers	Vaccine introduction and reimbursement decisions	Societal perspective	Literature review; stakeholder panel consultation (without patients/public)	7 attributes	No	Best worst scaling-derived weights from stakeholder panel (without patients/public)	NR
Schnipper et al. [12] 2016 (ASCO framework)	Global	Cancer treatments	Healthcare professionals Patients	Shared decision making with patients on cancer treatments	Patient's perspective	Stakeholder panel consultation (including patients)	3 attributes	No	Scoring by net health benefit algorithm developed by stakeholder panel (including patients)	NR
Verguet et al.[78] 2016 (Extended CEA)	NR	Health technologies	HTA policymakers	Health technology coverage and funding decisions	Patient's perspective	NR	4 attributes	Yes	Calculation of a financial protection ICER and display of the distribution of health benefits, financial risk protection and private expenditures crowded out by population stratum	NR
IOM/NAE[67,104] 2015 (SMART Vaccines Framework)	Global	Vaccines	HTA policymakers Healthcare professionals Industry Research units	Vaccine development and investment decisions	Societal perspective	Literature review; stakeholder panel consultation (without patients/public)	up to 35 attributes	Yes	Rank order centroids-derived weights from stakeholder panel (without patients/public)	A set of boundary values for different attributes in the framework
NHS[69] 2015 (CDF Prioritisation Tool)	UK	Cancer drugs	HTA policymakers	Cancer drug funding decisions	The health system perspective	Literature review; stakeholder panel consultation (including patients and public)	8 attributes	Yes	A prioritisation scoring tool developed by HTA	Aggregate score threshold

									policymakers and experts	
Miller et al. [53] 2015	US	Public health programs	HTA policymakers Public health officials Researchers	Nonclinical prevention program funding decisions	Societal perspective	Literature review	7 attributes	No	NR	NR
MSKCC[13] 2015 (Drug Abacus)	US	Cancer drugs	HTA policymakers Healthcare professionals	Value-based pricing for cancer drugs	NR	Public data companies sent into the FDA to obtain approval	8 attributes	No	User-defined weighting	Estimated price is the highest acceptable price for the product
NCCN[14] 2015 (NCCN evidence blocks)	Global	Cancer treatments	Healthcare professionals Patients	Shared decision making with patients on cancer treatments	Patient's perspective	NR	5 attributes	No	Healthcare professional judgements and discussions with patients	NR
Paulden et al. [55] 2015	Canada	Orphan drugs	HTA policymakers	Orphan drug coverage and reimbursement decisions	The health system perspective	Literature review; stakeholder panel consultation (including patients)	19 attributes	Yes	Net value calculation from stakeholder panel (including patients)	NR
Radaelli et al. [73] 2014 (VTS framework)	Italy	Health technologies	HTA policymakers	Health technology coverage and delisting decisions	NR	Literature review	21 attributes	NR	Weighting on a 1–8 scale for quantitative attributes and judgements for qualitative attributes from stakeholder panel (without patients/public)	NR
Venhorst et al. [75] 2014	Low and middle-income countries	Breast cancer treatments	HTA policymakers	Breast cancer control policy development	NR	Literature review; Delphi study within stakeholder panel (including patients)	10 attributes	NR	Weighting by a 1-5 Likert scales from stakeholder panel (including patients)	NR
Marsh et al. [52] 2013	UK	Public health programs	HTA policymakers	Funding for preventative health program	NR	Literature review; policymaker consultation and consensus	5 attributes	Yes	DCE-derived weights from policymakers	NR

Seigfried et al. [74] 2013	US	Biotechnology drugs	Hospital formulary committee	Formulary listing	NR	Literature review	20 attributes	No	Scoring by a value scorecard from healthcare professionals	NR
Anonychuk et al. [60] 2012	NR	Diagnostics	HTA policymakers Hospital decision-makers	Diagnostics coverage and reimbursement decisions	The health system perspective	NR	10 attributes	No	Expert judgements	NR
Golan and Hanson[43,86] 2012 (The Israel VfM framework)	Israel	Health technologies	HTA policymakers	Health technology coverage and funding decisions	Societal perspective	Literature review	4 attributes	No	PAPRIKA-derived weights from policymakers	NR
IOM[66] 2012	US	Public health programs	HTA policymakers	Community-based public health program funding decisions	Societal perspective	Expert consensus	4 attributes	Yes	Expert judgements	NR
Patel et al. [54] 2012 (VALUE Framework)	US	Health technologies	Educators and residency programs	Decision making on health technology use in clinical practice	NR	NR	5 attributes	No	Healthcare professional judgements and discussions with patients	NR
Winquist et al.,[76] 2012	Canada	Orphan drugs	HTA policymakers	Orphan drug coverage and reimbursement decisions	The health system perspective	stakeholder consensus (without patients/public)	5 attributes	Yes	Judgements from stakeholder panel (without patients/public)	NR
Youngkong et al. [59] 2012	Thailand	Health technologies	HTA policymakers	Public funding of health technologies	NR	Literature review; stakeholder panel consensus (including patients andpublic)	8 attributes	Yes	Equal weights were used by the stakeholder panel (including patients and the public)	A threshold of one-time per capita GDP per QALY gained.
Diaby and Lachaine[38] 2011	Côte ''Ivoire	Drugs	HTA policymakers	National formulary listing and drug reimbursement decisions	The health system perspective	Literature review; expert consensus	3 attributes	Yes	DCE-derived weights from experts	Selecting top X medicines to be reimbursed until the available budget is exhausted
Gibson et al. [42] 2011	Canada	Health service programs	HTA policymakers	Health service program funding decisions	NR	Literature review; LHIN stakeholder panel consensus	15 attributes	No	A criteria-based scoring tool from LHIN stakeholder	NR

						(including the public)			panels (including the public)	
Stafinski et al. [58] 2011	Canada	Health technologies	HTA policymakers	Health technology coverage and funding decisions	NR	Literature review; policymaker consensus	6 attributes	NR	Policymaker judgements	NR
Broqvist et al. [61] 2011 (The National Model for Transparent Prioritisation in Swedish Health Care)	Sweden	Health technologies	HTA policymakers	Health technology coverage and funding decisions	Societal perspective	NR	10 attributes	Yes	Priority level assigned by policymakers on a 1-10 scale against each attribute	NR
Houweling et al. [65] 2010	Netherlands	Vaccines	HTA policymakers	Vaccine coverage and reimbursement decisions	Societal perspective	Literature review	7 attributes	Yes	Policymaker and expert judgements by reference to hierarchical attributes	NR
Kroese et al. [68] 2010 (UKGTN framework)	UK	Genetic tests	HTA policymakers	Coverage decision for the NHS Directory of Molecular Genetic Testing	NR	Literature review; policymaker and expert consensus	5 attributes	No	Weighting by assigning percentage points from policymakers and experts	NR
Lee et al. [50] 2010	US	Diagnostics	HTA policymakers Healthcare professionals Researchers Patients	Public policy decisions and diagnostic coverage and funding decisions	NR	NR	3 attributes	No	NR	NR
Porter[72] 2010	NR	Health technologies	NR	Health technology value assessment	NR	NR	6 attributes	No	NR	NR
Maciosek et al. [51] 2009	US	Public health programs	HTA policymakers	Public health program coverage decisions	Societal perspective	Literature review; experience from policymakers	9 attributes	Yes	Three criteria combination approaches (ordered sequence, weighting system and simultaneous consideration) were discussed.	NR

Oregon HSC[70] 2009	US	Health service programs	HTA policymakers	Health service program coverage decisions	NR	stakeholder panel consensus (without patients/public)	8 attributes	Yes	Weights and scores assigned by stakeholder panel (without patients/public)	Selecting top X medicines to be reimbursed until the available budget is exhausted
Piso and Wild[57] 2011	NR	Vaccines	HTA policymakers	Vaccine coverage decisions	NR	Literature review	12 attributes	Yes	NR	NR
Browman et al. [35] 2008 (6-STEPPPs framework)	Canada	Cancer treatments	HTA policymakers	Cancer treatment funding decisions	NR	NR	8 attributes	No	Use of a scoring system and deliberation within stakeholder panel (including patients)	NR
Jehu-Appiah et al. [47] 2008	Ghana	Health technologies	HTA policymakers	Health technology coverage and reimbursement decisions	Societal perspective	Focus group of policymakers	5 attributes	Yes	DCE-derived weights from policymakers	NR

6-STEPPPs: Systematic Tool for Evaluating Pharmaceutical Products for Public Funding Decisions; **ASCO**: American Society of Clinical Oncology; **AVF**: Advance Value Framework; **BAS**: Balanced Assessment System; **BEACON**: Burden/target population, Environment, Affordability/value, Comparator, Outcomes, Number of studies/quality of evidence framework; **CADTH**: Canadian Agency for Drugs and Technologies in Health; **CCO**: Cancer Care Ontario; **CDF:** Cancer Drug Fund; **EBV**: evidence-based valuation; **DCE**: Discrete Choice Experiment; **ESMO-MCBS**: the European Society for Medical Oncology Magnitude of Clinical Benefit Scale; **EPI**: The Expanded Programme on Immunization; **EVIDEM**: the Evidence and Value: Impact on dEcision Making framework; **FDA**: Food and Drug Administration; **GDP**: gross domestic product; **GRA-QALY:** the Generalized Risk-Adjusted Quality-Adjusted Life Year; **GRADE EtD**: the Grading of Recommendations Assessment, Development and Evaluation Evidence to Decision framework; **HSTs**: Highly specialised medical technologies; **HTA**: health technology assessment; **ICER**: Institute for Clinical and Economic Review; **ISPOR**: International Society for Pharmacoeconomics and Outcomes; **IOM**: The Institute of Medicine; **LHIN**: Local Health Integration Network; **HSC**: Health Services Commission; **MSKCC**: Memorial Sloan Kettering Cancer Center; **NAE:** National Academy of Engineering; **NCCN**: National Comprehensive Cancer Network; **NHS**: National Health Service; **NR**: not reported; **OHE/EPEMED**: Office of Health Economics/The European Personalised Medicine Association; **OHTAC**: the Ontario Health Technology Advisory Committee; **PAPRIKA**: Potentially All Pairwise RanKings of all possible Alternatives; **pCODR**: the pan-Canadian Oncology Drug Review; **QALY:** Quality-Adjusted Life Year; **SMART Vaccines**: Strategic Multi-Attribute Ranking Tool for Vaccines; **UK**: The United Kingdom; **US**: The United States; **UKGTN**: The UK Genetic Testing Network; WHO: World Health Organization; **VALUE**: Validation and Variability, Affordability and Access, Long-term and Less side effects, Utility and Usability, Effectiveness and Errors; **VfM**: value for money; **VTS**: Valutazione delle Tecnologie Sanitarie; **WHSSC**: the Welsh Health Specialised Services Committee;

Footnotes:

*: The members on the stakeholder panel represent at least three groups of stakeholders, which include policymakers, experts (academics, methodologists, or healthcare professionals), patients, public, healthcare administrators, non-government organizations, social service workers and industry.

**:This framework uses the Generalized Risk-Adjusted QALY (GRA-QALY).

Appendices

Appendix 1. Search strategy

Databases: PubMed (January 1, 2008, to October 1, 2019)

Search strategy:

#33 Search #13 AND #20 AND #23 Filters: Publication date from 2008/01/01 to 2019/10/01 Sort by: Publication Date 4491

#23 Search framework* 255737

#20 Search #14 OR #15 OR #16 OR #17 OR #18 OR #19 854487

#19 Search "Insurance Coverage"[Mesh] 15840

#18 Search "Costs and Cost Analysis"[Mesh] 228541

#17 Search "Insurance, Health, Reimbursement"[Mesh] 44727

#16 Search "Reimbursement Mechanisms"[Mesh] 35959

#15 Search cost 826174

#14 Search reimburs* 45349

#13 Search #1 OR #2 OR #3 OR #4 OR #5 OR #6 OR #7 OR #8 OR #10 OR #11 OR #12 2579539

#12 Search "Decision Support Techniques"[Mesh] 74675

#11 Search "Decision Making, Organizational"[Mesh] 11049

#10 Search "Decision Making"[Mesh] 192290

#8 Search "Technology Assessment, Biomedical"[Mesh] 10835

#7 Search "decision* making" 198770

#6 Search decision* 429609

#5 Search "value* assessment*" 246

#4 Search "value* driven" 553

#3 Search "value* based framework" 10

#2 Search "value* based" 5412

#1 Search value* 2075815

Database: OVID EMBASE (January 1, 2008 to September 27, 2019)

Search Strategy:

1 value*.mp. (2635068)

2 value* based.mp. (10170)

3 value* driven.mp. (401)

4 value-based framework.mp. (19)

5 exp biomedical technology assessment/ (13808)

6 value* assessment.mp. (535)

7 decision*.mp. or exp decision support system/ or exp decision making/ (680424)

8 decision making.mp. (410972)

9 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 (3236351)

10 reimbursement.mp. or exp reimbursement/ (65914)

11 exp "cost"/ or exp "healthcarecost"/ (339044)

12 exp health insurance/ or exp value-based insurance design/ (253774)

13 10 or 11 or 12 (541836)

14 exp conceptual framework/ or framework.mp. (268877)

15 9 and 13 and 14 (3014)

16 limit 15 to yr="2008 -Current" (2591**)**

Database: Cochrane Library (January 1, 2008 to September 30, 2019)

Search strategy:

ID	Search	Hits
#1	MeSH descriptor: [Technology Assessment, Biomedical] explode all trees	139
#2	MeSH descriptor: [Decision Making] explode all trees	3960
#3	MeSH descriptor: [Decision Making, Organizational] explode all trees	42
#4	MeSH descriptor: [Decision Support Techniques] explode all trees	2420
#5	value*	158299
#6	"value* based"	296
#7	"value* based framework"	0
#8	"value* driven"	10
#9	"value* assessment*"	56
#10	decision*	33595
#11	"decision* making"	14454
#12	#1 OR #2 OR #3 OR #4 OR #5 OR #6 OR #7 OR #8 OR #9 OR #10 OR #11	184923
#13	reimburse*	1882
#14	COST	56991
#15	MeSH descriptor: [Reimbursement Mechanisms] explode all trees	234
#16	MeSH descriptor: [Insurance, Health, Reimbursement] explode all trees	263
#17	MeSH descriptor: [Costs and Cost Analysis] explode all trees	10038
#18	MeSH descriptor: [Insurance Coverage] explode all trees	69

#19 #13 OR #14 OR #15 OR #16 OR #17 OR #18 58460

#20 framework* 7173

#21 #12 AND #19 AND #20 with Cochrane Library publication date Between Jan 2008 and Oct 2019 849

Database: NIHR (January 1, 2008 to October 1, 2019)

Search strategy:

NIHR: Nov 20, 2018

value-based framework

1 (value*) 12051

2 ("value* based") 77

3 MeSH DESCRIPTOR Technology Assessment, Biomedical EXPLODE ALL TREES 560

4 ("value* based framework") 0

5 ("value* driven") 0

6 ("value* assessment*") 0

7 (decision*) 13380

8 ("decision* making") 1277

9 MeSH DESCRIPTOR Decision Making EXPLODE ALL TREES 447

10 MeSH DESCRIPTOR Decision Making, Organizational EXPLODE ALL TREES 17

11 MeSH DESCRIPTOR Decision Support Techniques EXPLODE ALL TREES 1629

12 #1 OR #2 OR #3 OR #4 OR #5 OR #6 OR #7 OR #8 OR #9 OR #10 OR #11 20844

13 (reimburs*) 1353

14 (cost) 22534

15 MeSH DESCRIPTOR Reimbursement Mechanisms EXPLODE ALL TREES 169

16 MeSH DESCRIPTOR Insurance, Health, Reimbursement EXPLODE ALL TREES 266

17 MeSH DESCRIPTOR Costs and Cost Analysis EXPLODE ALL TREES 17164

18 MeSH DESCRIPTOR Insurance Coverage EXPLODE ALL TREES 39

19 #13 OR #14 OR #15 OR #16 OR #17 OR #18 23220

20 (framework*) 927

21 #12 AND #19 AND #20 428

22 * FROM 2008 TO 2019 52821

23 #21 AND #22 220

Appendix 2. Classification of attributes of included value-based frameworks

1. **Health benefits of technology**

   It assesses health technology's efficacy, effectiveness, safety, or its impact on patient-reported outcomes. Terms used in existing VAFs include:

   - (comparative) clinical effectiveness, clinical outcomes, clinical impact, effect size, the magnitude of effect, etc.
   - safety and tolerability of the technology, unintended consequences, toxicity of treatment, side effects etc.
   - patient preferences or patient reported outcomes, utilization and patient adherence, quality of life, etc.
   - the magnitude of benefit and harm.

2. **Quality of evidence**

   It assesses the credibility and certainty of evidence related to the use of health technology. Terms used in existing VAFs include:

   - quality of evidence
   - uncertainty about the magnitude or durability of the long-term benefits of this intervention and the long-term risk of serious side effects of this intervention
   - completeness and consistency of documentation; and relevance and validity of documentation
   - certainty of benefit and harm and type of analyses conducted
   - etc.

3. **Cost-effectiveness**

   It assesses the resource use efficiency of health technology. Terms used in existing VAFs include:

   - cost-effectiveness
   - estimated incremental cost-effectiveness
   - impact on efficiency (cost-opportunity)
   - opportunity costs
   - cost-benefit
   - efficiency of vaccination
   - etc.

4. **Innovation**

It assesses the level of novelty and advancement in in terms of the mechanism of action or the approach/technique used to improve the properties, usage, and performance of health technology. Terms used in existing VAFs include:

- a novel mechanism of action or approach that will allow successful treatment of many patients for whom other available treatments have failed
- innovation profile of treatment
- scientific spillover
- vaccine characteristics and presentation, administration schedule including the number of doses and possible combinations
- companion tests of health technology
- whether or not medication is injected
- etc.

5. **Burden of disease**

It describes mortality, morbidity, and economic impact of the target disease.

Terms used in existing VAFs include:

- number of potential beneficiaries, burden of the condition, size of population, the proportion of population eligible for the intervention, prevalence, rarity etc.
- severity of disease, impact on suffering, health impact of disease, extent to which the disease is life-threatening or chronically debilitating without treatment etc.
- costs burden of the disease, social impact of disease, disease raises fear and stigma in the public etc.

6. **Unmet needs**

It describes availability or limitations of alternative health technologies. Terms used in existing VAFs include:

- availability of alternatives, limitations of alternative technologies in use, trail comparator, need for medical services etc.
- conformity of programs, expert consensus/clinical practice guidelines, background of assessment: other jurisdictions' decisions etc.

7. **Affordability**

It assesses costs and budget impact associated with the use of health technology.

Terms used in existing VAFs include:

- budget impact, financial impact on health system, impact on other spending, macroeconomic benefit etc.

- (net) cost of treatment, other medical cost, non-medical costs, resource requirements, cost savings within and outside healthcare system etc.
- affordability of regimen.

8. Ethics and equity

It describes the equity/ethical implications and impact associated with the use of health technology. Terms used in existing VAFs include:

- impact on equity and patient accessibility, reduction of health disparities across racial, ethnic, gender, socioeconomic or regional categories etc.
- vulnerability of population affected, benefits on socioeconomically disadvantaged populations.

9. Societal impact of health technology

It describes the sociocultural, organizational, legal and political implications associated with the use of health technology. Terms used in existing VAFs include:

- poverty reduction, reduced caregiver or family burden, improved productivity, impact on public health, prevention of future illness, humanistic improvement in benefit over comparator etc.
- historical, cultural and political context, acceptability within the political system, acceptability within stakeholders etc.
- coherence with national/regional planning, consistent with societal values, alignment with normative contextual criteria regarding population priorities, mandate and scope of health system, common goal and specific interests, social implications, organizational implications etc.
- feasibility within the health system, system capacity and appropriate use of intervention, strategic fit etc.

CHAPTER 3. IDENTIFING ATTRIBUTES FOR A VALUE ASSESSMENT FRAMEWORK IN CHINA: A QUALITATIVE STUDY

Status: This manuscript has been accepted by Pharmacoeconomics in December 2022. It is reproduced here with permission.

Identifying attributes for a value assessment framework in China: a qualitative study

Mengmeng Zhang[1], Yun Bao[2], Yi Yang[3], Melissa Kimber[1,4], Mitchell Levine[1,5], Feng Xie[1,6]

1. Department of Health Research Methods, Evidence, and Impact, McMaster University, Hamilton, Ontario, Canada
2. Institute of Clinical Research and Evidence Based Medicine, Gansu Provincial Hospital, Lanzhou, Gansu, China.
3. Key Lab of Health Technology Assessment, National Health Commission, Fudan University, Shanghai, China
4. Offord Centre for Child Studies, Department of Psychiatry and Behavioural Neurosciences, McMaster University, Hamilton, Ontario, Canada.
5. Division of Clinical Pharmacology and Toxicology, Department of Medicine, McMaster University, Hamilton, Ontario, Canada.
6. Centre for Health Economics and Policy Analysis, McMaster University, Hamilton, Ontario, Canada.

Correspondence:

Feng Xie, Department of Health Research Methods, Evidence, and Impact, Centre for Health Economics and Policy Analysis (CHEPA), McMaster University

1280 Main Street West, Hamilton, Ontario L8S 4L8, CANADA.

Email address: fengxie@mcmaster.ca Phone: 1-905-525-9140

Author Contributions: Concept and design: Zhang, Kimber, Xie

Acquisition of data: Zhang, Bao, Yang

Analysis and interpretation of data: Zhang, Bao, Yang, Kimber, Levine, Xie

Drafting of the manuscript: Zhang

Critical revision of the paper for important intellectual content: Zhang, Yang, Kimber, Levine, Xie

Supervision: Xie

Other: Supervisory committee: Kimber, Levine

Conflict of Interest Disclosure: The authors reported no conflicts of interest.

Funding: The authors received no financial support for this research.

Role of the Funder/Sponsor: Not applicable.

Ethical Approval: The study was granted ethical approval by the Hamilton Integrated Research Ethics Board (HiREB, Project No. 12993) on May 11, 2021, and was conducted in accordance with the Declaration of Helsinki.

Consent to participate: All participants provided informed consent.

Consent for publication: All authors and participants provided consent for publication.

Data Availability: The data are not available to protect the privacy of participants.

Highlights

- Despite extensive discussions advocating for value-based health technology assessment and policy making in China, no empirical study has been conducted in this country.
- Through interviews with 34 Chinese stakeholders and a review and analysis of 16 government documents, we have identified 12 value attributes measuring severity of disease, health benefit, safety, economic impact, innovation, organizational impact, health equity, and quality of evidence.
- These value attributes could be used for the development of a VAF to support transparent, consistent, and robust health technology value assessment in China.

Abstract

Background: Value assessment frameworks (VAF) are promising tools for measuring the value of health technologies and informing coverage policymaking. However, most published VAFs were developed for high-income countries. This study was aimed to identify value attributes as part of the development of a VAF in China.

Methods: We used the approach of qualitative description. Specifically, we conducted open-ended semi-structured interviews with Chinese stakeholders, as well as a review and analysis of publicly available government documents related to health technology assessment (HTA) and coverage policies in China. Conventional content analysis and the constant comparison technique were used to generate value attributes. Multiple criteria were used to determine the inclusion of a value attribute, with response levels of included attributes finalized via consensus meetings among the research team.

Results: Thirty-four stakeholders living or working in China completed the semi-structured interview. These stakeholders included policymakers (n=4), healthcare providers (n=8), HTA researchers (n=6), patients and members of the general public (n=9), and industry representatives (n=7). In addition, 16 government documents were included for analysis. 12 value attributes grouped in eight categories are included in the VAF: 1) severity of disease, 2) health benefit including survival, clinical outcomes, and patient-reported outcomes, 3) safety, 4) economic impact, including budget impact to payer and to patients, and cost-effectiveness, 5) innovation, 6) organizational impact, 7) health equity and 8) quality of evidence.

Conclusion: These twelve value attributes were identified for the development of a VAF to support health technologies value assessment and coverage policymaking in China.

Introduction

The requirement to meet rising healthcare needs with scarce resources, is shifting healthcare systems from "volume-driven" to "value-driven" services and funding models.[1, 2] In a value-driven system, healthcare choices and decisions are made based on the comprehensive assessment of health technologies (e.g. drugs, devices, medical or surgical procedures and health programs).[3] In response to this shift, a number of value assessment frameworks (VAFs) have recently been developed to support health technology assessment (HTA) and subsequent coverage policymaking. These VAFs facilitate transparent and consistent decision making, and promote the adoption and diffusion of innovative health technologies in healthcare systems.[4–7] Most existing VAFs were developed for high-income countries.[8] There is a lack of VAFs developed in low- and middle-income countries that account for population preferences and limited resources available in the local setting.

China is one of the most populous countries with 1.44 billion residents and 12% (~172 million) over 65 years old.[9] China's healthcare system is facing unprecedented challenges in meeting the healthcare needs of its population. It was estimated that the healthcare spending in China accounted for 6.6% of its gross domestic product (GDP) in 2019. The current healthcare delivery model is primarily volume-based and the expenditures are expected to exceed 9% of China's GDP by 2035.[10, 11] As part of latest healthcare reforms, the National Healthcare Security Administration (NHSA) was established to adopt a centralized approach to drug pricing, coverage, and decision making.[12, 13] The value of new health technologies is being considered, albeit

informally, in this process. Developing a framework to guide the value assessment of new technologies can support consistent and efficient coverage policymaking which is critical to the establishment of an accessible, equitable, and sustainable healthcare system for China.

Recently, the Evidence and Value: Impact on DEcisionMaking (EVIDEM) framework has been adapted to support HTA in China.[14, 15] The EVIDEM framework was developed through literature review with some attributes noted as context sensitive.[16] However, stakeholder engagement in the adaptation of the EVIDEM framework to the context of China was limited. The input from patients and the general public was missing in the adaptation. The objective of this study was to identify key value attributes for developing a VAF for China through interviews with Chinese stakeholders.

Methods

Overview

This study was conducted as part of the development of a VAF for HTA and coverage policymaking in China. We previously completed a systematic literature review to summarize existing VAFs which informed the present study design.[8] This study focused on identification and selection of value attributes for the VAF through incorporation of multiple stakeholders' perspectives. A future study will conclude the program of work via a survey among Chinese stakeholders to develop a VAF that includes all value identified and accounts for the dependence between value attributes and the uncertainty in coverage decision-making process.

Study design

We designed a qualitative study that was informed by the principles of qualitative description (QD) to elicit stakeholders' perspectives on important attributes for assessing the value of new health technologies.[17, 18] QD seeks to provide a rich description of a phenomenon, a process, or the perspectives and perceptions of people who have direct experience with the phenomenon of interest.[17, 18] A central element of QD is staying close to the data provided by participants and to generate an overarching description of the phenomenon without too much interpretation.[17] Thus, QD emphasizes the importance of collecting and collating perceptions of events or experiences from target populations to advance our understanding about health-related phenomenon, as well as healthcare planning or services.[19] It is a research design well regarded for addressing applied research questions with healthcare policy and practice relevance.[19]

Study setting and participants

Members of the public are the consumers of healthcare services and key drivers of health technology usage.[20] It is critical to engage them in the development of VAFs to align healthcare decisions with public preferences. However, the engagement of patients and members of the public was generally limited in existing VAFs.[8] Therefore, patients and participants recruited from the public (hereafter referred to as the public) are one of the key stakeholder groups for our study. We particularly considered factors that could impact the public's perspectives and expectation on new health technology in sampling and recruitment to reflect the diversity of perspectives, experience, and expertise. These factors

include the public's geographical region (e.g., Northwest China versus South China), residence (urban versus rural) and insurance type (the urban employee basic medical insurance (UEBMI) vs urban-rural residents basic medical insurance (URRBMI)).[21–24] There are seven geographical regions in China.[25] Considerable disparities exist in their levels of economic development and health investment, with East China and South China ranking highest, Northwest China and Southwest China ranking lowest and Northeast China, North China and Central China in the middle.[26] Meanwhile, substantial urban-rural differences in personal income and economic development still exist despite increased urbanization in China in recent years.[27] On the other hand, public health insurance programs are the major form of health insurance for people in China and cover over 95% of the population.[21] There are two public health insurance programs in China: UEBMI that provides coverage to working or retired urban residents in the formal sector; URRBMI (merged from the Newly Cooperative Medical Scheme and Urban Resident Basic Medical Insurance) that provides coverage to urban residents and rural residents who are not eligible for UEBMI.[21, 22, 28, 29] Public health insurance programs are operated and organized by the local government and there is substantial variation in the amount of funding and coverage available between the different public health insurance programs, which is further complicated by additional layer of funding and coverage availability in different regions.

Informed by the information above, as well as the methodological guidelines for qualitative inquiry, participants were sampled and recruited using purposeful sampling procedures. Specifically, we used criterion, maximum variation and snowball sampling techniques.[30] With respect to criterion sampling, policymakers, healthcare providers, industry

representatives, and academic researchers were asked to describe their experience in HTA or health technology-related decision or policymaking using a pre-developed screening questionnaire (see Appendix 1 in the electronic supplementary material (ESM)). The public were required to: 1) be older than 18 years and 2) be able to understand and communicate in Mandarin. We used the maximum variation sampling approach to ensure that selected policymakers, healthcare providers, industry representatives and academic researchers varied in the years of work experience, professional status (e.g., senior versus junior), expertise (e.g., physicians versus nurses), residence area and geographical regions. The public were selected in terms of variation in age, sex, residence area, geographical regions, insurance type, socioeconomic status (e.g., occupation, education, and work activity) and current health status (e.g., presence vs. absence of disease diagnosis). Snowball sampling supplemented our recruitment efforts via asking participants to link the interviewer to individuals who might be willing and able to participate. As is customary in inductive qualitative research, sampling, data collection, and data analysis happened concurrently. Therefore, sampling continued until data saturation was achieved where the amount, variation and depth of the data was deemed capable of adequately generating a comprehensive description of value attributes from multiple stakeholders.[31–33] Data saturation was determined via independent coding of the data by two coders, as well as consensus-based discussions amongst team experts in qualitative methods, HTA, and health policy. Given the descriptive aims of our work, as well as the inclusion of multiple stakeholders who are involved in various stages of HTA and policymaking, we expected

to achieve data saturation following the completion of 35 semi-structured interviews with 7~10 participants in each stakeholder group.

Due to the Coronavirus disease (COVID-19) pandemic, we used China's major social media platform (i.e., WeChat) as the primary recruitment tool and one of the interview platforms.[34] The lead researcher (MZ) screened and selected participants following the above sampling strategy.

Data collection

We conducted one-on-one open-ended semi-structured interviews with participants. Virtual web-based technology (i.e., WeChat) or online conferencing software (e.g., Microsoft Teams, Tencent Meeting) was used. Interviews focused on encouraging the participants to describe their perceptions about important attributes when evaluating the value of a new health technology or their perspectives on the characteristics that a health technology with high value should have.[35] At the end of each interview, participants were asked to name any relevant documentation that they consulted or felt relevant to the assessment of value for health technology. Documents recommended by interview participants were also reviewed by the study team.

Interviews were conducted between June 19, 2021, and October 7, 2021. All interviews were audio recorded and transcribed verbatim except for interviews with two policymakers at their request. Each of the interviews were rendered anonymous via the transcription process. Two interviewers examined the transcripts following predeveloped transcription guidelines to ensure the accuracy of transcriptions. The guidelines provided general

formatting rules of removing identifying information, capturing nuances (e.g., long pauses from participants), and highlighting strongly expressed opinions (e.g., raised voice added with italics to communicate emphasis). Field notes were also created by the interviewers after each interview to document contextual information and to capture reflective thoughts about interview content that was perceived to be relevant to the data analysis.

To facilitate the interview with stakeholders with different background and knowledge, we developed an interview guide for each stakeholder group (see Appendix 2 in the ESM). Trained qualitative interviewers pilot tested each version of the interview guide with senior researchers to ensure that questions were asked in an appropriate and consistent way to obtain the most relevant information. Plain language was used in the guide for the public.

Data analysis

The data analysis consisted of three stages. First, we used conventional content analysis and the constant comparison technique to generate relevant concepts and categories from the interview and the documents reviewed.[19, 36–38] The content analysis was carried out immediately after each interview so that emerging questions or issues can be incorporated in the subsequent interviews. Two coders reviewed the transcripts and government issued regulations and documents independently and identified key concepts that we described as value attributes. Value attributes were then organized into categories based on the content described. Solidifying the identification, definition, organization of value attributes was achieved via consensus among the team. Based on the value attributes identified in the first round of five interviews, the transcript and any new document files

suggested by the interviewee in each subsequent interview were added to the data set for analysis. The coders used the constant comparison technique to determine whether any new value attributes or categories needed to be generated.[37] Where any new attributes or categories were identified, the interviewers went back to previously coded data to ensure that they were coded in each transcript. This iterative coding process was supplemented via analytical memoing by the coders, which captured the generation and justification for the development of new attributes and their categories until data saturation was achieved.[39]

The second component of our analysis involved the use of multiple criteria to guide decisions related to retaining or dropping value attributes. The exclusion criteria were informed by the findings of our recently published systematic review of existing VAFs.[8] Some previous VAFs included societal context, such as political, historical and cultural milieu as contextual attributes and recommended measuring them qualitatively.[8, 16, 40–42] However, the qualitative measurement methods for these attributes or the approaches of incorporating the measurement results of these attributes into the decision-making process were unknown or not reported in these VAFs.[8, 16, 40–42] This could increase the risk for inconsistency in VAF application, as well as lack of transparency in decision makings; both of these potential procedural issues contradict VAFs' primary goals of accurate and reliable value assessment to inform healthcare decisions.[20] Thus, attributes that were not measurable quantitatively or qualitatively due to unclear definition(s) from the participants, or that may pose challenges for reliable measurement using currently available methods, were excluded.

Third, response levels for each attribute were generated through discussion and consensus among the research team. The discussion was informed by the suggestions from the interviewees, and the Grading of Recommendations Assessment, Development, and Evaluation (GRADE) system.[43, 44] In GRADE, quality of evidence for each outcome is divided into four levels: very low, low, moderate and high.[43] The magnitude of effect for a single outcome can be divided into four ranges (i.e., trivial, small, moderate, or large effect) by three thresholds (i.e., small, moderate or large effect threshold).[43] We adopted the four levels for quality of evidence and this four-range approach to define response levels of other attributes included in the VAF. Symbols and color coding were adopted for the response levels of to facilitate the understanding and use of the framework.

All transcripts, memos and documents from this study were managed using NVivo (Release 1.0 / March 18, 2020). Descriptive statistics and frequencies were used to analyse and present participants' socio-demographic characteristics which was performed using Excel.

Ethical approval and consent

Approval to conduct this study was granted by the Hamilton Integrated Research Ethics Board (HiREB). All participants provided informed consent to participate in this study. An honorarium was provided to each participant after the interview. We used a series of strategies for all phases of the study to promote the rigor and trustworthiness of our research and reporting procedures, which are outlined in Appendix 3 in the ESM.

Results

This study was reported following the Standards for Reporting Qualitative Research (SRQR) reporting guideline.[45]

Participant characteristics

A total of 34 online interviews were conducted and the mean duration of interview was 64 minutes (range: 23 – 95). The data saturation was achieved after 29 interviews (1 policymaker, 8 healthcare providers, 6 academic researchers, 9 public, and 5 industry representatives), but we conducted another 5 interviews (3 policymakers and 2 industry representatives) to ensure the inclusion of participants with various demographic characteristics in these two stakeholder groups (see Appendix 4 in the ESM). A total of 16 government issued documents were identified and analyzed (see Appendices 5 & 6 in the ESM). The policymakers were from hospital, provincial and national healthcare security administration agencies. Healthcare providers included physicians, nurses, and pharmacists. HTA researchers were from academia, consulting companies, or non-governmental organizations. Industry representatives were working in various departments including market access, research and development and health economics outcomes research in pharmaceutical or medical devices companies.

Table 1 presents the characteristics of the study participants. Participants resided in 13 different provinces that spanned all seven of the geographic regions in China; 50% of the participants identified as female. Most participants were from North or East China (n=24, 70.6%), living in urban areas (n=29, 85.3%), with UEBMI (n=29, 85.3%) and with a

bachelor's degree or higher (n=31, 91.2%). Out of the 9 public participants, most were from North or East China (n=5, 55.6%), living in urban areas (n=6, 66.7%), and with UEBMI (n=6, 66.7%). Within this group, nearly half of the participants identified as female (n=4, 44.4%). Of the 16 government issued documents reviewed, 15 (94.8%) were released in the last 5 years (i.e., between 2016 - 2021). Documents were published by the National Health Commission (n=6, 37.5%), the General Office of the State Council (n=5, 31.3%), the National Healthcare Security Administration (n=4, 25%), and the National Medical Products Administration (n=1, 6.2%).

Attribute identification and selection

Table 2 displays the descriptions of all the value attributes included in the VAF, as well as illustrative quotes from the coded data. A total of 12 value attributes grouped to eight categories are included:1) severity of disease, 2) health benefit, including survival, clinical outcomes, and patient-reported outcomes (PROs), 3) safety, 4) economic impact, including budget impact to payer, out-of-pocket costs to patients, and cost effectiveness, 5) innovation, 6) organizational impact, 7) health equity, and 8) quality of evidence. Appendix 7 in the ESM presents the generation of categories and value attributes in the form of a coding tree. All participants discussed the importance of health benefits, safety, economic impact, and health equity (see Appendix 4 in the ESM). Most participants discussed the current health system context and the potential organizational impact of new health technologies (n=31, 91.18%) and quality of evidence (n= 26, 76.47%). They believed that quality of evidence should be separately rated for each characteristic. Half of the participants discussed cost effectiveness (n=17, 50%), severity of disease (n=17, 50%) and

the value of innovation in addressing unmet needs (n=17, 50%). Most interviewees (n=28, 82.4%) believed that the rankings or the relative importance of the other attributes varied across diseases of different levels of severity, and that different priorities should be assigned to the disease for coverage decision making. All 12 value attributes were discussed across all stakeholder groups; the one exception was the attribute of 'cost-effectiveness,' which was not discussed by any participant from the public (see Appendix 8 in the ESM). The public participants emphasized the importance of health benefits, safety, and out-of-pocket costs to patients. One public participant discussed innovation in addressing unmet needs. When discussing quality of evidence, the public defined evidence as recommendations from healthcare providers and other patients. All 12 value attributes have been mentioned in the government-issued policy documents. However, only two documents (12.5%) discussed severity of disease.

Ethics and societal implications were mentioned in the interviews, but it was not clear whether and how to measure them. Ten participants (29.41%) discussed ethics. However, four of them did not give a clear description of ethics. The other six described ethics with substantial variation, ranging from healthcare professionals' behaviors to no harm to patients which overlapped safety. Societal implications discussed by participants were extremely broad including demographic, cultural, economic, legal, and political context in China. It was not clear whether or how to measure these implications in the value framework and therefore, they were excluded.

Attribute levels

We categorized the severity of disease into three levels to reflect life-threatening or critical disease, severe disease and moderate or mild disease as discussed by the participants. For quality of evidence, we used the four levels for high, moderate, low, and very low. For attributes measuring health benefits, safety, cost-effectiveness, innovation, and health equity, we used the four levels for excellent, good, fair and poor. For attributes measuring costs and organizational impact, we used the four levels of none, low, moderate, and high.

Discussion

This qualitative descriptive study has identified 12 important value attributes for a VAF for health technology value assessment and decision making in China. The included attributes represent a broad range of value components related to severity of disease, health benefit, safety, economic impact, innovation, organizational impact, health equity, and quality of evidence.

Using semi-structured interview and document analysis, this qualitative study involved multiple stakeholders including patients and members of the public, policymakers, healthcare providers, HTA researchers and industry representatives for attribute identification. We identified attributes that capture aspects important to the stakeholders in China for health technology value assessment and coverage decision making by 1) purposively selecting participants who have diverse background and experience with health technology use, assessment and coverage decision making in China, 2) inductively analyzing the participants' insightful and contextual descriptions and discussions and 3) deliberately supplementing and triangulating the interview data with review of government documents related to HTA and coverage policies. Among existing VAFs, the attributes

were often identified through literature review or selected by a few healthcare providers, health economists and policymakers without direct input from the public.[8, 14] Not doing so risks missing value attributes important for the public and healthcare providers, two key parties involved in health care decision making.

Similar to most existing VAFs, our VAF includes severity of disease, health benefit, safety, and quality of evidence.[8] However, there are important differences in measuring these attributes based on inputs from the qualitative study.

First, previous frameworks usually measure severity of disease as part of burden of disease along with unmet needs or size of population. Sometimes, they include both burden of disease and budget impact to payer, or both unmet needs and innovation.[7, 16, 46] There are overlaps between these attributes. For example, the budget impact to payer takes into account the size of population. The unmet needs has been one of the criteria to determine the novelty of a health technology.[47, 48] On the other hand, some multicriteria decision analysis (MCDA) frameworks included severity of disease alongside other attributes in the weighted-sum model[7, 49, 50] which assumes independence and compensation between attributes.[51] Therefore, the inclusion of disease severity in the weighted-sum model ignores the potential interactions and dependence between disease severity and other value attributes, which has been suggested by previous studies and our discussions with participants about the relative importance of attributes for diseases at different levels of severity.[51, 52] In our framework, severity of disease was used to construct disease scenarios at different levels of severity. Budget impact to payer incorporates size of

population while innovation incorporates unmet needs. In each scenario, the relative weights of the remaining attributes are to be determined separately.

Second, the health benefit of a health technology is measured through three value attributes in our framework: survival, clinical outcomes (excluding survival) and PROs. The value framework developed by the International Society for Pharmacoeconomics and Outcomes Research, however, includes quality-adjusted life years (QALY) as a core value attribute.[53] Although QALY was not included as a separate value attribute in our VAF, both survival and PROs (including health-related quality of life) were identified as important value attributes. This categorization was used because some interviewees did not mention QALY, which might be due to that they were not familiar with the concept of QALY. Another reason was that those participants who discussed QALY were concerned about the limitations of QALY in capturing value attributes such as equity.

Third, quality of evidence was included in some existing VAFs as an overall rating of quality of evidence on all attributes.[16, 50, 54] For example, quality of evidence was included as an attribute in the weighted-sum model alongside other attributes in the EVIDEM framework.[16] The relative importance of quality of evidence ratings for different attributes (e.g., clinical outcome vs economic impact) was left to the users' judgement.[16] It was not clear what value attributes to which quality of evidence should apply and how the overall quality of evidence rating was generated in EVIDEM. In our framework, we rate quality of evidence using the GRADE approach and incorporate it in the assessment of performance level for each attribute.[43]

There is no consensus on the inclusion of cost-effectiveness alongside costs and health benefits into a VAF.[51, 55–57] Some argue that cost-effectiveness overlaps with costs and effectiveness and suggest removing cost-effectiveness or costs and effectiveness.[51, 55] We include attributes on budget impact to payer, out-of-pocket costs to patients, and health benefit alongside cost-effectiveness in our framework. This was because cost-effectiveness was mentioned by interviewees from all stakeholder groups except those from the public. Besides, cost-effectiveness measures the marginal effect of a health technology vs the comparator which supplements, instead of replacing, the measures of cost and health outcomes in the value assessment.[58].

It has been increasingly recognized that value is multi-dimensional and value assessment has expanded beyond the current cost per QALY gained approach.[53, 59, 60] Even with modifiers, the cost per QALY gained method may still be limited in capturing all the dimensions of value and it is difficult to set appropriate thresholds to facilitate decision making.[61–63] MCDA is an alternative approach that has been proposed to measure value.[8, 62, 64] This approach is originated in the discipline of operational research and is concerned with decision making situations where multiple dimensions are to be combined or aggregated.[64] It has been increasingly explored in healthcare decision making and adopted or piloted by various HTA agencies and VAFs around the world.[8, 40, 46, 57, 65–67] The multiple value attributes identified in our study offer an opportunity to evaluate the utility of MCDA methods.[8, 62, 64] Subsequently, we will construct a survey using the identified value attributes among healthcare stakeholders in China. The survey will include hypothetical drugs described by the identified attributes experimentally

varying in their levels. Appendix 9 in the ESM gives an example of the value profile that could be used in the survey and in real-world decision-makings.

Our study has a few limitations. First, people from less developed regions in China were underrepresented in our study due to their limited access to internet and online data collection platforms where we posted our recruitment advertisement and performed the interview. Second, the perspectives of policymakers might be underrepresented due to restrictions on government officials from participating in research. We have conducted a document analysis to at least partly address this limitation. Third, ethics and societal implications were not included in our VAF due to their unclear definitions and difficulty to measuring either qualitatively or quantitatively. This might limit the capacity of the framework in capturing some ethical and societal concerns.

China has made considerable effort in improving patients' accessibility to quality healthcare services while striving for the efficiency of healthcare resource use. The government has developed and adopted various policies including the zero drug mark-up policy, reform of public hospital payment method, national health insurance negotiation and centralized drug procurement. [21, 68–70] The concept of value assessment in HTA has also been increasingly discussed and debated at the national level in China.[71] These policies and progress present both challenges and opportunities for the application of our VAF in China. Despite the rapid development of HTA in China in recent years, a few issues have yet to be addressed. These issues include lack of HTA researchers with sufficient training and experience, lack of a national HTA agency that produces and endorses HTA reports and lack of the translation of the HTA evidence to informing decision-makings on

the introduction or reimbursement of health technologies.[72, 73] Furthermore, the inclusion of most new drugs relies on the centralized drug procurement process that involves price negotiation between the representatives of pharmaceutical companies and NHSA.[74] Prioritization has been given to drugs for cancers, rare diseases, chronic diseases and children's diseases.[74] The decentralized HTA system and the emphasis on price, CEA results and certain diseases in the price negotiation process might result in the lack of relevant data to support the value assessment of health technologies on attributes such as health equity, innovation and organizational impact. Thus, the use of our VAF could be impacted. However, NHSA has recently adopted a scoring checklist to facilitate the assessment of health technologies across multiple dimensions which is similar to multi-criteria decision making (MCDM).[75] The National Health Commission has also released a series of national guidelines for the comprehensive clinical evaluation of drugs since last year.[76] The guidelines have included all the attributes in our VAF except for severity of disease. These changes could open up great opportunities to validate and apply our VAF in health technology value assessment and coverage decision-making in China to improve equity and accessibility of new health technologies.

Conclusions

Twelve value attributes were identified for the development of a VAF to support transparent, consistent, and robust health technology value assessment in China.

References

1. Centers for Medicare and Medicaid Services (2021) CMS' Value-Based Programs | CMS. https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/Value-Based-Programs/Value-Based-Programs. Accessed 16 Jan 2022

2. Miller HD (2009) From Volume To Value: Better Ways To Pay For Health Care. Health Affairs 28:1418–1428

3. National Library of Medicine HTA 101: II. FUNDAMENTAL CONCEPTS. https://www.nlm.nih.gov/nichsr/hta101/ta10104.html#Heading1. Accessed 29 Jun 2020

4. Institute for Clinical and Economic Review (2020) 2020 Value Assessment Framework | Topic. In: ICER. https://icer-review.org/topic/2020-value-assessment-framework/. Accessed 20 Aug 2020

5. Schnipper LE, Davidson NE, Wollins DS, et al (2016) Updating the American Society of Clinical Oncology Value Framework: Revisions and Reflections in Response to Comments Received. J Clin Oncol 34:2925–2934

6. Alonso-Coello P, Schünemann HJ, Moberg J, et al (2016) GRADE Evidence to Decision (EtD) frameworks: a systematic and transparent approach to making well informed healthcare choices. 1: Introduction. BMJ 353:i2016

7. Angelis A, Kanavos P (2017) Multiple Criteria Decision Analysis (MCDA) for evaluating new medicines in Health Technology Assessment and beyond: The Advance Value Framework. Soc Sci Med 188:137–156

8. Zhang M, Bao Y, Lang Y, Fu S, Kimber M, Levine M, Xie F (2021) What Is Value in Health and Healthcare? A Systematic Literature Review of Value Assessment Frameworks. Value in Health. https://doi.org/10.1016/j.jval.2021.07.005

9. United Nations Population Division (2020) World Population Prospects: 2019 Revision. https://population.un.org/wpp/. Accessed 25 Sep 2020

10. National Health Commission of the People's Republic of China (2020) 2019 Statistical Bulletin of China's Health Development. http://www.nhc.gov.cn/guihuaxxs/s10748/202006/ebfe31f24cc145b198dd730603ec4442.shtml. Accessed 17 Oct 2020

11. The World Bank Healthy China : Deepening Health Reform in China : Building High-Quality and Value-Based Service Delivery (Chinese). In: World Bank. https://documents.worldbank.org/en/publication/documents-reports/documentdetail. Accessed 28 Sep 2020

12. General Office of the State Council of the People's Republic of China (2019) The pilot of National Centralized Drug Procurement Policy on Drug Utilization and Drug Expenditures. http://www.gov.cn/zhengce/content/2019-01/17/content_5358604.htm. Accessed 15 Jul 2021.

13. General Office of the State Council of the People's Republic of China (2019) National Healthcare Security Administration and other eight Departments' Report on Expanding the Pilot of The National Centralized Drug Procurement Policy on Drug Utilization and Drug Expenditures. http://www.gov.cn/zhengce/zhengceku/2019-09/30/content_5456439.htm. Accessed 15 Jul 2021.

14. Geng J, Chen X, Yu X, Huang J, He D, Dong J (2018) Study on the evidence-based decision-making framework for reimbursement technologies in view of EVIDEM. Chinese Journal of Health Policy 11:50–54

15. Bao Y, Gao B, Meng M, Ge B, Yang Y, Ding C, Shi B, Tian L (2021) Impact on decision making framework for medicine purchasing in Chinese public hospital decision-making: determining the value of five dipeptidyl peptidase 4 (DPP-4) inhibitors. BMC Health Serv Res 21:1–7

16. Goetghebeur MM, Cellier MS (2018) Can reflective multicriteria be the new paradigm for healthcare decision-making? The EVIDEM journey. Cost Eff Resour Alloc 16:54

17. Bradshaw C, Atkinson S, Doody O (2017) Employing a Qualitative Description Approach in Health Care Research. Glob Qual Nurs Res. https://doi.org/10.1177/2333393617742282

18. Neergaard MA, Olesen F, Andersen RS, Sondergaard J (2009) Qualitative description - the poor cousin of health research? BMC Med Res Methodol 9:52

19. Sandelowski M (2000) Whatever happened to qualitative description? Res Nurs Health 23:334–340

20. Dubois RW, Westrich K (2019) As Value Assessment Frameworks Evolve, Are They Finally Ready for Prime Time? Value Health 22:977–980

21. Meng Q, Mills A, Wang L, Han Q (2019) What can we learn from China's health system reform? BMJ 365:l2349

22. Shi J, Liu G (2018) Chapter 9 Health Insurance and Payment System Reform in China. In: Risk Adjustment, Risk Sharing and Premium Regulation in Health Insurance Markets: Theory and Practice. Elsevier Science, pp 263–278

23. Chen J, Lin Z, Li L, et al (2021) Ten years of China's new healthcare reform: a longitudinal study on changes in health resources. BMC Public Health 21:2272

24. Zhao C, Wang C, Shen C, Wang Q (2018) China's achievements and challenges in improving health insurance coverage. Drug Discov Ther 12:1–6

25. Baidu. Geographical regions of People's Republic of China. https://baike.baidu.com/item/%E4%B8%AD%E5%9B%BD%E5%9C%B0%E7%90%86%E5%8C%BA%E5%88%92/4221764. Accessed 21 Mar 2022.

26. Wikipedia. List of GDP per capita in provincial administrative regions of the People's Republic of China. 2022. https://zh.wikipedia.org/w/index.php?title=%E4%B8%AD%E5%8D%8E%E4%BA%BA%E6%B0%91%E5%85%B1%E5%92%8C%E5%9B%BD%E7%9C%81%E7%BA%A7%E8%A1%8C%E6%94%BF%E5%8C%BA%E4%BA%BA%E5%9D%87%E5%9C%B0%E5%8C%BA%E7%94%9F%E4%BA%A7%E6%80%BB%E5%80%BC%E5%88%97%E8%A1%A8&oldid=71307927. Accessed 25 Apr 2022.

27. Yusuf S, Saich T (2008) China Urbanizes : Consequences, Strategies, and Policies. https://doi.org/10.1596/978-0-8213-7211-1

28. National Health Security Administration (2021) Report on the development of medical insurance system in 2020. http://www.gov.cn/xinwen/2021-03/08/content_5591551.htm. Accessed 1 Jul 2022

29. Tikkanen R, Osborn R, Mossialos E, Djordjevic A, Wharton GA (2020) China. https://www.commonwealthfund.org/international-health-policy-center/countries/china. Accessed 1 Jul 2022

30. Coyne IT (1997) Sampling in qualitative research. Purposeful and theoretical sampling; merging or clear boundaries? J Adv Nurs 26:623–630

31. Morse JM (2015) Critical Analysis of Strategies for Determining Rigor in Qualitative Inquiry. Qual Health Res 25:1212–1222

32. Vasileiou K, Barnett J, Thorpe S, Young T (2018) Characterising and justifying sample size sufficiency in interview-based studies: systematic analysis of qualitative health research over a 15-year period. BMC Medical Research Methodology 18:148

33. Saunders B, Sim J, Kingstone T, Baker S, Waterfield J, Bartlam B, Burroughs H, Jinks C (2018) Saturation in qualitative research: exploring its conceptualization and operationalization. Qual Quant 52:1893–1907

34. Wechat. WeChat - Free messaging and calling app. https://www.wechat.com/. Accessed 16 Jan 2022.

35. Van den Hoonaard DK (Deborah K 1951- (2019) Qualitative research in action : a Canadian primer / Deboroah K. van den Hoonaard: Oxford University Press; 2019.

36. Elo S, Kyngäs H (2008) The qualitative content analysis process. J Adv Nurs 62:107–115

37. Lewis-Beck M, Bryman A, Futing Liao T (2022) The SAGE Encyclopedia of Social Science Research Methods. https://doi.org/10.4135/9781412950589.

38. Bowen GA (2009) Document Analysis as a Qualitative Research Method. Qualitative Research Journal 9:27–40

39. Morse JM (1995) The Significance of Saturation. Qual Health Res 5:147–149

40. Radaelli G, Lettieri E, Masella C, Merlino L, Strada A, Tringali M (2014) Implementation of EUnetHTA core Model® in Lombardia: the VTS framework. Int J Technol Assess Health Care 30:105–112

41. Piso B, Wild C (2009) Decision support in vaccination policies. Vaccine 27:5923–5928

42. Gibson J, Mitton C, DuBois-Wing G (2011) Priority setting in Ontario's LHINs: ethics and economics in action. Healthc Q 14:35–43

43. Guyatt G, Oxman AD, Akl EA, et al (2011) GRADE guidelines: 1. Introduction-GRADE evidence profiles and summary of findings tables. J Clin Epidemiol 64:383–394

44. BMJ Best Practice. What is GRADE? https://bestpractice.bmj.com/info/toolkit/learn-ebm/what-is-grade/. Accessed 24 Mar 2022

45. O'Brien BC, Harris IB, Beckman TJ, Reed DA, Cook DA (2014) Standards for reporting qualitative research: a synthesis of recommendations. Acad Med 89:1245–1251

46. Jehu-Appiah C, Baltussen R, Acquah C, Aikins M, d'Almeida SA, Bosu WK, Koolman X, Lauer J, Osei D, Adjei S (2008) Balancing equity and efficiency in health priorities in Ghana: the use of multicriteria decision analysis. Value Health 11:1081–1087

47. Trueman P, Drummond M, Hutton J (2001) Developing guidance for budget impact analysis. Pharmacoeconomics 19:609–621

48. Hofmann S, Branner J, Misra A, Lintener H (2021) A Review of Current Approaches to Defining and Valuing Innovation in Health Technology Assessment. Value Health 24:1773–1783

49. Badia X, Aguarón A, Fernández A, et al (2019) Patient involvement in reflective multicriteria decision analysis to assist decision making in oncology. Int J Technol Assess Health Care 35:56–63

50. Guarga L, Badia X, Obach M, Fontanet M, Prat A, Vallano A, Torrent J, Pontes C (2019) Implementing reflective multicriteria decision analysis (MCDA) to assess orphan drugs value in the Catalan Health Service (CatSalut). Orphanet J Rare Dis 14:157

51. Marsh KD, Sculpher M, Caro JJ, Tervonen T (2018) The Use of MCDA in HTA: Great Potential, but More Effort Needed. Value Health 21:394–397

52. Nord E, Daniels N, Kamlet M (2009) QALYs: Some Challenges. Value in Health 12:S10–S15

53. Lakdawalla DN, Doshi JA, Garrison LP, Phelps CE, Basu A, Danzon PM (2018) Defining Elements of Value in Health Care-A Health Economics Approach: An ISPOR Special Task Force Report [3]. Value Health 21:131–139

54. Iskrov G, Miteva-Katrandzhieva T, Stefanov R (2016) Multi-Criteria Decision Analysis for Assessment and Appraisal of Orphan Drugs. Front Public Health 4:214

55. Dankó D, Molnár MP (2017) Balanced assessment systems revisited. J Mark Access Health Policy 5:1355190

56. Dankó D (2014) Health technology assessment in middle-income countries: recommendations for a balanced assessment system. J Mark Access Health Policy. https://doi.org/10.3402/jmahp.v2.23181

57. Marsh K, Lanitis T, Neasham D, Orfanos P, Caro J (2014) Assessing the value of healthcare interventions using multi-criteria decision analysis: a review of the literature. Pharmacoeconomics 32:345–365

58. Neumann PJ, Ganiats TG, Russell LB, Sanders GD, Siegel JE (eds) (2016) Cost-Effectiveness in Health and Medicine. https://doi.org/10.1093/acprof:oso/9780190492939.001.0001

59. Guindo LA, Wagner M, Baltussen R, Rindress D, van Til J, Kind P, Goetghebeur MM (2012) From efficacy to equity: Literature review of decision criteria for resource allocation and healthcare decisionmaking. Cost Effectiveness and Resource Allocation 10:9

60. Reed SD, Dubois RW, Johnson FR, Caro JJ, Phelps CE (2019) Novel Approaches to Value Assessment Beyond the Cost-Effectiveness Framework. Value Health 22:S18–S23

61. Lakdawalla DN, Phelps CE (2020) Health technology assessment with risk aversion in health. J Health Econ 72:102346

62. Phelps CE, Lakdawalla DN, Basu A, Drummond MF, Towse A, Danzon PM (2018) Approaches to Aggregation and Decision Making-A Health Economics Approach: An ISPOR Special Task Force Report [5]. Value Health 21:146–154

63. Asaria M, Griffin S, Cookson R (2016) Distributional Cost-Effectiveness Analysis: A Tutorial. Med Decis Making 36:8–19

64. Hansen PH, Devlin ND (2020) Multi-Criteria Decision Analysis (MCDA) in Healthcare Decision-Making. The Oxford Encyclopedia of Health Economics

65. Diaby V, Campbell K, Goeree R (2013) Multi-criteria decision analysis (MCDA) in health care: A bibliometric analysis. Operations Research for Health Care 2:20–24

66. Angelis A, Linch M, Montibeller G, Molina-Lopez T, Zawada A, Orzel K, Arickx F, Espin J, Kanavos P (2020) Multiple Criteria Decision Analysis for HTA across four EU Member States: Piloting the Advance Value Framework. Soc Sci Med 246:112595

67. Institute for Quality and Efficiency in Health Care (IQWiG). General Methods Version 6.0. 2020. https://www.iqwig.de/en/about-us/methods/methods-paper/. Accessed 3 Jun 2022.

68. The General Office of the State Council The General Office of the State Council on Promoting the Centralized Procurement of Drugs with Volume Opinions on normalization and institutionalization. http://www.gov.cn/zhengce/content/2021-01/28/content_5583305.htm. Accessed 22 Jan 2022.

69. National Health Commission (2016) Notice on National Negotiations for Centralized Drug Purchasing. http://www.nhc.gov.cn/yaozs/s3577/201605/15fb339b6b854b8981dee3306d76ce27.shtml. Accessed 22 Jan 2022.

70. National Health Security Administration (2021) Notice of the National Health Security Administration on Printing and Distributing the Three-Year Action Plan for the Reform of DRG/DIP Payment Methods. http://www.gov.cn/zhengce/zhengceku/2021-11/28/content_5653858.htm. Accessed 22 Jan 2022

71. National Health Commission of the People's Republic of China (2020) Announcement on the Public Solicitation of Opinions on the Guidelines for the Management of Comprehensive Clinical Evaluation of Drugs.

http://www.nhc.gov.cn/yaozs/s7656/202011/d11ddc32fae84121a0dfca36b015a31d.shtml. Accessed 2 Sep 2021

72. Chen Y, Chi X, He Y, Wei Y, Oortwijn W, Shi L (2019) Mapping of Health Technology Assessment in China: Situation Analysis and International Comparison. International Journal of Technology Assessment in Health Care 35:401–407

73. Shi L, Mao Y, Tang M, Liu W, Guo Z, He L, Chen Y (2017) Health technology assessment in China: challenges and opportunities. Global Health Journal 1:11–20

74. Si L, Xu L, Chen M, Jan S (2020) Using strategic price negotiations to contain costs and expand access to medicines in China. BMJ Global Health 5:e002256

75. National Health Security Administration (2022) Announcement of the National Health Security Administration on Publishing the "2022 National Basic Medical Insurance, Work Injury Insurance and Maternity Insurance Drug Catalog Adjustment Work Plan" and related documents. http://www.nhsa.gov.cn/art/2022/6/29/art_109_8342.html. Accessed 17 Dec 2022

76. National Health Commission of the People's Republic of China (2021) Management Guidelines for Comprehensive Clinical Evaluation of Drugs. http://www.nhc.gov.cn/cms-search/xxgk/getManuscriptXxgk.htm?id=532e20800a47415d84adf3797b0f4869. Accessed 30 Dec 2021

77. National Bureau of Statistics Gross domestic product. https://data.stats.gov.cn/easyquery.htm?cn=C01. Accessed 10 Jul 2022

List of tables

Table 1. Participant characteristics

Characteristics	All participants (n=34)
Age, n (%)	
18 – 39 years old	17 (50)
40 – 49 years old	9 (26.5)
50 years old and above	8 (23.5)
Female sex, n (%)	17 (50)
Stakeholder group	
The public	9 (26.5)
Policymakers	4 (11.8)
Healthcare providers	8 (23.5)
HTA researchers	6 (17.6)
Industry representatives	7 (20.6)
Regions of China, n (%)	
North or East China	24 (70.6)
All other regions of China (northeast, northwest, south, southwest, and central China)	10 (29.4)
Urban residents, n (%)	29 (85.3)
Highest education level, n (%)	
High school or lower	3 (8.8)
Bachelor's degree or higher	31 (91.2)
Insurance type, n (%)	
UEBMI	29 (85.3)
URRBMI	5 (14.7)
Work experience in the healthcare sector, n (%)	
1 – 9 years	11 (44)
10 years or above	14 (56)

Abbreviations: HTA: health technology assessment; NA: not applicable; UEBMI: the urban employee basic medical insurance; URRBMI: the urban-rural resident basic medical insurance.

Table 2. List of value attributes included in the value assessment framework

Categories	Value Attributes	Key quotations from participants	Response level	Descriptions
Severity of disease	Severity of disease	*"Although both drugs prolong patients' survival by three months…for example, one drug is for breast cancer and the other is for pancreatic cancer, then their values are different." (#2, healthcare provider)* *"There are considerable unmet needs for these severe or critical diseases…we should first cover the drugs for these severe or critical diseases." (#6, HTA researcher)* *"We can use different sets of weights for different diseases… we can categorize the disease to four types: 1) chronic diseases, 2) life-threatening or critical diseases, 3)non-life-threatening diseases and 4) other diseases like rabies." (#16, HTA researcher)* *"For different types of diseases…relatively speaking, health insurance agency considers more about drugs with higher financial risks, thus, (the government) places extra emphasis on severe or critical diseases." (#32, policymaker)* *"For life-threatening diseases without effective treatment, medicines or medical devices that are likely to be effective in improving clinical outcomes in early or mid-stage clinical trials can be approved with conditions." (Document #7)*	○ Life threatening or critical ○ Severe ○ Moderate or mild	- Life threatening/critical refers to diseases or conditions with immediate risk of death - Severe refers to diseases or conditions where treatment is needed, otherwise there will be risk of death - Moderate/mild refers to non-life-threatening diseases and may need care at outpatient or community level
Health benefit	Survival†	*"Taking surgery as an example, short-term benefits could include decreased complication rate or hospital stay; and*		

	Clinical outcomes (excluding survival)† Patient-reported outcomes†	*long-term benefits could be longer survival or lower recurrence rate." (#10, industry representative)* *"The most important thing to me is the clinical effectiveness… For example, statins are good. They stabilized my arterial plaques." (#17, the public)* *"A good drug should be able to relieve the patients' symptoms." (#20, healthcare provider)* *"I think a good drug should be able to make my life go back to normal, to reduce the pain and suffering and to relieve the symptoms." (#23, the public)* *"The most important thing for me is to extend my life…after using the drug, my life was saved, and it is already very good for me to be alive now." (#26, the public)* *"QALY, i.e., the quality-adjusted life year is a good indicator of clinical effectiveness…If we don't have QALY, then outcomes like time to recovery and changes in clinical outcomes, medical imaging results or blood test results can also be used (to assess clinical effectiveness)." (#31, industry representative)* *"Using the methods of Evidence-Based Medicine and Pharmacoeconomics, assess the safety, effectiveness, rationality, affordability and adherence of the drugs under negotiation." (Document #5)* *"The clinical benefit of the drug among the population should be assessed and determined using quantitative research methods. The key outcome measures include survival and quality of life…" (Document #16)*	 Excellent Good Fair Poor	E: Substantially longer survival/better outcomes than comparator G: Moderately longer survival/better outcomes than comparator F: Similar or non-inferior to comparator P: Substantially shorter survival/worse outcomes than comparator

Safety	Safety†	*"If the patient uses the drug, there could be some negative impacts on his/her physiological functions." (#3, HTA researcher)* *"I think safety is about adverse drug reactions and complications." (#7, industry representative)* *"When we are administering the drug to patients, we hope the drug can be very effective. However, the drug can also have side effects. We hope there can be as few side effects as possible, which is good for the patients." (#12, healthcare provider)* *"First, the drug should be safe and have no negative impacts on the patients' physiological functions." (#13, healthcare provider)* *"Hospitals need to establish a system to monitor and report adverse drug reactions, medication errors and medication incidents." (Document #6)* *"For drugs that are selected for centralized purchase, healthcare institutions should strengthen the monitoring of adverse drug reactions (ADRs) and report suspect ADRs following the national guideline on time…" (Document #12)*		
Economic impact	Budget impact to payer Out-of-pocket costs to patients	*"I do hope that the new drug can be a bit cheaper so that more patients can afford it." (#1, the public)* *"For countries without universal coverage, we need to consider the out-of-pocket costs to patients. The treatment of some diseases can bring considerable financial pressure on the patients." (#8, industry representative)* *"The costs are shared among all of us. The patients usually consider how much of the technology can be covered [by the*	N None L Low M Moderate H High	N: No cost L: Low cost M: Moderate cost H: High cost

		payer] and how much they need to pay out of pocket." (#11, healthcare provider) *"If I'm the decision maker, I would consider how large the population size is, what benefits I can bring to the people with the money I spend and whether the benefit is worth the money. (#24, HTA researcher)* *"Another thing is that the price is acceptable for ordinary people. If the drug is very expensive, ordinary people like me can't afford it." (#29, the public)* *"The budget impact analysis group and the Pharmacoeconomics group should be formed by experts recommended by local health insurance administration agencies. These two groups assess the budget impact and cost-effectiveness of the drugs under application." (Document #1)* *"The national essential drug list needs to be adjusted when ….2) there are changes in the incidence and prevalence of diseases in our country." (Document #3)* *"For the management of chronic diseases such as hypertension, diabetes and severe mental disorders, local governments should explore feasible methods to reduce patients' economic burden of drugs, enhance the public's sense of gain and ensure the essential drugs' impact on reducing drug costs and promoting rational drug use." (Document #8)*		
	Cost-effectiveness†	*"The health technology should be cost-effective, based on the results of cost minimization analysis or cost-effectiveness analysis." (#34, policymaker)*	E Excellent	E: < ¥ 80,000 per QALY, i.e., smaller than China's per-capita GDP in 2021‡

		"This is what Pharmacoeconomics studies…after we ensure its [the drug's) safety and effectiveness, its economic efficiency is also very important." (#27, healthcare provider) *"Drugs that have passed the initial review will also have to be economically efficient to enter the national medical insurance catalogue." (Document #2)*	G Good F Fair P Poor	G: < ¥ 160,000 per QALY, i.e., smaller than two times China's per-capita GDP in 2021 F: < ¥ 240,000 per QALY, smaller than three times China's per-capita GDP in 2021 P: > ¥ 240,000 per QALY, i.e., larger than three times China's per-capita GDP in 2021
Innovation	Innovation in addressing unmet needs beyond health benefit, safety, and economic impact†	*"They can give us healthcare professionals some new knowledge or provide new options in the management of some diseases where no available treatment exists or broaden our mind. Then I think they are of value to us." (#4, healthcare provider)* *"Some drugs can address the unmet medical needs where no drug is available for the treatment of a disease. Then these drugs are highly innovative." (#9, industry representative)* *"For some diseases, patients don't have any choices. There is no treatment available. Then it is not appropriate [to talk about economic efficiency] because this new drug is a breakthrough…the assessment of economic efficiency depends on the situation we are in." (#24, HTA researcher)* *"I have had this old drug. I had to take it three times a day. My doctor gave me this new drug. I only need to take one pill daily. I think this is good. This is what a good drug should be like." (#26, the public)*		E: Offering the first of its kind or the only option for a disease's diagnosis, prevention, or treatment G: Offering an improved option for a disease's diagnosis, prevention, or treatment (e.g., route, frequency, duration, and place of administration) F: Offering an option similar to currently available alternatives for a disease's diagnosis, prevention, or treatment P: Offering an option worse than currently available alternatives for a disease's diagnosis,

		"We need to consider whether it is urgent… for example, there aren't many drugs for children. Then new drugs for children will be considered first." (#30, policymaker) *"Exclusive drugs proceed to the negotiation stage. Non-exclusive drugs proceed to bidding stage." (Document #15)* *"Drugs included in the national insurance catalogue should be irreplaceable, affordable, safe and effective." (Document #13)* *"The government encourage drug innovations that have new treatment mechanisms, that can treat severe or life-threatening diseases or that can treatment rare diseases to promote the evolution of pharmaceutical technologies." (Document #4)*		prevention, or treatment (e.g., route, frequency, duration, and place of administration)
Organizational impact	Level of impact on the health system, healthcare facilities and healthcare providers (e.g., space, administration, personnel, and training)	*"There is a learning curve for new medical devices. It takes me time and practice to fully know how to use it [the medical device]." (#11, healthcare provider)* *"First, we need to know its [the device's] nature and feature. For example, does it use electricity? Does it use water? Is it giant or small?" (#19, policymaker)* *"We need to include the negotiated drugs [in the hospital]. However, there are limits on the number of drug types and the share of drug sales in the hospital." (#15, HTA researcher)* *"To improve the popularization of suitable health technologies, we need to have a team of experts and deliver training to healthcare professionals so that suitable health*	N None L Low M Moderate H High	**N:** None or negligible impact **L:** Low impact **M:** Moderate impact **H:** High impact

		technologies can be used in local hospitals." (Document #10) *"To ensure the missions of essential drugs, publicly funded hospitals should first select drugs within the national essential drug list when determining institutional formulary or drug catalogue." (Document #8)*		
Health equity	Improvement of health equity across populations (e.g., different race, age, gender, socio-economic status, or regions)†	*"We should give all patients the same right of accessing treatment. This should be taken into account when the new drug is approved, reimbursed, or introduced to hospitals or even when the drug is manufactured or developed." (#5, HTA researcher)* *"Some devices are expensive, and the hospital is not able to make a profit from it. However, from another point of view, it [the device] can satisfy the needs of some local patients. They don't have to go to other larger cities to get treatment. They can access the treatment in our local hospitals." (#18, healthcare provider)* *"When some ordinary people get sick, they don't have the ability to afford [the treatment]. I think the drug can be covered by the government so that it costs less, and we ordinary people can afford it." (#21, the public)* *"I'm not sure if this treatment can be accessible in more hospitals…It would be very helpful if we can get the treatment in our local hospital instead of hospitals in big cities." (#28, the public)* *"Manufactures should fulfil their responsibilities listed in the contracts to ensure the drug supply, especially the drugs supply to rural or inaccessible areas. (Document #11)*	E Excellent G Good F Fair P Poor	**E:** The new technology delivers much more benefits to socioeconomically disadvantaged patients **G:** The new technology delivers slightly more benefits to socioeconomically disadvantaged patients **F:** The new technology brings similar benefits to all patients (status quo) **P:** The new technology delivers more benefits to socioeconomically advantaged patients

		"We should always adjust the drug catalogue based on the affordability of the public and the government. By reducing drugs prices, we improve the accessibility and equity." (Document #2)		
Quality of evidence	Quality of evidence	*"Drugs covered by the insurance should be those recommended by guidelines…If the drug is chosen by most people and most doctors, then it is a good drug." (#1, the public)* *"I trust the doctors…I will just take the drugs recommended by the doctor rather than those on the advertisements." (#14, the public)* *"I will use drugs that are in the mid-range, which means they are accepted and popularized among us ordinary people." (#22, the public)* *"For the assessment of a technology's effectiveness and safety, we mainly follow the clinical experts' consensus which comes from two sources: physicians and the National Medical Products Administration. The report from NMPA is authoritative in evaluation of the randomized controlled trials to support the approval of the technology." (#33, policymaker)* *"For the registration of medical devices, the following documents should be submitted: … 7) documents to prove the safety and effectiveness of the devices." (Document #9)* *"…comprehensively use data from clinical trials, adverse drug reaction monitoring database, administrative database, real-word studies and literatures to quantitatively or qualitatively assess the drugs' safety, clinical effectiveness,*	⊕⊕⊕⊕ High ⊕⊕⊕◯ Moderate ⊕⊕◯◯ Low ⊕◯◯◯ Very low	**High:** We are very confident in the conclusion **Moderate:** We are moderately confident in the conclusion: the conclusion is likely to be true **Low:** Our confidence in the conclusion is limited: the conclusion may be true **Very low**: We have very little confidence in the conclusion: the conclusion is unlikely to be true

		economic efficiency, innovation, suitability, accessibility etc." (Document #14)		

Abbreviations: HTA: health technology assessment; GDP: gross domestic product; QALY: quality-adjusted life year.

†: Comparators should be clearly stated for this attribute.

‡: China's GDP per capita was ¥ 80,976 in 2021.[77] The value of GDP per capita should be updated to the GDP for the year when using the VAF.

Appendices

Appendix 1: Demographic questionnaires

In which province and city do you live?

_____

1. Do you work in the health sector?
   A. Yes
   B. No
   If A is selected, go to Question 3; If B is selected, go to question 9-12.

2. What is your occupation in the health sector?
   A. Healthcare policy makers
   B. Healthcare providers
   C. Health economics or health technology assessment researchers
   D. Pharmaceutical or medical device company
   E. Other, please specify: _____.

   If A is selected, go to Question 4 and 13
   If B is selected, go to Question 5
   If C is selected, go to Question 6 and 13
   If D is selected, go to Question 7 and 8

3. What is the level of your department?
   A. National
   B. Provincial
   C. Regional
   D. Other, please specify: _____.

4. What do you do as a healthcare provider?
   A. Physician
   B. Surgeon
   C. Pharmacist
   D. Nurse
   E. Other, please specify: _____.

5. Where do you work as a health economics or health technology assessment researcher?
   A. University or academic institute
   B. Consulting company
   C. Non-government organization
   D. Other, please specify:_____.

6. What type of health technologies is your job related to ?
   A. Drug
   B. Medical device
   C. Diagnostics
   D. Other, please specify:_____.

7. What department or area do your work in?
   A. Market access
   B. Insurance coverage
   C. Medical affairs
   D. Health economic and outcome research
   E. Other, please specify:_____.

8. Please indicate your occupation: _____.

   Please indicate your gender: _____.

9.

   Which category below includes your age?

10.

   A.18–- 29 B. 30–- 44 C. 45–- 64 D. 65 or above

11. What is the highest level of school you have completed or the highest degree you have received?
   A. Less than high school degree
   B. High school degree or equivalent
   C. Some college but no degree
   D. Bachelor's degree
   E. Graduate degree (Master or Doctor degree)

12. How long have you been involved in health technology decision making or health technology assessment ?
   A.1-3 years B. 4-6 years C. 7-10 years D. 10 years or above

13. What is your professional status?
   A. Junior B. Moderate C. Senior

Appendix 2: Interview guide

- ***For policymakers, healthcare providers, HTA researchers and industry employees***

Information about this interview protocol: This interview guide gives the interviewer an idea about what we would like to learn about stakeholders' perspectives regarding important issues/characteristics to consider during the decision-making processes. It acts as a prompt, reminding the interviewer of necessary topics to cover, questions to ask and areas to probe. All the questions are open-ended and the exact wording during interviews might change. Short questions could be used to make sure the interviewer understand what the participant said or to seek more information. These questions include "So, you are saying that…?" "Please tell me more about…?" or "Why do you think…?"

Rapport building:

- Brief introduction of the researcher and the research.

Thank you very much for agreeing to take this interview. My name is _____ and I am a researcher from _____. My research focuses on health technology assessment and patient reported outcomes measurement.

Like other countries in the world, China's healthcare expenditures are rising while healthcare resources are limited. A value assessment framework (VAF) will provide a structured and promising tool to measure and communicate the value of health technologies (e.g., drugs, diagnostics, devices etc.) in healthcare decision makings. While some VAFs have been developed in other countries, a VAF representing healthcare stakeholders', especially the public's, perspectives in China have not been developed. The objective of this study is to develop a VAF in China. In the interview, we are hoping to learn about your thoughts regarding important issues/characteristics to consider when you evaluate the value of a drug or a health technology.

- Informed consent, interviewee's permission to record the interview and the management of collected data.

Before we start, I would like to remind you that if you have any further questions about the letter of information or this study, you can contact me anytime. I would be happy to answer your questions.

For the questions in this interview, there are no right or wrong answers and all your answers will be kept confidential. During the interview, if there are any questions that you feel uncomfortable answering or you would prefer not to answer, you may skip over that section or stop the interview. You will receive an honorarium of ¥ 100 in the form of cash or gift cards at the end of this interview.

Now I will turn on the recorder and state my name, the date, and your participant ID.

[interviewer to turn on the recorder, stating interviewer initials, date and time, and participant ID#]

Interview questions:

- **Opening questions:**

1. *I would like to ask you to briefly introduce yourself and your experience with health technology value assessment and decision making.*
   Probes:
   a. *There are different types of health technologies, please describe for me the type of health technology that your work / research focuses on.*
   b. *Please describe for me the disease or condition that your work / research focuses on.*
   c. *Please describe for me your responsibilities in coverage policy making (only for policymakers).*

- **Core questions:**

2. *What do you think are the important characteristics to consider when evaluating the value of a health technology?*
   Probes:
   a. *Probe on the definitions of dimensions or factors discussed by the interviewee. For example, if the interviewee mentions "make patients feel better", we could ask "what do you mean by 'make patients feel better'"? or "How do you define 'make patients feel better'"?*
   b. *Probe on each dimension or factor the interviewee mentions.*
   c. *If the interviewee mentions a term or a concept that is not familiar to the interviewer or most people, ask the interviewee to clarify or elaborate. For example, can you please elaborate on "scientific spillover"? Or can you please explain "scientific spillover" to me?*
3. *How would you rank the importance of these issues/characteristics for value assessment of health technologies?*
   Probes:
   a. *How will the order of these issues vary for different types of diseases (e.g., critical vs mild)?*
   b. *How will the order of these issues vary for different types of health technologies (e.g., drugs vs devices)?*
   c. *How will the order of these issues vary for different levels of decision makers (e.g., national vs individual)?*
4. *In a value assessment framework, the dimensions or factors need to be combined to guide decision making. Existing frameworks combine the factors using methods of deliberation, or weighting or both. Which method do you prefer?*

Probes:

a. *Which method do you think is appropriate for China?*
b. *Which stakeholder group(s) do you think should be involved in the process?*
c. *What role should each stakeholder group play? Why?*

5. *What is your perspective on health technology assessment in China?*
   Probes:
   a. *What is your perspective on the health technology assessment of Traditional Chinese Medicine?*
   b. *What is your perspective on the health technology assessment of generic drugs in China?*
   c. *What is your perspective on the opinion that health technology assessment in China needs to be different from other countries (e.g., the UK, the US or Canada)?*

- **Wrapping-up questions**

6. *Is there anything we might have forgotten? Is there anything else do you think it is important for health technology value assessment and decision making?*

- **Closing remarks:**

Thank you again for participating in our study. As a reminder, you will receive your ¥ 100 reward in the form of cash or a gift card of your local supermarket, which one would you prefer?

I would like to remind you that there might be another interview in the future if further information is needed and I will book an appointment with you in advance. I think that is it from my end. Do you have any other questions or comments?

Thank you very much.

Now I will now turn off the recording.

[Note to interviewer: end recording]

- ***For patients and members of the general public***

Information about this interview protocol: This interview guide gives the interviewer an idea about what we would like to learn about stakeholders' perspectives regarding important issues/characteristics to consider during the decision-making processes. It acts as a prompt, reminding the interviewer of necessary topics to cover, questions to ask and areas to probe. All the questions are open-ended and the exact wording during interviews might change. Short questions could be used to make sure the interviewer understand what the participant said or to seek more information. These questions include "So, you are saying that…?" "Please tell me more about…?" or "Why do you think…?"

Rapport building:

- Brief introduction of the researcher and the research.

Thank you very much for agreeing to take this interview. My name is _____ and I am a researcher from _____. My research focuses on health technology assessment and patient reported outcomes measurement.

Like other countries in the world, China's healthcare spendings are rising while healthcare resources are limited. In this context, our health system focuses more on a drug's value or usefulness/worth rather than its quantity used in healthcare facilities. For today's interview, we would like to discuss with you what a high-value or useful drug should be like from your point of view.

- Informed consent, interviewee's permission to record the interview and the management of collected data.

Before we start, I would like to remind you that if you have any further questions about the letter of information or this study, you can contact me anytime. I would be happy to answer your questions.

For the questions in this interview, there are no right or wrong answers and all your answers will be kept confidential. During the interview, if there are any questions that you feel uncomfortable answering or you would prefer not to answer, you may skip over that section or stop the interview. You will receive an honorarium of ¥ 100 in the form of cash or gift cards at the end of this interview.

Now I will turn on the recorder and state my name, the date, and your participant ID.

[interviewer to turn on the recorder, stating interviewer initials, date and time, and participant ID#]

Interview questions:

- **Opening questions:**

1. *I would like to ask you to briefly introduce yourself and your experience with the use of health technologies such as drugs, devices, or medical procedures.*
   Probes:
   a. *Does your work involve the healthcare sector?*
   b. *What drugs / devices / procedures are you using / going through currently or recently? (e.g., drugs: Lipitor, Adalat; devices: cardiac pacemaker, coronary stent; procedures: CT scan, colonoscopy)*
   c. *If you are not using / going through any drugs / devices / procedures, what about your family or friends? Can you share with me some of their experiences?*

- **Core questions:**

2. *What do you think are the important characteristics of a high-value or useful drug? Or what do you think a high-value or useful drug should be like?*
   Probes:
   a. *Probe on the definitions of dimensions or factors discussed by the interviewee. For example, if the interviewee mentions "make me feel better", we could ask "what do you mean by 'make you feel better'"? or "How do you define 'make you feel better'"?*
   b. *Probe on each dimension or factor the interviewee mentions.*
   c. *If the interviewee mentions a term or a concept that is not familiar to the interviewer or most people, ask the interviewee to clarify or elaborate. For example, can you please explain to me what is "..."?*
3. *How would you rank the importance of these issues/characteristics for value assessment of health technologies?*
   Probes:
   a. *How will the order of these issues vary for different types of diseases (e.g., critical vs mild)?*
   b. *How will the order of these issues vary for different types of health technologies (e.g., drugs vs devices)?*
   c. *How will the order of these issues vary for different levels of decision makers (e.g., national vs individual)?*
4. *To compare the value or usefulness of different drugs, the dimensions or factors we just talked about need to be combined to help us make decisions. We can combine the factors using methods of discussion or scoring or both discussion and scoring. Which method do you prefer?*
   Probes:
   a. *Which method do you think is appropriate for China?*
   b. *Who do you think should be involved in the process?*
   c. *What role should each group of people play? Why?*
5. *What do you think needs to be done so that everyone can have access to high-value or useful drugs?*

Probes:

a. *How do you think we should evaluate the value of Traditional Chinese Medicine? How different should the evaluation be compared to Western Medicine?*
b. *What is your perspective on the evaluation of generic drugs' value in China?*
c. *What is your perspective on the opinion that drug evaluation in China needs to be different from other countries (e.g., the UK, the US or Canada)?*

- **Wrapping-up questions**
  a. *Is there anything we might have forgotten? Is there anything else do you think it is important for health technology value assessment and decision making?*
- **Closing remarks:**

Thank you again for participating in our study. As a reminder, you will receive your ¥ 100 reward in the form of cash or a gift card of your local supermarket, which one would you prefer?

I would like to remind you that there might be another interview in the future if further information is needed and I will book an appointment with you in advance. I think that is it from my end. Do you have any other questions or comments?

Thank you very much.

Now I will now turn off the recording.

[Note to interviewer: end recording]

Appendix 3:

Table S1. Strategies to promote rigor of this study

Criteria	Strategy to achieve rigor	Research phase	Actions taken in the study
Credibility	Reflexivity	All phases of study	The lead researcher (MZ) will maintain reflexive journals along this study to document and assess the impact of her perspective, position and presence on research process.
	Triangulation (data source and investigator triangulation)	Data collection and analysis	Data source triangulation will be performed using data collected from different groups of stakeholders. Investigator triangulation will be performed by two interviewers with different experiences involved in this study.
	Member checking	Data collection and analysis	Key themes and issues that emerge in the study will be discussed and confirmed in interviews with subsequent interviewees.
	Peer examination	Data collection and analysis	This process involves researchers discussing insights and problems within the research team. Since the interviews will be performed using Mandarin, all the transcripts will be in Mandarin with 10 of them translated to English for the experienced researcher to check and discuss. The coding system will be developed in Mandarin by two interviewers and subsequently translated to English for the research team to review and refine.
	Interviewing process	Data collection	Interview techniques such as establishing rapport, reframing questions and participants' responses and seeking validation of answers will be used during interview. An internally consistent interview guide will be used for all interviews.
Transferability	Reflexivity	All phases of study	As detailed above.
	Comparison of sample to demographic data	Data collection and analysis	The characteristics of informants will be compared to the demographic information on each stakeholder group (Appendix 1).
	Rich descriptions of participants' perspectives	Data collection and analysis	Rich description of the participants' experience and perspectives will be provided in our study.
Dependability	Triangulation	Data collection	By collecting data from multiple sources, the dependability of data is promoted.
	Stepwise replication	Data analysis	Each interviewer (MZ and YB) will independently code a sample of transcripts to identify codes and discuss within research team to achieve consensus around code labels and code definitions.
	Peer examination	Protocol development	Decisions with respect to the sampling and data collection in this study will be discussed within the research team.
Confirmability	Reflexivity	All phases of study	As detailed above.
	Triangulation	Data collection	As detailed above.

Appendix 4:

Table S2. Saturation grid for interviewees (n = 34)

Categories and value attributes	Interview number†																																	
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34
Severity of disease	×	×	×			×	×		×	×						×	×	×	×					×	×		×		×	×		×		
Health benefit	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×
Survival		×			×	×	×	×	×	×	×				×										×	×	×	×	×	×		×	×	
Clinical outcomes (excluding survival)	×		×	×	×	×	×	×		×	×	×	×	×	×	×	×	×	×	×	×	×			×	×	×	×	×	×	×	×	×	×
Patient-reported outcomes	×	×	×	×		×		×	×		×				×	×			×	×		×	×		×	×	×	×	×	×	×		×	
Safety	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×
Economic impact	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×
Budget impact to payer	×	×	×	×	×	×	×		×	×	×			×		×		×	×				×	×	×			×		×	×	×		×
Out-of-pocket costs to patients	×		×	×	×	×	×	×	×	×	×	×	×	×	×		×		×	×	×	×	×	×	×	×	×	×	×	×	×	×		×
Cost-effectiveness		×	×		×	×	×		×						×	×			×					×	×		×			×	×	×	×	×
Innovation		×		×	×	×			×	×						×			×	×			×	×	×		×			×	×	×	×	
Organizational impact	×	×	×	×	×		×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×		×	×		×	×	×	×	×
Health equity	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×
Quality of evidence	×		×	×	×	×	×	×		×	×			×		×	×	×	×	×		×		×	×		×	×	×	×	×	×	×	×

× indicates the concept observed.

† The interview number is the order of the participants being interviewed.

Appendix 5:

Table S3 List of included documents

No.	Title	Department	Release date
1	Administrative measures for the payment of medical devices included in the basic medical insurance [1]	National Healthcare Security Administration	2021-11-19
2	Administrative Measures for the National Essential Medical Insurance Drug List [2]	National Health Commission of the People's Republic of China	2021-11-15
3	Interpretation of the drug list approved through preliminary form review for the 2021 adjustment of national medical insurance drug list [3]	National Healthcare Security Administration	2021-07-30
4	Management Guidelines for the Comprehensive Clinical Evaluation of Drugs [4]	National Health Commission of the People's Republic of China	2021-07-21
5	Work Plan for the Adjustment of the National Medical Insurance Drug List in 2021 [5]	National Healthcare Security Administration	2021-06-30
6	Regulations on the Supervision and Administration of Medical Devices [6]	The General Office of the State Council	2021-03-18
7	The General Office of the State Council's Opinion on Promoting the normalization and institutionalization of Centralized Procurement of Drugs [7]	The General Office of the State Council	2021-01-28
8	Interim Administration Measures for the National Essential Medical Insurance Drug List[8]	National Healthcare Security Administration	2020-07-30
9	Drug Administration Law of the People's Republic of China [9]	National Medical Products Administration	2019-08-27
10	Notice on carrying out drug use monitoring and comprehensive clinical evaluation of drugs [10]	National Health Commission of the People's Republic of China	2019-04-09

11	Notice on further strengthening the management of the allocation and use of essential drugs in public healthcare institutions [11]	National Health Commission of the People's Republic of China	2019-01-17
12	The General Office of the State Council's Opinions on improving the Administration of National Essential Medical Insurance [12]	The General Office of the State Council	2018-09-19
13	Opinions on Deepening the Reform of the Review and Approval System and Encouraging the Innovation of Drugs and Medical Devices [13]	The General Office of the State Council	2017-10-08
14	Notice on strengthening pharmaceutical administration and the pharmaceutical service model transition [14]	National Health Commission of the People's Republic of China	2017-07-12
15	Notice on Administration of the Centralized Procurement of Drugs under Negotiated Agreement [15]	National Health Commission of the People's Republic of China	2016-05-20
16	Guiding Opinions of the Ministry of Health on Strengthening the Promotion of Appropriate Health Technology [16]	The General Office of the State Council	2008-04-29

Appendix 6:

Table S4. Saturation grid for documents included in the study (n = 16)

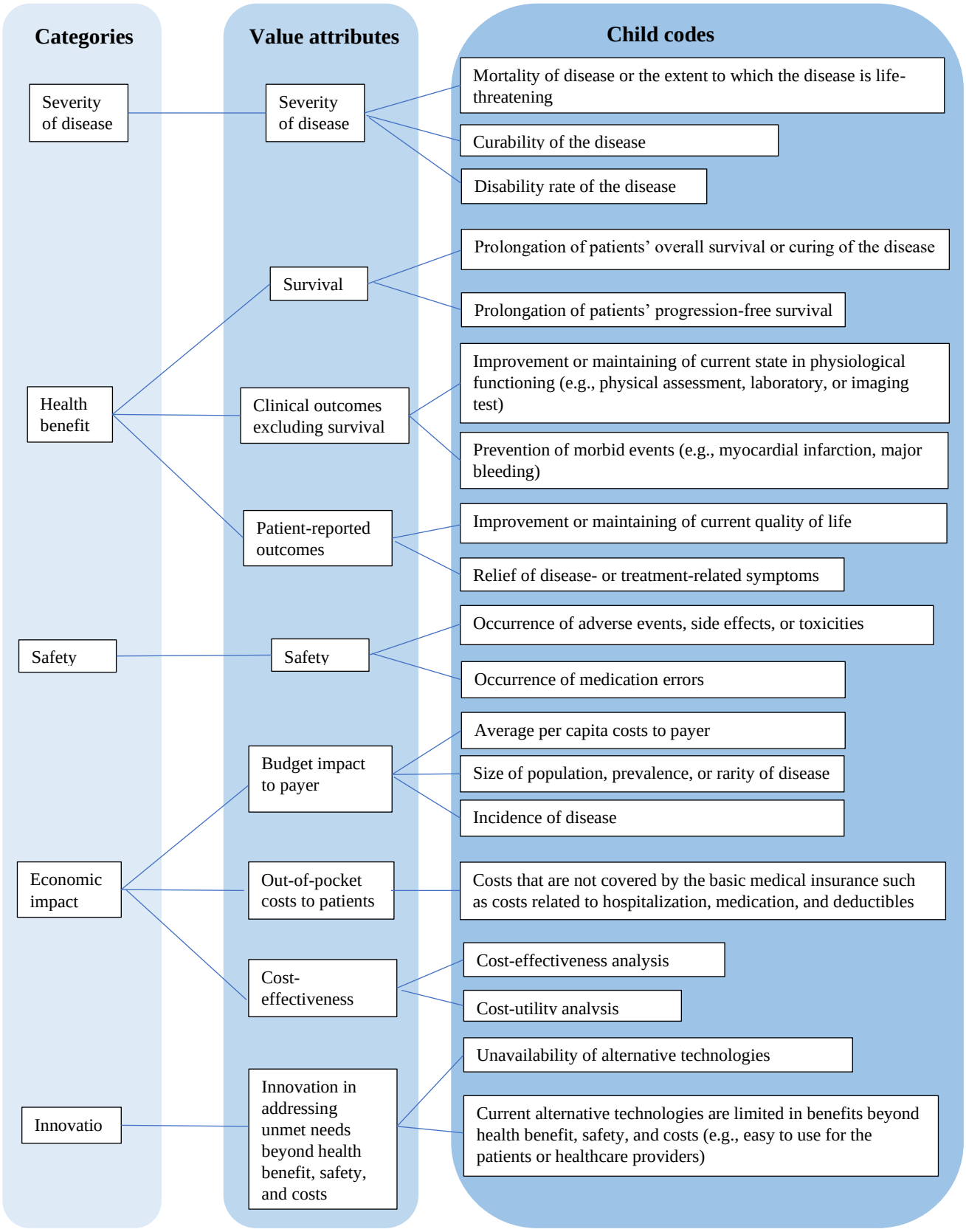
Categories and value attributes	Document number†															
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Severity of disease		×											×			
Health benefit‡	×		×	×	×	×		×	×	×	×	×	×	×	×	×
Survival				×												
Clinical outcomes (excluding survival)				×												
Patient-reported outcomes				×												
Safety		×		×	×	×	×		×					×		
Economic impact	×	×	×	×	×		×	×	×		×	×			×	×
Budget impact to payer		×		×			×	×							×	×
Out-of-pocket costs to patients	×	×		×	×				×		×	×			×	×
Cost-effectiveness		×	×	×	×						×	×				
Innovation	×	×		×			×	×	×				×			
Organizational impact		×	×	×			×	×	×		×	×	×	×	×	×
Health equity	×	×	×	×					×	×	×	×	×		×	
Quality of evidence		×		×	×	×		×		×			×			

× indicates the concept observed.

† Documents are listed in the order of release dates from most recent to oldest

‡ Most documents mentioned health benefit without specifying the dimensions of health benefit

Appendix 7:


Categories
Value attributes
Child codes
Severity of disease
Severity of disease
Mortality of disease or the extent to which the disease is life-threatening
Curability of the disease
Disability rate of the disease
Health benefit
Survival
Prolongation of patients' overall survival or curing of the disease
Prolongation of patients' progression-free survival
Clinical outcomes excluding survival
Improvement or maintaining of current state in physiological functioning (e.g., physical assessment, laboratory, or imaging test)
Prevention of morbid events (e.g., myocardial infarction, major bleeding)
Patient-reported outcomes
Improvement or maintaining of current quality of life
Relief of disease- or treatment-related symptoms
Safety
Safety
Occurrence of adverse events, side effects, or toxicities
Occurrence of medication errors
Economic impact
Budget impact to payer
Average per capita costs to payer
Size of population, prevalence, or rarity of disease
Incidence of disease
Out-of-pocket costs to patients
Costs that are not covered by the basic medical insurance such as costs related to hospitalization, medication, and deductibles
Cost-effectiveness
Cost-effectiveness analysis
Cost-utility analysis
Innovatio
Innovation in addressing unmet needs beyond health benefit, safety, and costs
Unavailability of alternative technologies
Current alternative technologies are limited in benefits beyond health benefit, safety, and costs (e.g., easy to use for the patients or healthcare providers)

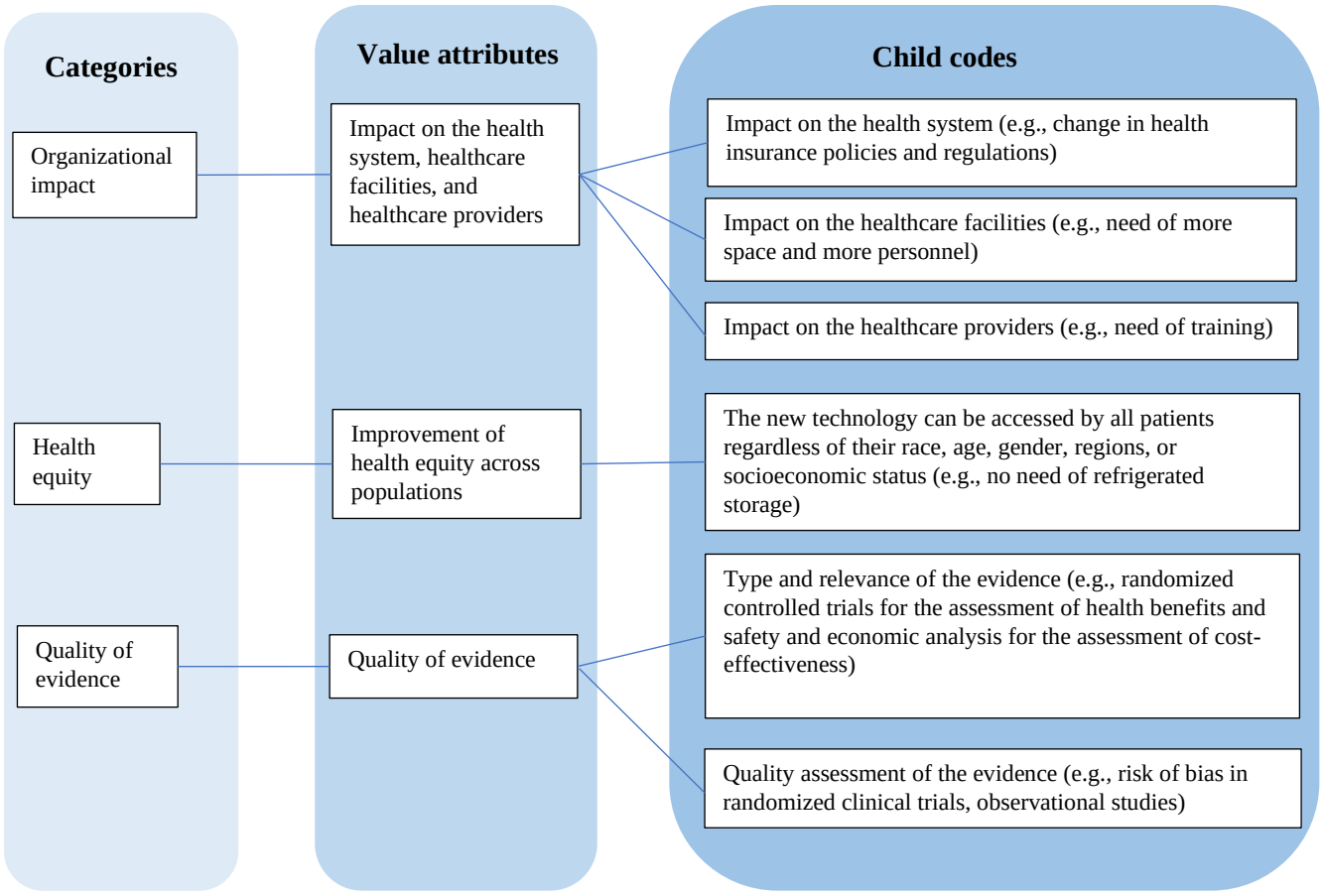


Figure S1. Coding tree representing the categories, value attributes, and child codes.

Appendix 8:

Table S5. Distribution of value attributes across stakeholder groups

Categories and value attributes	Stakeholder group				
	The public	Healthcare providers	HTA researchers	Industry representatives	Policymakers
Severity of disease	×	×	×	×	×
Health benefit	×	×	×	×	×
Survival	×	×	×	×	×
Clinical outcomes (excluding survival)	×	×	×	×	×
Patient-reported outcomes	×	×	×	×	×
Safety	×	×	×	×	×
Economic impact	×	×	×	×	×
Budget impact to payer	×	×	×	×	×
Out-of-pocket costs to patients	×	×	×	×	×
Cost-effectiveness		×	×	×	×
Innovation†	×	×	×	×	×
Organizational impact	×	×	×	×	×
Health equity	×	×	×	×	×
Quality of evidence	×	×	×	×	×

† Mentioned once by the public

Appendix 9: An example for the value profile of a hypothetical drug

Suppose that we have a new drug that is for the treatment of **a severe disease (e.g., chronic obstructive pulmonary disease)**. The drug has the following characteristics:

Attributes	Performance	Description
Survival	E	The new drug provides substantially longer survival than the comparator
Clinical outcomes	E	The new drug delivers substantially better clinical outcomes (e.g., blood pressure) than the comparator
Patient-reported outcomes	E	The new drug delivers substantially better patient-reported outcomes (e.g., quality of life) than the comparator
Safety	E	The new drug is substantially safer than the comparator
Budget impact to payer	H	The budget impact to payer with the reimbursement of the new drug is high (both the size of population and the annual costs per capita considered)
Out-of-pocket costs to patients	N	There are no out-of-pocket costs to patients with the reimbursement of the new drug
Cost-effectiveness	G	The incremental cost-effectiveness ratio of the new drug vs the comparator is lower than 160 thousand RMB
Innovation	G	The new drug offers an improved option for the disease's diagnosis, prevention, or treatment (e.g., lower administration, better ease of use)
Organizational impact	N	The new drug has none or negligible impact on the health system (e.g., no extra space required in the hospital, no extra personnel training)
Health equity	G	The new technology delivers slightly more benefits to socioeconomically disadvantaged patients vs advantaged patients than the comparator

Key:

E	Excellent	G	Good	F	Fair	P	Poor
N	None	L	Low	M	Moderate	H	High

Please note that the drug's performance on each attribute has incorporated both the drugs' effect size and quality of evidence for that attribute.

From your perspective, the value of the drug is:

Low value					Fair					High value
0	1	2	3	4	5	6	7	8	9	10
☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

The drug should be:

☐ Covered by national medical insurance

☐ Negotiated for coverage by national medical insurance

☐ Not covered by national medical insurance

References:

1. National Healthcare Security Administration (2021) Administrative measures for the payment of medical devices included in the basic medical insurance. http://www.nhsa.gov.cn/art/2021/11/19/art_113_7352.html. Accessed 1 Jul 2022

2. National Health Commission of the People's Republic of China (2021) Administrative Measures for the National Essential Drug List. http://www.nhc.gov.cn/wjw/yjzj/202111/068c31b85cb7486b9f77057b3e358aae.shtml. Accessed 28 Dec 2021

3. National Healthcare Security Administration (2021) Interpretation of the publicity of the drug list in the 2021 national medical insurance drug list adjustment through preliminary form review. http://www.nhsa.gov.cn/art/2021/7/30/art_62_5683.html. Accessed 30 Dec 2021

4. National Health Commission of the People's Republic of China (2021) Management Guidelines for Comprehensive Clinical Evaluation of Drugs. http://www.nhc.gov.cn/cms-search/xxgk/getManuscriptXxgk.htm?id=532e20800a47415d84adf3797b0f4869. Accessed 30 Dec 2021

5. National Healthcare Security Administration (NHSA) (2021) Work Plan for the Adjustment of the National Medical Insurance Drug List in 2021. http://www.nhsa.gov.cn/art/2021/6/30/art_62_5386.html. Accessed 29 Dec 2021

6. The General Office of the State Council (2021) Regulations on the Supervision and Administration of Medical Devices. http://www.gov.cn/zhengce/content/2021-03/18/content_5593739.htm. Accessed 29 Dec 2021

7. The General Office of the State Council (2021) The General Office of the State Council's opinion on Promoting the normalization and institutionalization of Centralized Procurement of Drugs. http://www.gov.cn/zhengce/content/2021-01/28/content_5583305.htm. Accessed 28 Dec 2021

8. National Healthcare Security Administration (2020) Interim Measures for the Administration of Basic Medical Insurance Medications. http://www.gov.cn/gongbao/content/2020/content_5547646.htm. Accessed 29 Dec 2021

9. National Medical Products Administration (2019) Drug Administration Law of the People's Republic of China. https://www.nmpa.gov.cn/directory/web/nmpa/zhuanti/ztypglf/ypglfzyxx/20190827083801685.html. Accessed 28 Dec 2021

10. National Health Commission of the People's Republic of China (2019) Notice on carrying out drug use monitoring and comprehensive clinical evaluation of drugs. http://www.nhc.gov.cn/cms-search/xxgk/getManuscriptXxgk.htm?id=31149bb1845e4c019a04f30c0d69c2c9. Accessed 30 Dec 2021

11. National Health Commission of the People's Republic of China (2019) Notice on further strengthening the management of the allocation and use of essential drugs in public medical institutions. http://www.nhc.gov.cn/cms-search/xxgk/getManuscriptXxgk.htm?id=b3f6fb3f55314a7faff97386908bd4f4. Accessed 30 Dec 2021

12. The General Office of the State Council (2018) The General Office of the State Council's Opinions on improving the national essential medicine system. http://www.gov.cn/zhengce/content/2018-09/19/content_5323459.htm. Accessed 28 Dec 2021

13. The General Office of the State Council (2017) Opinions on Deepening the Reform of the Review and Approval System and Encouraging the Innovation of Drugs and Medical Devices. http://www.gov.cn/zhengce/2017-10/08/content_5230105.htm. Accessed 29 Dec 2021

14. National Health Commission of the People's Republic of China (2017) Notice on strengthening pharmaceutical affairs management and transforming the pharmaceutical service model. http://www.nhc.gov.cn/cms-search/xxgk/getManuscriptXxgk.htm?id=b44339ebef924f038003e1b7dca492f2. Accessed 30 Dec 2021

15. National Health Commission of the People's Republic of China (2016) Centralized procurement of negotiated drugs. http://www.nhc.gov.cn/cms-search/xxgk/getManuscriptXxgk.htm?id=15fb339b6b854b8981dee3306d76ce27. Accessed 30 Dec 2021

16. The General Office of the State Council (2008) Guiding Opinions of the Ministry of Health on Strengthening the Promotion of Appropriate Health Technology. http://www.nhc.gov.cn/bgt/pw10803/200804/7a8706d0d4e04612b0dec3547b8bb7b0.shtml. Accessed 30 Dec 2021

CHAPTER 4. DEVELOPING THE SCORING FUNCTIONS FOR A VALUE ASSESSMENT FRAMEWORK IN CHINA: A FACTORIAL SURVEY

Status: This manuscript has been under preparation for publication at the time of writing the thesis.

Developing the scoring functions for a value assessment framework in China: a factorial survey

Mengmeng Zhang[1], Yi Yang[2], Yun Bao[3], Melissa Kimber[1,4], Mitchell Levine[1,5], Feng Xie[1,6]

1. Department of Health Research Methods, Evidence, and Impact, McMaster University, Hamilton, Ontario, Canada
2. Key Lab of Health Technology Assessment, National Health Commission, Fudan University, Shanghai, China
3. Institute of Clinical Research and Evidence Based Medicine, Gansu Provincial Hospital, Lanzhou, Gansu, China.
4. Offord Centre for Child Studies, Department of Psychiatry and Behavioural Neurosciences, McMaster University, Hamilton, Ontario, Canada.
5. Division of Clinical Pharmacology and Toxicology, Department of Medicine, McMaster University, Hamilton, Ontario, Canada.
6. Centre for Health Economics and Policy Analysis, McMaster University, Hamilton, Ontario, Canada.

Correspondence:

Feng Xie, Department of Health Research Methods, Evidence, and Impact,

Centre for Health Economics and Policy Analysis (CHEPA), McMaster University

1280 Main Street West, Hamilton, Ontario L8S 4L8, CANADA.

Email address: fengxie@mcmaster.ca Phone: 1-905-525-9140

Highlights

- We conducted a factorial survey in which 240 value profiles with 10 attributes experimentally varying in their levels were randomly assigned to 365 healthcare stakeholders in China to assess the value of hypothetical drugs and making insurance coverage recommendations.
- In the scoring functions for severe/critical disease or mild/moderate disease, health benefits and safety carried higher weights than other attributes. The value and probability of receiving insurance coverage are higher for attribute profiles for severe/critical disease than for mild/moderate disease.
- With the scoring functions, the VAF can be used to estimate the value of a technology and the probability of entering negotiation or receiving coverage in China.

Abstract

Objective: The recently developed value assessment framework (VAF) for China comprises severity of disease, health benefits, safety, economic impact, innovation, organizational impact, health equity and quality of evidence. This study was aimed to develop the scoring functions for this framework.

Methods: We implemented a factorial survey among Chinese healthcare stakeholders from July 2022 to September 2022. 240 hypothetical drug value profiles described by the VAF were grouped into 60 blocks and randomly assigned to respondents. Each respondent was assigned with one block, each presented in three disease scenarios of different levels of severity. For each profile, responses were asked to assess the value on a scale from 0 (lowest) to 10 (highest), and make one of the three insurance recommendations: cover, to be negotiated for coverage, or reject. Linear and logistic mixed-effects models were used to develop scoring functions for aggregating the value attributes.

Results: 365 respondents participated in our survey. 3,968 responses from 331 respondents were included into the analysis. Most of the included respondents were under the age of 45 (n = 256, 77.3%), females (n = 208, 62.8%), living in urban areas (n = 296, 89.4%), and with a bachelor's degree or higher (n = 303, 91.5%). Health benefits and safety carried more weights than other attributes in the scoring functions across disease scenarios. The value and probability of receiving insurance coverage for the attribute profiles for severe/critical disease were higher than for mild/moderate disease.

Conclusion: The scoring functions of the VAF can be used to assess the value of a drug and its probability of receiving insurance coverage in China.

Introduction

Scarce resources, an aging population, and rising health care demands and expenditures have driven the shift of healthcare systems from being "volume-based" to "value-based".[1,2] In a value-based system, health care choices and decisions are made based on the comprehensive assessment of health technologies (e.g. drugs, devices, medical or surgical procedures and health programs).[3] For this important transition, value assessment frameworks (VAFs) have emerged as promising tools to measure the value of health technologies and support insurance coverage decision making.[4] Beyond the conventional cost-effectiveness analysis (CEA, with or without the quality-adjusted life years (QALY) metric), existing VAFs capture broader concerns such as burden of disease, affordability and societal impact.[5] The inclusion of broader aspects of value calls for innovative approaches to the aggregation of multiple value attributes.[6]

There are three major approaches used in existing VAFs.[5] The first approach is deliberation in which decision makers achieve consensus through comparisons and discussions of the performance of various health technologies on value attributes. Intuitive and heuristic as deliberation is, the process could be relatively informal and unstructured which could result in a lack of transparency and consistency in decision making.[7,8] The second approach is expanding beyond the traditional CEA measures of health gains and costs. A few CEA-based frameworks integrate the attributes such as insurance value, severity of disease, and equity into the value assessment by calculating incremental costs per unit of financial risk protection benefits, incremental costs per generalized risk-adjusted QALY and displaying distributions of health gains across different groups of people.[9–11] However, these modified CEA frameworks may be limited in incorporating some value attributes such as innovation and organizational impact and setting appropriate thresholds to facilitate decision making.[6] Another approach is the multi-criteria

decision analysis (MCDA) which originates in the discipline of operational research where multiple attributes are to be combined or aggregated for decision making.[12] MCDA involves various methods that differ in how the attributes are combined. The most commonly used MCDA method is the weighted-sum model where each health technology is assigned a numerical value based on the combination of the technology's performance on the attributes and the weights of the attributes through a value function or scoring algorithm.[5] The weights of the attributes in those MCDA-based VAFs were elicited using various techniques including direct rating techniques (e.g., rating on a Likert scale) and choice experiments (e.g., discrete choice experiments).[5] Compared to deliberation and modified CEA, MCDA facilitates transparency in the decision-making process, bypasses cognitive errors made in intuitive judgements, and allows for the deliberation on both quantitative (e.g., CEA or MCDA results) and qualitative information (e.g., fear of contagion).[6–8] MCDA has been increasingly explored in healthcare decision making and by various HTA agencies.[5,13–18] However, using the single MCDA value index and pre-set threshold(s) to support decision making may not account for the uncertainty in real decision-making contexts even when integrated with deliberation.[19,20]

The objective of this study was to develop the scoring functions to aggregate the attributes included in a VAF for China.[21]

Methods

Overview

We used the method of factorial survey to elicit healthcare stakeholders' preferences for the attributes and develop the scoring functions.[22,23] Factorial survey, also known as vignette experiment or experimental vignette methodology, is a method to assess how people make judgements about multi-dimensional phenomena.[22,23] Respondents are confronted with

descriptions of hypothetical situations or objects (i.e., vignettes) and asked to simultaneously evaluate all the attributes and make judgements based on trade-offs which resembles real decision making process.[24,25] The vignettes consist of multiple attributes or characteristics experimentally varying in their levels to describe the situation or object.[22] This method is more intuitive and less abstract for respondents than direct rating techniques used in many MCDA-based VAFs where respondents are asked to rate or rank the relative importance of the attributes using scales or point allocation techniques.[5,12,26–28] In addition, the use of vignettes could lead to more subtle questioning and reduce the risk of social desirability bias compared to direct rating techniques.[22] Respondents are thus more likely to give their honest responses in a factorial survey.[22,25] On the other hand, factorial survey is similar to choice experiments in the use of situational descriptions such as the discrete choice experiments (DCE) adopted by a few VAFs.[18,29] However, factorial survey involves absolute evaluation of a single vignette instead of comparisons and choices between vignettes.[22,25] Thus, it allows for the incorporation of multiple questions about the single vignette. For example, what is the value of the drug under assessment? Should the drug be covered by the insurance? In contrast, DCE only asks respondents to state their choices.[22] Consequently, broader and deeper insights about the multi-dimensional assessment can be gained through factorial survey.[22,25]

Attributes included in the VAF

The attributes to construct vignettes in the factorial survey were the 12 attributes included in the VAF for China (see Appendix 1).[21] These attributes were deemed important for health technology value assessment and decision making from the perspectives of Chinese stakeholders through qualitative interviews and document analysis.[21] Most participants in our previous qualitative interviews discussed that the ranking order of attributes could change for diseases of different

levels of severity and that quality of evidence was relevant to various attributes.[21] Thus, we constructed three disease scenarios in the VAF and factorial survey: 1) life-threatening or critical diseases (e.g., acute stroke); 2) severe diseases with no immediate risk of dying but could be life-threatening if left untreated (e.g., chronic obstructive pulmonary disease); and 3) moderate or mild diseases (e.g., mild hyperlipidemia).[21] For quality of evidence, there are four levels of high, moderate, low and very low following the Grading of Recommendations, Assessment, Development and Evaluations (GRADE) system.[30] We assumed that quality of evidence was already integrated into the response levels or performance of the hypothetical drugs on the remaining 10 attributes in the value profiles presented to respondents. Each of the remaining 10 attributes had four response levels. For attributes measuring health benefits, safety, cost-effectiveness, innovation, and health equity, we used the four levels of excellent, good, fair, and poor. For attributes measuring costs and organizational impact, we used the four levels of none, low, moderate, and high. (Appendix 1). The vignettes in our study were the value profiles of hypothetical drugs which summarized their performance on the 10 attributes. Appendix 2 presents an example of the disease scenario and the hypothetical drug value profile.

Survey design

We used a fractional factorial design for the survey.[22] With 10 attributes and four levels in each attribute, there are 4^{10} value profiles that can be theoretically described by the VAF. It is not feasible to administer this full factorial in a survey. A selected subset or fraction of the full factorial saves resources and provides sufficient information.[31] The sufficiency of the information provided by the fractional factorial design or the goodness of the design can be measured by design efficiency.[22] A commonly used measure of design efficiency is the determinant efficiency (D-efficiency).[22] It has a range of 0 to 100 with 100 indicating the most efficient design.[22] The value

of 100 is only possible when the design is completely balanced (levels of single attributes occurs with the same frequency) and orthogonal (uncorrelation between any two attributes) so that all parameters can be estimated precisely.[22] However, it is often impossible for a fractional design to be completely balanced and orthogonal and thus to reach a value of 100.[22] A sample with a D-efficiency value above 80 is considered reasonable with sufficient orthogonality and level balance.[22,23] Auspurg et al. found that for a factorial survey with 9 attributes, a sample of 200 vignettes is needed to ensure that D-efficiency is higher than the value of 90.[22] Given that we have 10 attributes for each disease scenario, we drew a random sample of 240 value profiles with a D-efficiency of 80.2 using computer algorithms ("AlgDesign" package in R).[32] Appendix 3 presents the design process of the survey.

Sampling and respondents

Previous research has found that each respondent should receive no more than 10 vignettes to avoid fatigue and learning effects and each vignette needs to be assessed by at least 5 respondents to achieve sufficient statistical power.[22] Therefore, the 240 value profiles were evenly divided into 60 blocks with four value profiles in each block. Each respondent was randomly assigned one block that was presented separately in three disease scenarios (each respondent completed 12 value assessment). The target sample size was 300 respondents.

We recruited respondents through professional networks as well as WeChat, the most popular social media platform in China.[33] Respondents who age 18 years old or above, can read Chinese and have access to the internet on their mobile devices or computers were eligible. We recruited patients and members of the public (hereafter referred to as the public), healthcare providers, academic researchers, industry representatives and policymakers who were representative of their stakeholder group in terms of age, gender, education, and residence area (Appendix 4).[34] All

respondents except the public were recruited through professional networks of healthcare providers, health economists and policymakers, while the public was recruited through WeChat. Snowball sampling was used by asking the respondents to share the post to their social networks.[34]

Respondents who 1) completed the survey in less than 1 minute or more than 24 hours; or 2) gave the same answers to all 12 value assessments were excluded. These invalid responses indicate low data quality due to the respondents' lack of attention.

Survey administration and data collection

The survey was administered in Mandarin using an online survey platform Wenjuan Xing ("Survey Star") which can be circulated directly through social media platforms.[35]

The online anonymous survey comprised of three sections. The first section was the letter of information where there was a brief introduction about the survey and followed by the informed consent. The second section contained the survey instructions, a warm-up value profile, and 12 value profiles for assessment. Respondents were asked to rate the value for a hypothetical drug presented in each value profile on a scale of 0-10 with 0 lowest value and 10 highest. The respondents were also asked to make a coverage decision for each profile: cover, to be negotiated for coverage, or reject. These tasks resembled the real coverage decision-making process in China.[36] In this section, the respondents can review and change their answers if needed. The third section recorded the demographic information (e.g., age, sex, and education of the respondents (Appendix 5).

Pilot testing

We conducted a pilot test and a preliminary analysis of 10% of the target sample size (n=30) to assess the feasibility of the survey. In the pilot test, we asked respondents' perceptions about the

survey task were added at the end of the second section (Appendix 3). The time the respondents needed to complete the survey was recorded. Analysis of the pilot test data showed that 83.3% (n = 25) and 86.7% (n=26) of the respondents understood their task in the survey and could easily answer the questions, respectively. It took the respondents around 7 minutes on average to complete the survey.

Data analysis

The demographic characteristics of respondents and the distribution of the value profile blocks were described using frequencies (percentages), means (standard deviations (SDs)) or medians (interquartile ranges (IQRs)) as appropriate. We adopted the technique of mixed-effects modeling as recommended by Auspurg.[22] Given that each respondent completed multiple value profiles, the respondent was included as the random effect.[37] During the analysis we found that there were no significant difference in attribute coefficients or odds ratios estimates between the second and third best response levels in each disease scenario, between severe and critical disease scenarios or between "to be negotiated" and "cover". Therefore, we collapsed the attribute levels into "excellent", "good or fair" and "poor" or "none", "low or moderate" and "high", disease scenarios into "mild or moderate" versus "severe or critical ", and the decision categories into "to be rejected" versus "to be negotiated or cover (not to be rejected).

In the linear mixed-effects models, value scores were used as the dependent variable and the 10 attributes as the independent variables. In the logistic mixed-effects model, the decision was included as the dependent variable and the 10 attributes as the independent variables. The response levels of "poor" or "high" were used as the reference levels in the models. We assessed the overall fit of the models using the Akaike information criteria (AIC) and R^2. The collinearity was assessed using the tolerance statistics with a tolerance of larger than 0.20 acceptable. C statistic of larger

than 0.8 was acceptable for the logistic regression. The coefficients, unadjusted odds ratios (OR) and 95% confidence intervals (CI) were estimated.

The final scoring algorithm comprised the linear model and the logistic model which are used for estimating the value score and the probability of negotiation or coverage for all value profiles described by the VAF, respectively. The validity of the scoring functions was assessed by comparing the mean value and probability of negotiation or coverage over the sum of response levels on the ten attributes. We used the sum of attribute levels as a proxy indication for overall value of the profiles. The hypothesis was that the higher the sum of attribute levels, the better the drug value, and hence the higher the value estimate and the probability of negotiation or coverage. We also compared the differences in value estimate and probability between disease scenarios and presented them against the sum of attribute levels.

The significance level of 0.05 was used for all statistical tests. All the data analyses were performed using R statistical software, version 4.2.1.[38]

Ethics approval

The approval to conduct this study was obtained from the Hamilton Integrated Research Ethics Board (HiREB, Project No. #14710). Participation in this survey was completely voluntary and anonymous. There was no payment or reimbursement for participating in the study.

Results

Respondents characteristics

A total of 365 respondents participated in the survey from July to September 2022. 408 responses from 34 respondents were not valid and thus excluded. One question was not displayed properly on the survey platform and four blank responses to this question were excluded. Thus, 331

respondents with 3,968 responses were included for the analysis. The mean time to complete the task was 10 minutes 50 seconds (standard deviation: 12 minutes 7 seconds). Table 1 presents the demographic characteristics of the respondents in the survey. The excluded respondents were older (45 years and above: 41.2% vs 22.6%), more likely to live in rural areas (32.4% vs 10.6%), and receive lower education (high school and under: 38.3% vs 4.5%) compared to the included respondents. Of the 331 respondents included in the analysis, most were under the age of 45 (n = 256, 77.3%), females (n = 208, 62.8%), living in urban areas (n = 296, 89.4%), from North or East China (n = 228, 68.9%), and with a bachelor's degree or higher (n = 303, 91.5%). Out of the 90 respondents from the general public, most had the urban employee basic medical insurance (n = 48, 53.3%). Most policymakers were hospital administrators (n = 10, 83.3%). Most healthcare providers were pharmacists (n = 78, 65.5%) and from tertiary hospitals (n = 78, 65.5%). HTA researchers were mainly from academia (n = 42, 66.7%). Respondents from the industry were mainly from departments of health economics and outcomes research (n = 18, 38.3%), or market access and reimbursement (n = 11, 23.4%). Out of the 240 respondents from the health care sector, half had worked in the area for 7 years or more (n = 122, 50.6%). Appendix 6 presents the distribution of the value profile blocks among the respondents.

Linear and logistic mixed-effects models

Tables 2 & 3 present the coefficient estimates of the linear and logistic mixed-effects models in different disease scenarios. For mild/moderate disease, survival (excellent: β=1.48, 95% CI 1.28 – 1.68; OR 7.57, 95% CI 5.08 – 11.27; fair or good: β=0.95, 95% CI 0.66 – 1.24; OR 4.06, 95% CI 2.47 – 6.68), clinical outcomes (excellent: β=1.12, 95% CI 0.91 – 1.33; OR 5.47, 95% CI 3.64 – 8.22; fair or good: β=0.72, 95% CI 0.46 – 0.99; OR 2.45, 95% CI 1.55 – 3.87), patient-reported outcomes (excellent: β=0.95, 95% CI 0.75 – 1.15; OR 4.36, 95% CI 2.95 – 6.44; fair or good:

β=0.57, 95% CI 0.31 – 0.84; OR 2.43, 95% CI 1.53 – 3.87) and safety (excellent: β=0.99, 95% CI 0.79 – 1.2; OR 3.46, 95% CI 2.37 – 5.08; fair or good: β=0.55, 95% CI 0.28 – 0.82; OR 1.93, 95% CI 1.22 – 3.05) had higher coefficients and odds ratios than the other attributes. For severe/critical disease, these attributes also had higher coefficients and odds ratios than the other attributes: survival (excellent: β=1.92, 95% CI 1.74 – 2.1; OR 17.09, 95% CI 10.01 – 29.19; fair or good: β=1.05, 95% CI 0.79 – 1.32; OR 6.08, 95% CI 3.45 – 10.73), clinical outcomes (excellent: β=1.17, 95% CI 0.98 – 1.36; OR 6.07, 95% CI 3.77 – 9.78; fair or good: β= 0.77, 95% CI 0.52 – 1.01; OR 3.07, 95% CI 1.81 – 5.2), patient-reported outcomes (excellent: β=1.13, 95% CI 0.95 – 1.32; OR 4.43, 95% CI 2.81 – 7; fair or good: β=0.66, 95% CI 0.41 – 0.9; OR 2.12, 95% CI 1.27 – 3.55) and safety (excellent: β=1, 95% CI 0.81 – 1.19; OR 4.91, 95% CI 3.11 – 7.77; fair or good: β=0.55, 95% CI 0.31 – 0.8; OR 2.18, 95% CI 1.28 – 3.71). Appendix 7 presents the coefficient estimates of the logistic mixed-effects model. Appendix 8 presents the value scoring functions and gives an example of value and probability estimation.

Figures 1a & 1b show that the value score of a health technology increases over the sum of attribute levels. The value scores of drugs for the same sum of attribute levels were generally higher for severe/critical diseases than those for mild/moderate disease. Figure 1c & 1d show that the probability of entering negotiation or receiving coverage also increases over the sum of attribute levels. The probabilities of negotiation or coverage of drugs for the same sum of attribute levels were generally higher for severe/critical diseases compared with mild/moderate disease. The difference in mean probabilities of entering negotiation or coverage between the two disease scenarios was the largest when the sum of attribute levels was 17.

Discussion

Using the methods of factorial survey, we developed the scoring functions for the aggregation of attributes included in the VAF for China. In either disease scenario (mild/moderate disease versus severe/critical disease), health benefits and safety carry more weights than the other attributes. The scoring functions can be used to estimate the value score of any value profile described by the VAF as well as its probability of entering negotiation or receiving coverage.

A few existing MCDA VAFs adopted the weighted-sum model and used aggregate value scores and thresholds to facilitate decision making.[39,40] Despite the improved efficiency and transparency in these VAFs, this approach has limitations in supporting decision making. For example, Iskrov et al. developed an MCDA VAF for orphan drugs in which 70 and 50 points were the thresholds for reimbursement, conditional reimbursement, and possible rejection.[39] These thresholds were generated through stakeholder consensus but have not been validated and are considered rigid in supporting coverage decision making.[39], existing VAFs such as the HTA value framework developed by the Ontario Health Technology Advisory Committee (OHTAC) also chose not to use a quantitative model with explicit attribute weights and decision rules to avoid the rigidity in the MCDA approach.[41] Instead, they endorsed a multi-criteria decision making (MCDM) process in which value attributes were used to organize and guide deliberations and decision making on health technologies.[41] In our framework, we adopted the weighted-sum model to develop the value scoring functions.[12] However, we chose not to estimate any thresholds for decision making in the VAF. Instead, the VAF generates both value estimates and the probability of entering negotiation or receiving insurance coverage to support the decision making. Notably, the estimated probability provides decision makers with information about the likelihood of a health technology being covered. This reflects the uncertainty in real coverage negotiation and approval process in China

–a health technology with higher value may still be rejected. A strength of these scoring functions is the use of the MCDA approach to aggregate the value attributes and standardize the decision-making process while avoiding being too prescriptive in its application in decision making process.

Another strength of our study is that we considered the potential dependence between severity of disease and all other attributes that assess the performance of the technology by developing different value scoring functions for different disease scenarios. In contrast, some MCDA VAFs included severity of disease in the weighted-sum model and used the same set of attribute weights for diseases of different levels of severity.[8,42,43] Severity of disease has been recognized as an important consideration when setting priorities and assessing the value of health technologies in a number of countries.[44] Previous research has shown that the value of a particular health gain from relatively more severe disease is more likely to be deemed greater than the health gain from less severe disease.[45] Respondents tend to give priority to the relatively more severely ill.[46] Our survey also found that the value of a drug was higher for severe/critical disease than for mild/moderate disease, everything else equal. Moreover, we found that with the increase of the sum of attribute levels, the difference in the mean value scores between the two disease scenarios increased (the monotonic curve as presented in Figure 1b); but the difference in mean probability of negotiation or coverage first increased and then decreased with the largest difference occurring when sum of attribute levels was 17 (the bell-shaped curve as presented in Figure 1d). This may suggest that: 1) the better a drug performs on value attributes, the more valuable it is for severe/critical disease than for mild/ moderate disease; 2) the better a drug performs on value attributes, the increase of its probability of negotiation or coverage was larger for severe or critical disease first and then larger for mild or moderate disease; and 3) for drugs with moderate degree of performance on value attributes, it is critical to take severity of disease into account when making coverage

decisions. However, it is important to note that Figures 1b and 1d only display the relationship between differences in the means of value scores and probabilities of negotiation or coverage versus the sum of attribute levels. The variance of the difference in value scores and probabilities of negotiation or coverage between disease scenarios was not accounted for.

Currently, the National Healthcare Security Administration (NHSA) in China updates the National Medical Insurance drug list annually.[36] Drugs that are newly approved or not included in the drug list previously may be covered, enter the negotiation, or be rejected by the NHSA based on deliberation of the drugs' performance on multiple attributes by a committee comprised of physicians, health economists and policymakers.[36] Although some attributes such as innovation, health equity and organizational impact are included in the appraisal, they are assessed qualitatively.[36] In addition, dependence between severity of disease and other attributes is implicitly considered in the decision-making process with certain diseases such as cancer, children's diseases and rare diseases being given priority.[47] Our VAF provides the first formal MCDA value assessment and decision support system for China . It has the potential to improve the transparency, efficiency, consistency and robustness in health value assessment and decision-making process in China.

Our study has some limitations. First, the number of value profiles (i.e., 240 out of 4^{10} value profiles) and respondents in the survey were relatively small. In addition, people from rural areas, with less education and aged 65 years and older were under-represented. Thus, the prediction precision of our framework could be affected. Second, the order of value profiles for each respondent was not randomized due to the difficulty of implementing it on the survey platform. Order effect, which means the respondents' responses are affected by the order of value profiles, might happen.[22] Third, we collapsed disease scenarios, attribute levels and decision categories due

to non-significant differentiations between some categories. Despite that we designed the survey to resemble real decision-making processes, the collapsing may indicate potential redundancy of these variables' categorization.

Conclusions

Using the methods of factorial survey, the scoring functions for the VAF in China was developed. Using the developed VAF. The functions can be used to estimate the value of drugs and their probabilities of entering negotiation or being approved for coverage under different disease scenarios in China.

References

1. Centers for Medicare and Medicaid Services. CMS' Value-Based Programs | CMS, https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/Value-Based-Programs/Value-Based-Programs (2021, accessed 16 January 2022).

2. Miller HD. From Volume To Value: Better Ways To Pay For Health Care. *Health Affairs* 2009; 28: 1418–1428.

3. National Library of Medicine. HTA 101: II. FUNDAMENTAL CONCEPTS, https://www.nlm.nih.gov/nichsr/hta101/ta10104.html#Heading1 (accessed 29 June 2020).

4. Neumann PJ, Willke RJ, Garrison LP. A Health Economics Approach to US Value Assessment Frameworks-Introduction: An ISPOR Special Task Force Report [1]. *Value Health* 2018; 21: 119–123.

5. Zhang M, Bao Y, Lang Y, et al. What Is Value in Health and Healthcare? A Systematic Literature Review of Value Assessment Frameworks. *Value Health* 2022; 25: 302–317.

6. Phelps CE, Lakdawalla DN, Basu A, et al. Approaches to Aggregation and Decision Making-A Health Economics Approach: An ISPOR Special Task Force Report [5]. *Value Health* 2018; 21: 146–154.

7. Angelis A. Evaluating the Benefits of New Drugs in Health Technology Assessment Using Multiple Criteria Decision Analysis: A Case Study on Metastatic Prostate Cancer With the Dental and Pharmaceuticals Benefits Agency (TLV) in Sweden. *MDM Policy & Practice* 2018; 3: 2381468318796218.

8. Angelis A, Kanavos P. Multiple Criteria Decision Analysis (MCDA) for evaluating new medicines in Health Technology Assessment and beyond: The Advance Value Framework. *Social Science & Medicine* 2017; 188: 137–156.

9. Asaria M, Griffin S, Cookson R. Distributional Cost-Effectiveness Analysis: A Tutorial. *Med Decis Making* 2016; 36: 8–19.

10. Lakdawalla DN, Phelps CE. Health technology assessment with risk aversion in health. *J Health Econ* 2020; 72: 102346.

11. Verguet S, Kim JJ, Jamison DT. Extended Cost-Effectiveness Analysis for Health Policy Assessment: A Tutorial. *Pharmacoeconomics* 2016; 34: 913–923.

12. Hansen PH, Devlin ND. Multi-Criteria Decision Analysis (MCDA) in Healthcare Decision-Making. In: *The Oxford Encyclopedia of Health Economics*. Oxford University Press, http://www.oxfordreference.com/view/10.1093/acref/9780190902179.001.0001/acref-9780190902179-e-98 (2020, accessed 6 March 2022).

13. Diaby V, Campbell K, Goeree R. Multi-criteria decision analysis (MCDA) in health care: A bibliometric analysis. *Operations Research for Health Care* 2013; 2: 20–24.

14. Angelis A, Linch M, Montibeller G, et al. Multiple Criteria Decision Analysis for HTA across four EU Member States: Piloting the Advance Value Framework. *Soc Sci Med* 2020; 246: 112595.

15. Institute for Quality and Efficiency in Health Care (IQWiG). General Methods Version 6.0, https://www.iqwig.de/en/about-us/methods/methods-paper/ (2020).

16. Jehu-Appiah C, Baltussen R, Acquah C, et al. Balancing equity and efficiency in health priorities in Ghana: the use of multicriteria decision analysis. *Value Health* 2008; 11: 1081–1087.

17. Marsh K, Lanitis T, Neasham D, et al. Assessing the value of healthcare interventions using multi-criteria decision analysis: a review of the literature. *Pharmacoeconomics* 2014; 32: 345–365.

18. Radaelli G, Lettieri E, Masella C, et al. Implementation of EUnetHTA core Model® in Lombardia: the VTS framework. *Int J Technol Assess Health Care* 2014; 30: 105–112.

19. Marsh KD, Sculpher M, Caro JJ, et al. The Use of MCDA in HTA: Great Potential, but More Effort Needed. *Value in Health* 2018; 21: 394–397.

20. Howard R. Decision Analysis: Applied Decision Theory. *SDG*, https://sdg.com/publications/decision-analysis-applied-decision-theory/ (1966, accessed 27 September 2022).

21. Zhang M, Bao Y, Yang Y, et al. Identifying Attributes for a Value Assessment Framework in China: A Qualitative Study. *Pharmacoeconomics* 2023; 1–17.

22. Auspurg K, Hinz T. *Factorial Survey Experiments*. SAGE Publications, Inc, https://us.sagepub.com/en-us/nam/factorial-survey-experiments/book240309 (2015, accessed 11 May 2022).

23. Wallander L. 25 years of factorial surveys in sociology: A review. *Social Science Research* 2009; 38: 505–520.

24. Tata V, Varisco TJ, Du N, et al. Clinical decision making involving prescription drug monitoring programs: A factorial, vignette-based study among student pharmacists. *J Am Pharm Assoc (2003)* 2021; 61: 316–324.

25. Beyer H, Liebe U. Three experimental approaches to measure the social context dependence of prejudice communication and discriminatory behavior. *Soc Sci Res* 2015; 49: 343–355.

26. Badia X, Aguarón A, Fernández A, et al. Patient involvement in reflective multicriteria decision analysis to assist decision making in oncology. *Int J Technol Assess Health Care* 2019; 35: 56–63.

27. Choi SE, Berkowitz SA, Yudkin JS, et al. Personalizing Second-Line Type 2 Diabetes Treatment Selection: Combining Network Meta-analysis, Individualized Risk, and Patient Preferences for Unified Decision Support. *Med Decis Making* 2019; 39: 239–252.

28. Anderson P, Webb P, Groves S. Prioritisation of specialist health care services; not NICE, not easy but it can be done. *Health Policy* 2017; 121: 978–985.

29. Winquist E, Bell CM, Clarke JTR, et al. An evaluation framework for funding drugs for rare diseases. *Value Health* 2012; 15: 982–986.

30. Guyatt GH, Oxman AD, Vist GE, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ* 2008; 336: 924–926.

31. Antony J. 7 - Fractional Factorial Designs. In: Antony J (ed) *Design of Experiments for Engineers and Scientists (Second Edition)*. Oxford: Elsevier, pp. 87–112.

32. Wheeler B. AlgDesign: Algorithmic Experimental Design, https://CRAN.R-project.org/package=AlgDesign (2022, accessed 27 September 2022).

33. Wikipedia. WeChat. *Wikipedia*, https://en.wikipedia.org/w/index.php?title=WeChat&oldid=1112307059 (2022, accessed 27 September 2022).

34. Crouse T, Lowe PA. The SAGE Encyclopedia of Educational Research, Measurement, and Evaluation. Thousand Oaks,: SAGE Publications, Inc. Epub ahead of print 27 September 2022. DOI: 10.4135/9781506326139.

35. Mei B, Brown GTL. Conducting Online Surveys in China. *Social Science Computer Review* 2018; 36: 721–734.

36. National Healthcare Security Administration (NHSA). Work Plan for the Adjustment of the National Medical Insurance Drug List in 2022, http://www.nhsa.gov.cn/art/2021/6/30/art_62_5386.html (2022, accessed 20 January 2023).

37. Baillie M, Cessie S le, Schmidt CO, et al. Ten simple rules for initial data analysis. *PLOS Computational Biology* 2022; 18: e1009819.

38. R Core Team. R: A language and environment for statistical computing., https://www.R-project.org/ (2020).

39. Iskrov G, Miteva-Katrandzhieva T, Stefanov R. Multi-Criteria Decision Analysis for Assessment and Appraisal of Orphan Drugs. *Front Public Health* 2016; 4: 214.

40. Dankó D, Molnár MP. Balanced assessment systems revisited. *J Mark Access Health Policy* 2017; 5: 1355190.

41. Krahn M, Miller F, Bayoumi A, et al. DEVELOPMENT OF THE ONTARIO DECISION FRAMEWORK: A VALUES BASED FRAMEWORK FOR HEALTH TECHNOLOGY ASSESSMENT. *Int J Technol Assess Health Care* 2018; 34: 290–299.

42. Guarga L, Badia X, Obach M, et al. Implementing reflective multicriteria decision analysis (MCDA) to assess orphan drugs value in the Catalan Health Service (CatSalut). *Orphanet J Rare Dis* 2019; 14: 157.

43. EVIDEM. Evidence and Value: Impact on DEcision Making (EVIDEM) 10th Edition, https://www.evidem.org/ (2017, accessed 11 February 2020).

44. Skedgel C, Henderson N, Towse A, et al. Considering Severity in Health Technology Assessment: Can We Do Better? *Value Health* 2022; 25: 1399–1403.

45. Nord E, Daniels N, Kamlet M. QALYs: Some Challenges. *Value in Health* 2009; 12: S10–S15.

46. Shah KK. Severity of illness and priority setting in healthcare: a review of the literature. *Health Policy* 2009; 93: 77–84.

47. Si L, Xu L, Chen M, et al. Using strategic price negotiations to contain costs and expand access to medicines in China. *BMJ Global Health* 2020; 5: e002256.

List of figures and tables

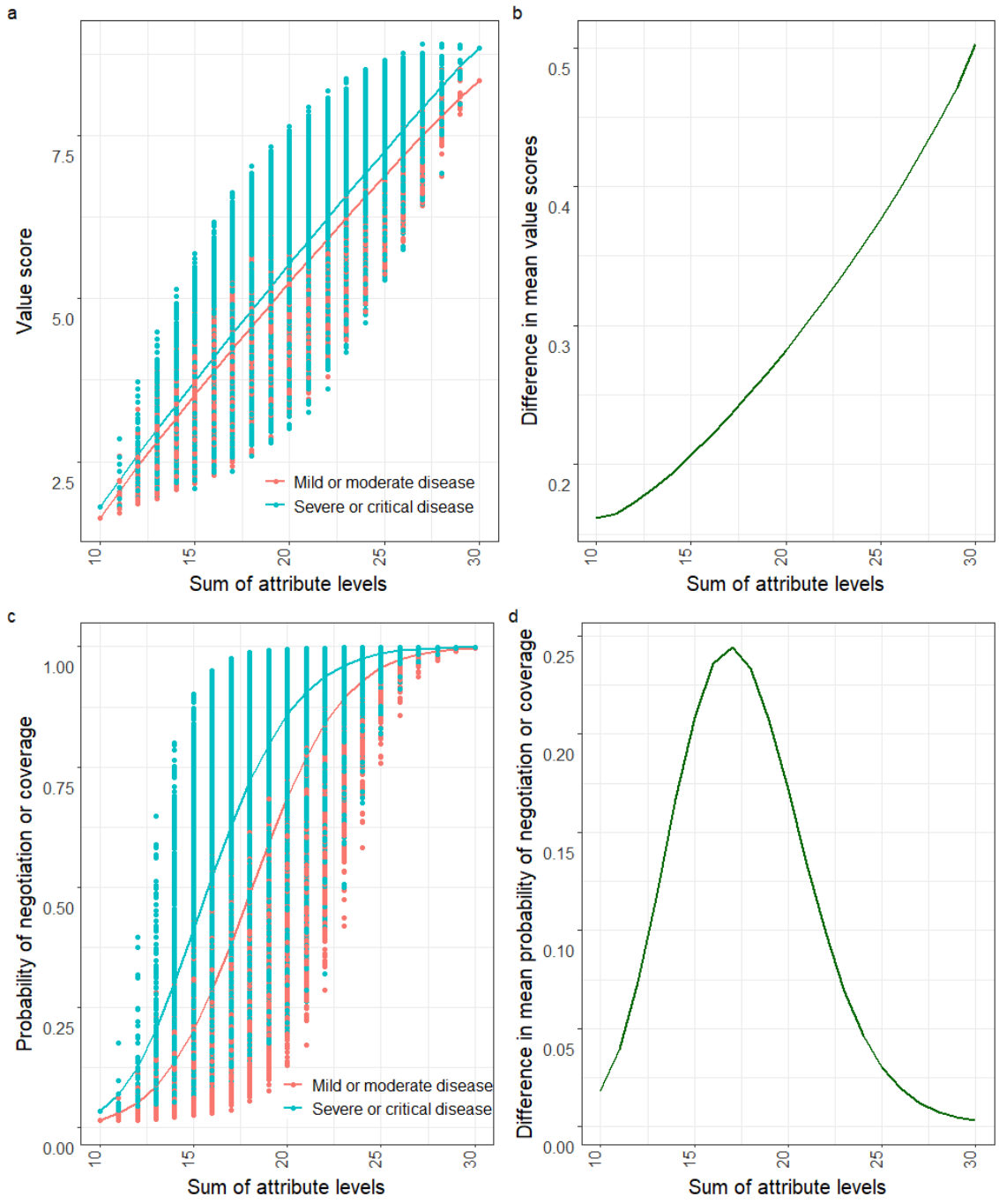


Figure 1. Relationship between the sum of attribute levels and a health technology's value score and probability of negotiation or coverage

Table 1. Characteristics of respondents in the survey

Characteristic	All respondents (n = 365)	Respondents included in the analysis (n = 331)	Respondents excluded from the analysis (n = 34)	P-value
Age, n (%)				0.005
18 – 29 years	118 (32.3%)	113 (34.1%)	5 (14.7%)	
30 - 44 years	158 (43.3%)	143 (43.2%)	15 (44.1%)	
45 – 64 years	81 (22.2%)	70 (21.1%)	11 (32.4%)	
65 years and above	8 (2.2%)	5 (1.5%)	3 (8.8%)	
Male, n (%)	138 (37.8%)	123 (37.2%)	15 (44.1%)	0.46
Rural area, n (%)	46 (12.6%)	35 (10.6%)	11 (32.4%)	0.001
Regions, n (%)				<0.001
North China	131 (35.9%)	123 (37.2%)	8 (23.5%)	
East China	113 (31.0%)	105 (31.7%)	8 (23.5%)	
Middle and South China	52 (14.2%)	36 (10.9%)	16 (47.1%)	
Southwest China	38 (10.4%)	36 (10.9%)	2 (5.9%)	
Northwest China	13 (3.6%)	13 (3.9%)	0	
Northeast China	12 (3.3%)	12 (3.6%)	0	
Other regions	6 (1.6%)	6 (1.8%)	0	
Stakeholder groups, n (%)				<0.001
The public	110 (30.1%)	90 (27.2%)	20 (58.8%)	
Policymaker	12 (3.3%)	12 (3.6%)	0	
Healthcare provider	132 (36.2%)	119 (36.0%)	13 (38.2%)	
HTA researcher	63 (17.3%)	63 (19.0%)	0	
Industry representative	48 (13.1%)	47 (14.2%)	1 (2.9%)	
Highest education, n (%)				<0.001
Middle school and under	3 (0.8%)	1 (0.3%)	2 (5.9%)	
High school and equivalent	25 (6.8%)	14 (4.2%)	11 (32.4%)	
College or equivalent	17 (4.7%)	13 (3.9%)	4 (11.8%)	
Bachelor's degree	118 (32.3%)	107 (32.3%)	11(32.4%)	
Graduate degree	202 (55.3%)	196 (59.2%)	6 (17.6%)	
Insurance type of the public, n (%)				0.90
Urban employee basic medical insurance	58 (52.7%)	48 (53.3%)	10 (50%)	
Urban-rural resident basic medical insurance	45 (40.9%)	36 (40%)	9 (45%)	
Other	7 (6.4%)	6 (6.7%)	1 (5%)	
Level of administration among policymakers, n (%)				NA
National or provincial administration	2 (16.7%)	2 (16.7%)	0	
Hospital administration	10 (83.3%)	10 (83.3%)	0	
Type of healthcare providers, n (%)				0.52
Physician	25 (18.9%)	21 (17.6%)	4 (30.8%)	
Nurse	18 (13.6%)	15 (12.6%)	3 (23.1%)	
Pharmacist	84 (63.6%)	78 (65.5%)	6 (46.2%)	
Other	5 (3.8%)	5 (4.2%)	0	

Level of hospitals among healthcare providers, n (%)				0.30
Tertiary	85 (64.4%)	78 (65.5%)	7 (53.8%)	
Secondary	39 (29.5%)	34 (28.6%)	5 (38.5%)	
Primary	2 (1.5%)	1 (0.8%)	1 (7.7%)	
Other	1 (0.8%)	1 (0.8%)	0	
Unknown	5 (3.8%)	5 (4.2%)	0	
Type of HTA researchers' affiliations, n (%)				NA
Universities / academic institutions	42 (66.7%)	42 (66.7%)	0	
Consulting company	6 (9.5%)	6 (9.5%)	0	
Non-profit HTA organizations	9 (14.3%)	9 (14.3%)	0	
Other	3 (4.8%)	3 (4.8%)	0	
Unknown	3 (4.8%)	3 (4.8%)	0	
Work area of industry representatives, n (%)				0.82
Market access and reimbursement	11 (22.9%)	11 (23.4%)	0	
Medical affairs	7 (14.6%)	7 (14.9%)	0	
Health economics and outcomes research	19 (39.6%)	18 (38.3%)	1 (100%)	
Other	11 (22.9%)	11 (23.4%)	0	
Work experience in the health sector of stakeholders except the public, n (%)				0.23
1-3 years	75 (33.3%)	73 (30.3%)	2 (14.3%)	
4-6 years	47 (18.4%)	46 (19.1%)	1 (7.1%)	
7-10 years	31 (12.2%)	28 (11.6%)	3 (21.4%)	
10 years and more	102 (40%)	94 (39.0%)	8 (57.1%)	

Abbreviations: HTA: health technology assessment

Table 2. Coefficient estimates (95% CI) of the linear mixed-effects models in different disease scenarios

Value attributes	Mild/moderate disease	P value	Severe/critical disease	P value
Intercept	1.64 (1.26, 2.02)	<0.001	1.8 (1.46, 2.15)	<0.001
Survival				
Poor	Ref		Ref	
Fair or good	0.95 (0.66, 1.24)	<0.001	1.05 (0.79, 1.32)	<0.001
Excellent	1.48 (1.28, 1.68)	<0.001	1.92 (1.74, 2.1)	<0.001
Clinical outcomes (excluding survival)				
Poor	Ref		Ref	
Fair or good	0.72 (0.46, 0.99)	<0.001	0.77 (0.52, 1.01)	<0.001
Excellent	1.12 (0.91, 1.33)	<0.001	1.17 (0.98, 1.36)	<0.001
Patient-reported outcomes				
Poor	Ref		Ref	
Fair or good	0.57 (0.31, 0.84)	<0.001	0.66 (0.41, 0.9)	<0.001
Excellent	0.95 (0.75, 1.15)	<0.001	1.13 (0.95, 1.32)	<0.001
Safety				
Poor	Ref		Ref	
Fair or good	0.55 (0.28, 0.82)	<0.001	0.55 (0.31, 0.8)	<0.001
Excellent	0.99 (0.79, 1.2)	<0.001	1 (0.81, 1.19)	<0.001
Budget impact to payer				
High	Ref		Ref	
Moderate or low	0.14 (-0.13, 0.41)	0.32	0.14 (-0.11, 0.39)	0.27
None	0.34 (0.13, 0.54)	<0.001	0.15 (-0.04, 0.33)	0.13
Out-of-pocket costs to patients				
High	Ref		Ref	
Moderate or low	0.22 (-0.08, 0.51)	0.15	0.02 (-0.25, 0.29)	0.88
None	0.23 (0.02, 0.43)	0.03	0.25 (0.06, 0.44)	0.01
Cost-effectiveness				
Poor	Ref		Ref	
Fair or good	0.14 (-0.16, 0.43)	0.37	0.24 (-0.03, 0.51)	0.08
Excellent	0.39 (0.19, 0.59)	<0.001	0.43 (0.25, 0.61)	<0.001

Innovation				
Poor	Ref		Ref	
Fair or good	0.34 (0.08, 0.6)	0.01	0.24 (0, 0.48)	0.05
Excellent	0.58 (0.37, 0.79)	<0.001	0.61 (0.42, 0.8)	<0.001
Organizational impact				
High	Ref		Ref	
Moderate or low	0.38 (0.11, 0.65)	0.01	0.13 (-0.12, 0.38)	0.3
Excellent	0.22 (0.02, 0.42)	0.03	0.12 (-0.06, 0.31)	0.19
Health equity				
Poor	Ref		Ref	
Fair or good	0.08 (-0.22, 0.37)	0.62	0.31 (0.04, 0.58)	0.03
Excellent	0.42 (0.21, 0.62)	<0.001	0.26 (0.07, 0.45)	0.01

Abbreviations: CI: confidence interval.

Table 3. Odds ratios estimates (95% CI) of the logistic mixed-effects models in different disease scenarios

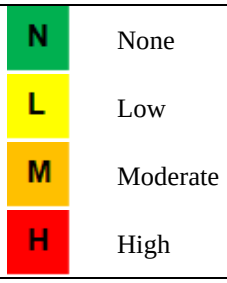
Value attributes	Mild/moderate disease	P value	Severe/critical disease	P value
Intercept	0.02 (0.01, 0.04)	<0.001	0.04 (0.02, 0.08)	<0.001
Survival				
Poor	Ref		Ref	
Fair or good	4.06 (2.47, 6.68)	<0.001	6.08 (3.45, 10.73)	<0.001
Excellent	7.57 (5.08, 11.27)	<0.001	17.09 (10.01, 29.19)	<0.001
Clinical outcomes (excluding survival)				
Poor	Ref		Ref	
Fair or good	2.45 (1.55, 3.87)	<0.001	3.07 (1.81, 5.2)	<0.001
Excellent	5.47 (3.64, 8.22)	<0.001	6.07 (3.77, 9.78)	<0.001
Patient-reported outcomes				
Poor	Ref		Ref	
Fair or good	2.43 (1.53, 3.87)	<0.001	2.12 (1.27, 3.55)	0.004
Excellent	4.36 (2.95, 6.44)	<0.001	4.43 (2.81, 7)	<0.001
Safety				
Poor	Ref		Ref	
Fair or good	1.93 (1.22, 3.05)	0.005	2.18 (1.28, 3.71)	0.004
Excellent	3.46 (2.37, 5.08)	<0.001	4.91 (3.11, 7.77)	<0.001
Budget impact to payer				
High	Ref		Ref	
Moderate or low	1.53 (0.96, 2.44)	<0.001	1.53 (0.88, 2.66)	0.13
None	2.13 (1.47, 3.1)	<0.001	1.46 (0.96, 2.21)	0.07
Out-of-pocket costs to patients				
High	Ref		Ref	
Moderate or low	0.99 (0.6, 1.65)	0.98	1.32 (0.73, 2.38)	0.36
None	1.21 (0.84, 1.74)	0.31	1.54 (1.01, 2.36)	0.05
Cost-effectiveness				
Poor	Ref		Ref	
Fair or good	0.91 (0.55, 1.49)	0.70	1.25 (0.7, 2.24)	0.45
Excellent	1.75 (1.23, 2.5)	0.002	1.79 (1.2, 2.69)	0.005

Innovation				
Poor	Ref		Ref	
Fair or good	1.84 (1.18, 2.88)	0.007	1.53 (0.92, 2.53)	0.10
Excellent	2.49 (1.7, 3.66)	<0.001	2.03 (1.3, 3.15)	0.002
Organizational impact				
High	Ref		Ref	
Moderate or low	1.39 (0.87, 2.21)	0.17	1.09 (0.64, 1.87)	0.74
Excellent	1.46 (1.02, 2.09)	0.04	1.21 (0.8, 1.84)	0.37
Health equity				
Poor	Ref		Ref	
Fair or good	1.35 (0.83, 2.2)	0.23	1.55 (0.89, 2.71)	0.12
Excellent	1.51 (1.05, 2.18)	0.03	1.35 (0.88, 2.06)	0.17

Appendices

Appendix 1:

Table S1. List of value attributes included in the value assessment framework

Categories	Value Attributes	Response level	Descriptions
Severity of disease	Severity of disease	○ Life threatening or critical ○ Severe ○ Moderate or mild	- Life threatening/critical refers to diseases or conditions with immediate risk of death - Severe refers to diseases or conditions where treatment is needed, otherwise there will be risk of death - Moderate/mild refers to non-life-threatening diseases and may need care at outpatient or community level
Health benefit	Survival†		E: Substantially longer survival/better outcomes than comparator G: Moderately longer survival/better outcomes than comparator F: Similar or non-inferior to comparator P: Substantially shorter survival/worse outcomes than comparator
	Clinical outcomes (excluding survival) †		
	Patient-reported outcomes†		
Safety	Safety†		
Economic impact	Budget impact to payer		N: No impact/cost L: Low impact/cost (less than the per capita disposable income, ¥35,128 in 2021 in China)[52] M: Moderate impact/cost (similar to the per capita disposable income, ¥35,128 in 2021 in China)[52] H: High impact/cost (higher than the per capita disposable income, ¥35,128 in 2021 in China)[52]
	Out-of-pocket costs to patients		
	Cost-effectiveness†		E: < ¥ 80,000 per QALY, i.e., smaller than China's per-capita GDP in 2021‡ G: < ¥ 160,000 per QALY, i.e., smaller than two times China's per-capita GDP in 2021 F: < ¥ 240,000 per QALY, smaller than three times China's per-capita GDP in 2021 P: > ¥ 240,000 per QALY, i.e., larger than three times China's per-capita GDP in 2021
Innovation	Innovation in addressing unmet needs beyond health benefit, safety, and costs†		E: Offering the first of its kind or the only option for a disease's diagnosis, prevention, or treatment G: Offering an improved option for a disease's diagnosis, prevention, or treatment (e.g., route, frequency, duration, and place of administration)

			F: Offering an option similar to currently available alternatives for a disease's diagnosis, prevention, or treatment P: Offering an option worse than currently available alternatives for a disease's diagnosis, prevention, or treatment (e.g., route, frequency, duration, and place of administration)
Organizational impact	Level of impact on the health system, healthcare facilities and healthcare providers (e.g., space, administration, personnel, and training)	N None L Low M Moderate H High	**N:** None or negligible impact **L:** Low impact **M:** Moderate impact **H:** High impact
Health equity	Improvement of health equity across populations (e.g., different race, age, gender, socio-economic status, or regions) †	E Excellent G Good F Fair P Poor	**E:** The new technology delivers much more benefits to socioeconomically disadvantaged patients **G:** The new technology delivers slightly more benefits to socioeconomically disadvantaged patients **F:** The new technology brings similar benefits to all patients (status quo) **P:** The new technology delivers more benefits to socioeconomically advantaged patients
Quality of evidence	Quality of evidence	⊕⊕⊕⊕ High ⊕⊕⊕◯ Moderate ⊕⊕◯◯ Low ⊕◯◯◯ Very low	**High:** We are very confident in the conclusion **Moderate:** We are moderately confident in the conclusion: the conclusion is likely to be true **Low:** Our confidence in the conclusion is limited: the conclusion may be true **Very low**: We have very little confidence in the conclusion: the conclusion is unlikely to be true

Abbreviations: HTA: health technology assessment; GDP: gross domestic product; QALY: quality-adjusted life year.

†: Comparators should be clearly stated for this attribute.

‡: China's GDP per capita was ¥ 80,976 in 2021. The value of GDP per capita should be updated to the GDP for the year when using the VAF.

Appendix 2: An example of the disease scenario and the hypothetical drug value profile

Suppose that we have a new drug that is for the treatment of **a severe disease (e.g., chronic obstructive pulmonary disease)**. The drug has the following characteristics:

Attributes	Performance	Description
Survival	E	The new drug provides substantially longer survival than the comparator
Clinical outcomes	E	The new drug delivers substantially better clinical outcomes (e.g., blood pressure) than the comparator
Patient-reported outcomes	E	The new drug delivers substantially better patient-reported outcomes (e.g., quality of life) than the comparator
Safety	E	The new drug is substantially safer than the comparator
Budget impact to payer	H	The budget impact to payer with the reimbursement of the new drug is high (both the size of population and the annual costs per capita considered)
Out-of-pocket costs to patients	N	There are no out-of-pocket costs to patients with the reimbursement of the new drug
Cost-effectiveness	G	The incremental cost-effectiveness ratio of the new drug vs the comparator is lower than 160 thousand RMB
Innovation	G	The new drug offers an improved option for the disease's diagnosis, prevention, or treatment (e.g., lower administration, better ease of use)
Organizational impact	N	The new drug has none or negligible impact on the health system (e.g., no extra space required in the hospital, no extra personnel training)
Health equity	G	The new technology delivers slightly more benefits to socioeconomically disadvantaged patients vs advantaged patients than the comparator

Key:

E	Excellent	G	Good	F	Fair	P	Poor
N	None	L	Low	M	Moderate	H	High

Please note that the drug's performance on each attribute has incorporated both the drugs' effect size and quality of evidence for that attribute.

From your perspective, the value of the drug is:

Low value					Fair					High value
0	1	2	3	4	5	6	7	8	9	10
☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

The drug should be:

☐ Covered by national medical insurance

☐ Negotiated for coverage by national medical insurance

☐ Not covered by national medical insurance

Appendix 3:

The design and construction of the survey

Informed by the MCDA/MCDM methods, the factorial experiment methods, and our study results in the first stage of our program of study, we conducted an online survey among healthcare stakeholders in China to derive weights for attributes except severity of disease and quality of evidence.[12,22,35] The following section describes the design and administration process of our survey.

1. Summary of scenarios, attributes, and response levels in the survey

As described in the manuscript, we constructed three disease scenarios in the survey, i.e., 1) immediately life-threatening or critical diseases or conditions such as acute stroke, 2) severe diseases with no immediate risk of dying but could be life threatening if left untreated such as chronic obstructive pulmonary disease, and 3) moderate or mild diseases such as mild hyperlipidemia. Since the performance of the hypothetical health technology on each attribute was modified by the quality of evidence, quality of evidence was not included in the value profile. Thus, there were 10 attributes in each value profile including survival, clinical outcomes, patient-reported outcomes, safety, budget impact to payer, out-of-pocket costs to patients, cost-effectiveness, innovation, organizational impact, and health equity. Each of these 10 attributes has 4 response levels (see Appendix 1). Response levels of the attribute were coded using the numbers 1-4 with 4 representing excellent or none and 1 representing poor or high (Appendix 8). Figure S1 shows the attributes, response levels, coded levels, and corresponding descriptions.

Attribute	Response level	Coded level	Descriptions
survival	Excellent	4	Substantially longer survival than comparator
survival	Good	3	Moderately longer survival than comparator
survival	Fair	2	Similar or non-inferior to comparator with respect to survival
survival	Poor	1	Substantially shorter survival than comparator
clinical_outcome	Excellent	4	Substantially better than comparator
clinical_outcome	Good	3	Moderately better than comparator
clinical_outcome	Fair	2	Similar or non-inferior to comparator
clinical_outcome	Poor	1	Substantially worse than comparator

Figure S1. Example of the attributes, coded response level and descriptions

2. A D-efficient sample of value profiles was selected from the value profile population

The total value profile population includes 4^{10} = 1,048,576 value profiles which are combinations of different response levels of the 10 attributes. Using the AlgDesign package in R, we searched a D-efficient sample from the value profile population.[35] D-efficient design provides both orthogonality (uncorrelation between any two attributes) and level balance (levels of single attributes occurs with the same frequency) in the sample.[24] Auspurg et al. found that for a factorial survey with 9 dimensions (attributes), a sample of 200 vignettes (in our case, value profiles) is needed to ensure that D-efficiency is higher than the value of 90.[35] Given that a D-efficiency of 80 is considered reasonable and that we have 10 attributes in each scenario, we drew a sample of 240 value profiles from the value profile population. Figure S2 shows an

example of the value profile sample and blocking. For example, in value profile No. 49, the response level of survival is 4 which means the new technology is substantially better than the comparator with respect to survival prolongation.

ids | **Attributes and coded levels in value profiles**

block_id	value_profile_id	survival	clinical_outcome	patient_reported_outco	safety	costs to payer	costs to patients	cost-effectiveness	innovation	organizational impact	health equity
1	49	4	4	4	3	1	1	2	4	1	3
1	244	2	4	4	4	3	2	2	3	3	2
1	879	2	3	3	4	3	4	4	3	3	3
1	12435	2	2	2	3	3	4	2	2	3	2
2	...										

Figure S2. An example of the value profile sample

3. Partitioning the value profile samples into value profile blocks

Previous research has shown that no more than 10 vignettes should be used per respondent to avoid fatigue and learning effect.[35] Thus, we partitioned the value profile sample into 60 blocks with 4 value profiles in each block and ask each respondent to complete 1 block for all 3 disease scenarios (each respondent will complete 12 value profiles). We used the AlgDesign package in R to complete the partitioning. Figure S2 shows an example of the value profile sample and blocking. For example, there are four value profiles in block No. 1 which include value profile No. 49, 244, 879 and 12435.

4. Random assignment of the value profile blocks to respondents

The blocks of value profiles were assigned to respondents randomly on the platform of Wen Juan Xing. The blocks were distributed on the "first come, first serve" principle where the first (nth) respondent will receive the first (nth) block (Figure S3). Figure S3 shows an example of the random assignment of the value profile blocks to respondents. For example, the second respondent will receive the second block which is block No. 10 that includes value profiles No. 38235, 79876, 96969 and 102344.

ids and order of blocks | **Attributes and levels within 1st value profile** | **Attributes and levels within 2nd value profile**

participant_id	block_order	block_id	1st_profile_id	survival_1	clinical_outcome_1	...	2nd_profile_id	survival_2	clinical_outcome_2	...	3rd_profile_id	...	4th_profile_id	...
1	1	1	49	4	4	...	244	2	4	...	879	...	12435	...
2	2	10	38235	3	3	...	79876	2	3	...	96969	...	102344	...
3	3	30	117	1	2	...	567	3	4	...	65432	...	112344	...
4	4	1	244	...	...	...	12435	...	...	...	879	...	49	...
5	5	4	...	...	...	...	...	...	...	...	...	...	...	...

Figure S3. Block assignment and data collection form showing random assignment of the value profile blocks to respondents

5. Presentation of the value profile in a tabular manner

The value profile was presented in a tabular manner as in Figure 1 in the manuscript. The tabular presentation would be more straightforward than text value profiles where the information is

presented in short text story. We drafted the fixed description texts for all respondents, translate the numeric code into the descriptions and definitions for each response level of the attributes and insert the descriptions for the corresponding attribute levels into the fixed description texts. Figure S4 illustrates the process of how we constructed and presented the value profiles for each respondent. Appendix 2 presents an example of the value profile.

6. Additional questions in the pilot survey:

The following questions were added to the pilot survey to understand participants' perceptions about the survey.

Please indicate your level of agreement of the following statement based on your experience completing the task:

Statement	*Level of agreement*
I can understand my task in this survey	□ *Strongly disagree* □ *Disagree* □ *Neutral* □ *Agree* □ *Strongly agree*
I can easily answer the questions	□ *Strongly disagree* □ *Disagree* □ *Neutral* □ *Agree* □ *Strongly agree*

If you have any comments, please provide below:

_____

block_id	value_profile_id	survival	clinical_outcome	patient_reported_outco	safety	costs to payer
1	49	4	4	4	3	1
1	244	2	4	4	4	3
1	879	2	3	3	4	3
1	12435	2	2	2	3	3
2	...					

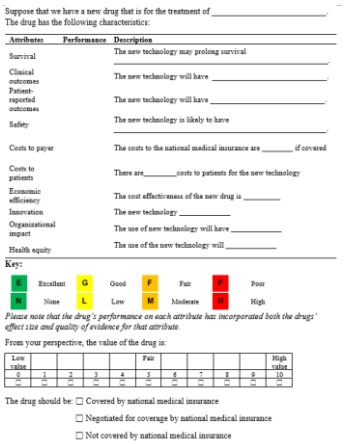
A D-efficient sample of the value profiles will be selected, and the selected 240 value profiles will be partitioned into 60 blocks using computer algorithms

participant_id	block_order	block_id	1st_profile_id	survival_1	clinical_outcome_1	...	2nd_prc
1	1	1	49	4	4	...	
2	2	10	38235	3	3	...	
3	3	30	117	1	2	...	
4	4	1	244	...	...	...	
5	5	4	...	...	...	...	...

The 60 blocks will be randomly assigned to respondents. The nth respondent will receive the nth block.

participant_id	block_order	block_id	1st_profile_id	survival_1	clinical_outcome_1	...	2nd_prc
1	1	1	49	4	4	...	
2	2	10	38235	3	3	...	
3	3	30	117	1	2	...	
4	4	1	244	...	...	...	
5	5	4	...	...	...	...	...

Attribute	Response level	Coded level	Descriptions
survival	Excellent	4	Substantially longer survival than comparator
survival	Good	3	Moderately longer survival than comparator
survival	Fair	2	Similar or non-inferior to comparator with respect to survival
survival	Poor	1	Substantially shorter survival than comparator
clinical_outcome	Excellent	4	Substantially better than comparator
clinical_outcome	Good	3	Moderately better than comparator
clinical_outcome	Fair	2	Similar or non-inferior to comparator
clinical_outcome	Poor	1	Substantially worse than comparator


Suppose that we have a new drug that is for the treatment of _____.
The drug has the following characteristics:

Attributes	Performance	Description
Survival		The new technology may prolong survival _____.
Clinical outcomes		The new technology will have _____.
Patient-reported outcomes		The new technology will have _____.
Safety		The new technology is likely to have _____.
Costs to payer		The costs to the national medical insurance are _____ if covered
Costs to patients		There are_____costs to patients for the new technology
Economic efficiency		The cost effectiveness of the new drug is _____
Innovation		The new technology _____
Organizational impact		The use of new technology will have _____
Health equity		The use of the new technology will _____

Key:
E Excellent G Good F Fair P Poor
N None L Low M Moderate H High

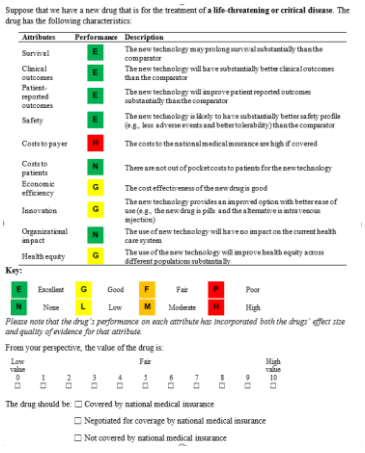
Please note that the drug's performance on each attribute has incorporated both the drugs' effect size and quality of evidence for that attribute.

From your perspective, the value of the drug is:

Low value					Fair					High value
0 ☐	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐	6 ☐	7 ☐	8 ☐	9 ☐	10 ☐

The drug should be: ☐ Covered by national medical insurance
☐ Negotiated for coverage by national medical insurance
☐ Not covered by national medical insurance

The numeric code in the block assignment form will be translated into the descriptions and definitions for each response level of the attributes. Then the descriptions for the corresponding attribute levels will be inserted into the fixed description texts.


Suppose that we have a new drug that is for the treatment of **a life-threatening or critical disease**. The drug has the following characteristics:

Attributes	Performance	Description
Survival	E	The new technology may prolong survival substantially than the comparator
Clinical outcomes	E	The new technology will have substantially better clinical outcomes than the comparator
Patient-reported outcomes	E	The new technology will improve patient reported outcomes substantially than the comparator
Safety	E	The new technology is likely to have substantially better safety profile (e.g., less adverse events and better tolerability) than the comparator
Costs to payer	H	The costs to the national medical insurance are high if covered
Costs to patients	N	There are not out of pocket costs to patients for the new technology
Economic efficiency	G	The cost effectiveness of the new drug is good
Innovation	G	The new technology provides an improved option with better ease of use (e.g., the new drug is pills and the alternative is intravenous injection)
Organizational impact	N	The use of new technology will have no impact on the current health care system
Health equity	G	The use of the new technology will improve health equity across different populations substantially

Key:
E Excellent G Good F Fair P Poor
N None L Low M Moderate H High

Please note that the drug's performance on each attribute has incorporated both the drugs' effect size and quality of evidence for that attribute.

From your perspective, the value of the drug is:

Low value					Fair					High value
0 ☐	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐	6 ☐	7 ☐	8 ☐	9 ☐	10 ☐

The drug should be: ☐ Covered by national medical insurance
☐ Negotiated for coverage by national medical insurance
☐ Not covered by national medical insurance

Figure S4. The process of constructing the value profiles

Appendix 4:

Table S2. Distribution of survey participants

Stakeholder group	Characteristics of participants
Policy makers	- National /provincial / municipal / hospital - 18 – 29 years old / 30 – 44 years old / 45 – 64 years old / over 65 years old - Male / female
Health care providers	- Urban / rural - Physicians / pharmacists / nurses - 18 – 29 years old / 30 – 44 years old / 45 – 64 years old / over 65 years old - Male / female
Academic researchers	- University or academic institute / consulting company or non-government organization - 18 – 29 years old / 30 – 44 years old / 45 – 64 years old / over 65 years old - Male / female
Industry representatives	- Insurance coverage / market access / other departments - Drugs / medical device and diagnostics - 18 – 29 years old / 30 – 44 years old / 45 – 64 years old / over 65 years old - Male / female
Members of the public	- Urban / rural - College or above / high school or below 18 – 29 years old / 30 – 44 years old / 45 – 64 years old / over 65 years old - Male / female

Appendix 5: Demographic questionnaires

14. In which province and city do you live?

    _____

15. Do you work in the health sector?
    C. Yes
    D. No
    If A is selected, go to Question 3; If B is selected, go to question 9-12.

16. What is your occupation in the health sector?
    A. Health care policy makers
    B. Health care providers
    C. Health economics or health technology assessment researchers
    D. Pharmaceutical or medical device company
    E. Other, please specify: _____.

    If A is selected, go to Question 4
    If B is selected, go to Question 5
    If C is selected, go to Question 6
    If D is selected, go to Question 7
    If E is selected, go to Question 9-12

17. What is the level of your department?
    E. National
    F. Provincial
    G. Regional
    H. Other, please specify: _____.

    Go to Question 10-14

18. What do you do as a health care provider?
    F. Physician
    G. Surgeon
    H. Pharmacist
    I. Nurse
    J. Other, please specify: _____.

    Go to Question 10-14

19. Where do you work as a health economics or health technology assessment researcher?
    E. University or academic institute
    F. Consulting company
    G. Non-government organization
    H. Other, please specify:_____.

Go to Question 10-14

20. What type of health technologies is your job related to ?
    E. Drug
    F. Medical device
    G. Diagnostics
    H. Other, please specify:_____.

    Go to Question 8

21. What department or area do your work in?
    F. Market access
    G. Insurance coverage
    H. Medical affairs
    I. Health economic and outcome research
    J. Other, please specify:_____.

    Go to Question 10-14

22. Please indicate your occupation: _____.

23. Please indicate your gender: _____.

24. Which category below includes your age?
    A.18 - 29 B. 30 - 44 C. 45 - 64 D. 65 or above

25. What is the highest level of school you have completed or the highest degree you have received?
    A. Less than high school degree
    B. High school degree or equivalent
    C. Some college but no degree
    D. Bachelor's degree
    E. Graduate degree (Master or Doctor degree)

26. How long have you been involved in health technology decision making or health technology assessment ?
    A.1-3 years B. 4-6 years C. 7-10 years D. 10 years or above

27. What is your professional status?
    A. Junior B. Moderate C. Senior

Appendix 6:

Table S3. Distribution of the 60 blocks among respondents

Block No.	All respondents (n = 362)	Respondents included in the analysis (n = 328)	Respondents excluded from the analysis (n = 34)		Block No.	All respondents (n = 362)	Respondents included in the analysis (n = 328)	Respondents excluded from the analysis (n = 34)
1	7	6	1		31	5	4	1
2	4	4	0		32	3	2	1
3	5	5	0		33	5	5	0
4	9	9	0		34	11	10	1
5	6	5	1		35	10	10	0
6	5	5	0		36	7	6	1
7	7	6	1		37	4	4	0
8	4	3	1		38	5	5	0
9	7	7	0		39	7	5	2
10	2	2	0		40	7	7	0
11	4	2	2		41	6	4	2
12	11	10	1		42	11	11	0
13	6	6	0		43	2	1	1
14	7	7	0		44	8	7	1
15	4	4	0		45	6	5	1
16	5	5	0		46	4	4	0
17	7	7	0		47	9	8	1
18	9	8	1		48	7	6	1
19	7	5	2		49	12	8	4
20	3	3	0		50	3	3	0
21	4	3	1		51	6	5	1
22	3	2	1		52	8	6	2
23	8	8	0		53	3	3	0
24	9	8	1		54	5	5	0
25	7	7	0		55	2	2	0
26	5	5	0		56	8	8	0
27	5	5	0		57	4	4	0
28	7	6	1		58	8	9	0
29	6	6	0		59	9	9	0
30	1	1	0		60	6	6	0

Appendix 7:

Table S4. Coefficient estimates (95% CI) of the logistic mixed-effects models in different disease scenarios

Value attributes	Mild/moderate disease	P value	Severe/critical disease	P value
Intercept	-4.13 (-4.9, -3.35)	<0.001	-3.35 (-4.15, -2.55)	<0.001
Survival				
Poor	Ref		Ref	
Fair or good	1.4 (0.91, 1.9)	<0.001	1.81 (1.24, 2.37)	<0.001
Excellent	2.02 (1.62, 2.42)	<0.001	2.84 (2.3, 3.37)	<0.001
Clinical outcomes (excluding survival)				
Poor	Ref		Ref	
Fair or good	0.9 (0.44, 1.35)	<0.001	1.12 (0.59, 1.65)	<0.001
Excellent	1.7 (1.29, 2.11)	<0.001	1.80 (1.33, 2.28)	<0.001
Patient-reported outcomes				
Poor	Ref		Ref	
Fair or good	0.89 (0.43, 1.35)	<0.001	0.75 (0.24, 1.27)	0.004
Excellent	1.47 (1.08, 1.86)	<0.001	1.49 (1.03, 1.95)	<0.001
Safety				
Poor	Ref		Ref	
Fair or good	0.66 (0.2, 1.11)	0.005	0.78 (0.25, 1.31)	0.004
Excellent	1.24 (0.86, 1.62)	<0.001	1.59 (1.13, 2.05)	<0.001
Budget impact to payer				
High	Ref		Ref	
Moderate or low	0.43 (-0.04, 0.89)	<0.001	0.42 (-0.13, 0.98)	0.13
None	0.76 (0.38, 1.13)	<0.001	0.38 (-0.04, 0.79)	0.07
Out-of-pocket costs to patients				
High	Ref		Ref	
Moderate or low	-0.01 (-0.52, 0.5)	0.98	0.28 (-0.32, 0.87)	0.36
None	0.19 (-0.17, 0.55)	0.31	0.43 (0.01, 0.86)	0.05
Cost-effectiveness				

Poor	Ref		Ref	
Fair or good	-0.1 (-0.59, 0.4)	0.70	0.22 (-0.36, 0.81)	0.45
Excellent	0.56 (0.2, 0.91)	0.002	0.58 (0.18, 0.99)	0.005
Innovation				
Poor	Ref		Ref	
Fair or good	0.61 (0.16, 1.06)	0.007	0.42 (-0.08, 0.93)	0.10
Excellent	0.91 (0.53, 1.3)	<0.001	0.71 (0.26, 1.15)	0.002
Organizational impact				
High				
Moderate or low	0.33 (-0.14, 0.79)	0.17	0.09 (-0.45, 0.63)	0.74
Excellent	0.38 (0.02, 0.74)	0.04	0.19 (-0.22, 0.61)	0.37
Health equity				
Poor				
Fair or good	0.3 (-0.19, 0.79)	0.23	0.44 (-0.12, 1)	0.12
Excellent	0.41 (0.04, 0.78)	0.03	0.3 (-0.13, 0.72)	0.17

Appendix 8: Value scoring algorithm and an example of value score and probabilities estimation:

Abbreviations:

BI: budget-impact to payer; CE: cost-effectiveness; HE: health equity; OI: organizational impact; OOP: out-of-pocket costs to patients; PROs: patient-reported outcomes.

Based on the coefficient estimates in Table 2 and Table S4, the value scoring algorithms and logit functions for the two disease scenarios are:

1) **Mild/moderate disease**

Value_mild = Intercept + $\sum\beta*X$ =

$1.64 + 0*\text{Survival}_{\text{poor}} + 0.95*\text{Survival}_{\text{fair or good}} + 1.48*\text{Survival}_{\text{excellent}} + 0*\text{Clinical}_{\text{poor}} + 0.72*\text{Clinical}_{\text{fair or good}} + 1.12*\text{Clinical}_{\text{excellent}} + 0*\text{PROs}_{\text{poor}} + 0.57*\text{PROs}_{\text{fair or good}} + 0.95*\text{PROs}_{\text{excellent}} + 0*\text{Safety}_{\text{poor}} + 0.55*\text{Safety}_{\text{fair or good}} + 0.99*\text{Safety}_{\text{excellent}} + 0*\text{BI}_{\text{high}} + 0.14*\text{BI}_{\text{moderate or low}} + 0.34*\text{BI}_{\text{none}} + 0*\text{OOP}_{\text{high}} + 0.22*\text{OOP}_{\text{moderate or low}} + 0.23*\text{OOP}_{\text{none}} + 0*\text{CE}_{\text{poor}} + 0.14*\text{CE}_{\text{fair or good}} + 0.39*\text{CE}_{\text{excellent}} + 0*\text{Innovation}_{\text{poor}} + 0.34*\text{Innovation}_{\text{fair or good}} + 0.58*\text{Innovation}_{\text{excellent}} + 0*\text{OI}_{\text{high}} + 0.38*\text{OI}_{\text{moderate or low}} + 0.22*\text{OI}_{\text{none}} + 0*\text{HE}_{\text{poor}} + 0.08*\text{HE}_{\text{fair or good}} + 0.42*\text{HE}_{\text{excellent}}$

Logit (Y=1, negotiation or coverage)_mild = Intercept + $\sum\beta*X$ =

$-4.13 + 0*\text{Survival}_{\text{poor}} + 1.4*\text{Survival}_{\text{fair or good}} + 2.02*\text{Survival}_{\text{excellent}} + 0*\text{Clinical}_{\text{poor}} + 0.9*\text{Clinical}_{\text{fair or good}} + 1.7*\text{Clinical}_{\text{excellent}} + 0*\text{PROs}_{\text{poor}} + 0.89*\text{PROs}_{\text{fair or good}} + 1.47*\text{PROs}_{\text{excellent}} + 0*\text{Safety}_{\text{poor}} + 0.66*\text{Safety}_{\text{fair or good}} + 1.24*\text{Safety}_{\text{excellent}} + 0*\text{BI}_{\text{high}} + 0.43*\text{BI}_{\text{moderate or low}} + 0.76*\text{BI}_{\text{none}} + 0*\text{OOP}_{\text{high}} + (-0.01)*\text{OOP}_{\text{moderate or low}} + 0.19*\text{OOP}_{\text{none}} + 0*\text{CE}_{\text{poor}} + (-0.1)*\text{CE}_{\text{fair or good}} + 0.56*\text{CE}_{\text{excellent}} + 0*\text{Innovation}_{\text{poor}} + 0.61*\text{Innovation}_{\text{fair or good}} + 0.91*\text{Innovation}_{\text{excellent}} + 0*\text{OI}_{\text{high}} + 0.33*\text{OI}_{\text{moderate or low}} + 0.38*\text{OI}_{\text{none}} + 0*\text{HE}_{\text{poor}} + 0.3*\text{HE}_{\text{fair or good}} + 0.41*\text{HE}_{\text{excellent}}$

2) **Severe or critical disease**

Value_severe = Intercept + $\sum\beta*X$ =

$1.8 + 0*\text{Survival}_{\text{poor}} + 1.05*\text{Survival}_{\text{fair or good}} + 1.92*\text{Survival}_{\text{excellent}} + 0*\text{Clinical}_{\text{poor}} + 0.77*\text{Clinical}_{\text{fair or good}} + 1.17*\text{Clinical}_{\text{excellent}} + 0*\text{PROs}_{\text{poor}} + 0.66*\text{PROs}_{\text{fair or good}} + 1.13*\text{PROs}_{\text{excellent}} + 0*\text{Safety}_{\text{poor}} + 0.55*\text{Safety}_{\text{fair or good}} + 1*\text{Safety}_{\text{excellent}} + 0*\text{BI}_{\text{high}} + 0.14*\text{BI}_{\text{moderate or low}} + 0.15*\text{BI}_{\text{none}} + 0*\text{OOP}_{\text{high}} + 0.02*\text{OOP}_{\text{moderate or low}} + 0.25*\text{OOP}_{\text{none}} + 0*\text{CE}_{\text{poor}} + 0.24*\text{CE}_{\text{fair or good}} + 0.43*\text{CE}_{\text{excellent}} + 0*\text{Innovation}_{\text{poor}} + 0.24*\text{Innovation}_{\text{fair or good}} + 0.61*\text{Innovation}_{\text{excellent}} + 0*\text{OI}_{\text{high}} + 0.13*\text{OI}_{\text{moderate or low}} + 0.12*\text{OI}_{\text{none}} + 0*\text{HE}_{\text{poor}} + 0.31*\text{HE}_{\text{fair or good}} + 0.26*\text{HE}_{\text{excellent}}$

Logit (Y=1, negotiation or coverage)_severe = Intercept + $\sum\beta*X$ =

-3.35 + 0*Survival $_{poor}$ + 1.81*Survival $_{fair\ or\ good}$ + 2.84*Survival $_{excellent}$ + 0*Clinical $_{poor}$ + 1.12*Clinical $_{fair\ or\ good}$ + 1.8*Clinical $_{excellent}$ + 0*PROs $_{poor}$ + 0.75* PROs $_{fair\ or\ good}$ + 1.49* PROs $_{excellent}$ + 0*Safety $_{poor}$ + 0.78*Safety $_{fair\ or\ good}$ + 1.59*Safety $_{excellent}$ + 0*BI $_{high}$ + 0.42* BI $_{moderate\ or\ low}$ + 0.38*BI $_{none}$ + 0*OOP $_{high}$ + 0.28*OOP $_{moderate\ or\ low}$ + 0.43*OOP $_{none}$ + 0*CE $_{poor}$ + 0.22* CE $_{fair\ or\ good}$ + 0.58*CE $_{excellent}$ + 0*Innovation $_{poor}$ + 0.42*Innovation $_{fair\ or\ good}$ + 0.71*Innovation $_{excellent}$ + 0*OI $_{high}$ + 0.09*OI $_{moderate\ or\ low}$ + 0.19*OI $_{none}$ + 0*HE $_{poor}$ + 0.44* HE $_{fair\ or\ good}$ + 0.3*HE $_{excellent}$

In the above functions, X refers to the dummy variables included in the model which can take the value 0 or 1. Take survival for an example: if the drug's performance on survival is poor, then Survival $_{poor}$ = 1, Survival $_{fair\ or\ good}$ = 0, Survival $_{excellent}$= 0 (see below table).

	Survival $_{poor}$	Survival $_{fair\ or\ good}$	Survival $_{excellent}$
Poor	1	0	0
Fair or good	0	1	0
Excellent	0	0	1

Suppose that we have a technology with the following response levels on value attributes:

Survival: fair or good, **clinical outcomes**: excellent, **PROs**: fair or good, **safety**: fair or good, **budget impact to payer**: poor, **out-of-pocket costs**: poor, **cost-effectiveness**: fair or good, **innovation**: fair or good, **organizational impact**: poor, **health equity**: fair or good.

Using the value scoring algorithms above, we can estimate the value score of the hypothetical technology:

1) For mild/moderate disease:

Value_mild = 1.64 + 0.95 + 1.12 + 0.57 + 0.55 + 0 + 0 + 0.14 + 0.34 + 0 + 0.08 = 5.39

2) For severe/critical disease:

Value_severe = 1.80 + 1.05 + 1.17 + 0.66 + 0.55 + 0 + 0 + 0.24 + 0.24 + 0 + 0.31 = 6.02

Using the logit functions above, we can estimate the probability of negotiation or coverage:

1) Mild/moderate disease:

Logit (Y=1, negotiation or coverage)_mild = Log (P (Y = 1) / P (Y = 0)) =

Log (P (Y = 1) / (1 – P (Y = 1))) = -4.13 + 1.40 + 1.70 + 0.88 + 0.66 + 0 + 0 + (-0.09) + 0.61 + 0 + 0.30 = 1.33

P (Y = 1, negotiation or coverage)_mild = e^1.33 / (1 + e^1.33) = 0.79

2) Severe/critical disease :

Logit (Y=1, negotiation or coverage)_severe = Log (P (Y = 1) / P (Y = 0)) =

Log (P (Y = 1) / (1 – P (Y = 1))) = -3.35 + 1.81 + 1.80 + 0.75 + 0.78 + 0 + 0 + 0.22 + 0.42 + 0 + 0.44 = 2.87

P (Y=1, negotiation or coverage)_mild = e^2.86 / (1 + e^2.86) = 0.95

CHAPTER 5. CONCLUSION

This chapter concludes the thesis by summarizing the main findings and discussing policy implications of the developed VAF. The limitations of this thesis and opportunities for future research in the field of VAF development are also discussed.

Through a systematic literature review of existing VAFs in Chapter 2, we found that most VAFs were developed for high-income countries with limited involvement of the public in the process of framework development.[1] There were substantial variations in the value attributes included, approaches to attribute identification and aggregation, perspectives and decision criteria.[1] The use of MCDA in VAF has increased in recent years.[1] These findings present a full picture of existing VAFs and has informed the design of the two subsequent projects described in Chapters 3 and 4. In Chapter 3, we identified 12 value attributes through open-ended semi-structured one-on-one interviews with 34 Chinese stakeholders and a review and analysis of 16 government documents using the approach of QD.[2] The stakeholders include policymakers, healthcare providers, HTA researchers, patients and members of the general public, and industry representatives. The attributes encompass a wide range of value dimensions related to severity of disease, health benefit, safety, economic impact, innovation, organizational impact, health equity, and quality of evidence.[2] Based on these identified value attributes, we conducted an online factorial survey to generate the value scoring functions for a VAF in China in Chapter 4. We found that survival, clinical outcome, patient-reported outcomes, and safety were more important for value assessment than other attributes across disease scenarios. The value of a drug and its probability of entering negotiation or being covered by the national medical insurance in China increase with the drug's higher degree of performance on value attributes or for

severer diseases. Using the developed VAF, the value of a given drug and its probabilities of entering negotiation or receiving coverage can be estimated to inform coverage decision-making in China.

Based on these findings, this thesis could contribute to the measurement and application of value assessment in HTA and coverage decision making in the following ways. First, this thesis confirmed that value is multi-dimensional.[1,3–5] In recent years, researchers and a number of regulatory authorities recognized that there are concerns that the conventional CEA and QALY cannot capture (e.g., equity and innovation).[6–9] Value elements such as value of hope, scientific spillover and insurance value have been proposed and discussed for value assessment by HTA organizations and researchers.[1,5,10,11] The National Health Security Agency (NHSA) in China has also included innovation and health equity into the coverage decision making process recently.[12] The qualitative study provided empirical evidence on value attributes that are considered important by Chinese healthcare stakeholders.[2] Second, this thesis revealed that interactions between the severity of disease and the remaining value attributes exist and thus, should be incorporated in health technology value assessment and coverage decision making. In this thesis, severity of disease was not included in the weighted-sum model in parallel to other attributes as in some existing VAFs.[13–15] Instead, it was used to construct different disease scenarios within which different value scoring functions were generated. Previous research has shown that the same health gains for individuals with different levels of medical needs (i.e., severity of disease) are not of the same value to the society.[6,16] Specifically, society tends to value interventions for individuals who are in more urgent need of medical care.[6,17] In

this thesis, participants discussed about the different ranking orders of value attributes for diseases with different levels of severity in the QD study.[2] In the survey, there were differences in mean estimated value scores and mean probability of negotiation or coverage between different disease scenarios for health technologies with the same sum of attribute levels. The use of different disease scenarios in the VAF aligned with stakeholders' perspectives and the society's valuation of health technologies for different groups of population.[6,17] Third, this thesis demonstrated the feasibility of using MCDA for value attribute aggregation and presented an approach to accommodate the uncertainty of coverage decision making in an MCDA-based VAF. Adopted by a number of VAFs, MCDA combines value attributes quantitatively and generates value scores to support decision making.[1,13,15,18] It could promote the transparency of framework development and application but was criticized for the difficulty and rigidity of using value scores and thresholds to facilitate decision making.[1,11,15,19] Through the joint use of estimated value scores and probability of negotiation or coverage by insurance, the developed VAF avoids this limitation of MCDA and provides flexibility to the coverage decision making process.

For the large number of health technologies with varying levels of benefits and risks to be assessed in China each year, the developed VAF provides a formal framework that can be used to assist value assessment. The rapid development of HTA and the increasing discussions of VAF and MCDA in China have presented opportunities to the application of the VAF to support coverage decision making.[20,21] With the estimated value and probabilities of negotiation or coverage, the VAF has the potential to improve the

transparency, efficiency, consistency and robustness in health value assessment and decision-making process in China.

Next, we will validate and promote the use of the developed VAF among decision makers in China. The assessment reports for drugs that have been evaluated by the NHSA in the recent years will be used. These drugs' value profiles, value scores and probabilities of entering negotiation or receiving coverage will be generated using the developed VAF and compared with the decisions made by the NHSA. We will also regularly update the value score and probability estimates for health technologies to incorporate new evidence and modify the framework to account for new attributes or perspectives.

This thesis has several limitations. First, there was a lack of representativeness of people from rural areas, with less education and aged 65 years and older in the QD and factorial survey studies. In 2017, 41.5% of China's population lived in rural areas and 11.4% aged 65 years and above.[22] In 2020, only 8% of the population had bachelor's degree and 1% had master's or doctoral degree.[23] Further research with the aim to improve the representativeness of respondents is needed. Second, the hypothetical health technologies used in the survey were only drugs. Other types of health technologies such as medical devices were not incorporated into the survey due to 1) the consideration that most stakeholders, especially the public have more knowledge and experience about drugs compared to other types of health technologies and therefore can relatively easily understand and complete the survey tasks; 2) the considerable differences between other types of health technologies and drugs with regard to their mechanisms of action, administration methods, and life cycles etc.; 3) the substantial heterogeneity among the

same type of technologies (e.g., medical devices: hearing aids vs CT scanners) in terms of the operator, and procedure that may impact the benefits of the health technology; 4) a lack of consensus on the HTA methods of other types of health technologies in China and many other countries; and 5) the separate reimbursement mechanisms for other types of health technologies in China.[24–26] Modifications to the VAF may be needed if there is evidence showing different preferences for other types of health technologies. Third, ethics and societal implications were excluded from the VAF due to their unclear definitions and difficulty to measuring either qualitatively or quantitatively. Future research could attempt to engage stakeholders to have in-depth discussion on pertinent ethical and societal issues (e.g., the fear of contagion during the COVID-19 pandemic) that may arise with the use of a health technology.

It is important to note that these limitations are related to some of the challenges in the development and application of VAF for health technology value assessment and decision making both in China and other countries. These challenges include the insufficient engagement of the public, lack of assessment methodologies for innovative technologies, and the difficulty of defining and measuring some of the value components.[27–29] Public engagement in the processes of framework development and application has been a key and frequently stated goal of VAFs.[27] The insufficient public engagement could result in failure of capturing outcomes that are important to patients and thus lack of evidence on patient-reported outcomes.[27,28] Various strategies have been used for public engagement in health technology value assessment and decision making, which include, but are not limited to, public consultation, focus group, and including the public in the research team.[4,30,31]

However, these strategies often require that participants have certain levels of knowledge to comprehend the topic and the access to the platform or information. The experience of individuals with low education or limited access to the resources might be under-represented. For example, the EVIDEM framework was adapted to China through discussions among a small group of policymakers, physicians, and health economists. The public and industry representatives were not engaged.[32,33] Effective strategies to improve the public' access to the participation platform and promote the engagement of a representative group from the general public are needed. Another challenge is related to the emergence of diverse innovative technologies, e.g., immunotherapy, medical 3D printing, robotic surgery and medical wearable devices.[29] They impose ongoing challenges to both health technology value assessment and coverage decision makings.[34] Methodologically, it is difficult to measure these innovations. For example, the impact of 3D printing on surgical time and precision is hard to measure due to the customized nature of the technology.[35] Ethically, it may not be feasible to test the safety of some technologies in humans, for example, the bioprinting technology.[35] These concerns or issues contribute to the inadequacy of evidence and uncertainty when using VAFs to support decision making. Future research is needed to develop or identify appropriate outcome measures to standardize their assessment. As for China, the lack of local data (e.g., the costs of these innovative technologies) adds another layer of difficulty to the assessment of these technologies.[36] Finally, controversies and difficulties exist with respect to the definition and measurement of some value attributes. For example, ethical considerations were included as one decision criterion in the weighted-sum model in the framework developed

by Youngkong et al.[37] In their framework, ethics was related to the rarity of the disease and the prevalence of the disease among the poor. The rarer the disease is and the poorer the patients are, the higher value score the technology has on ethics. In contrast, ethics was included in the EVIDEM framework by reflecting the ethical foundation for each attribute.[38] Three aspects were included for the ethical consideration: alleviation or prevention of patient suffering, prioritization of those who are worst off while ensuring greatest good for greatest number and ensuring sustainability.[38] In our framework, however, ethics tend to be related to clinical trials and healthcare professionals' behaviours.[2] It is unclear what are the appropriate methods to describe the ethics issue and to incorporate it into value assessment. Further research in this area is warranted.

The development of our VAF was capitalized on the knowledge from the development of existing VAFs, and based on the extensive engagement with multiple stakeholders through both qualitative and quantitative research. The joint use of the value estimate and probability prediction in the VAF incorporates the complexity and uncertainty of real coverage decision making in China. The developed VAF could be a useful tool to facilitate transparent, efficient, consistent, and robust health technology value assessment and coverage decision making in China.

References

1. Zhang M, Bao Y, Lang Y, et al. What Is Value in Health and Healthcare? A Systematic Literature Review of Value Assessment Frameworks. *Value Health*. 2022;25(2):302-317. doi:10.1016/j.jval.2021.07.005

2. Zhang M, Bao Y, Yi Y, Kimber M, Levine M, Xie F. Identifying attributes for value assessment framework in China: a qualitative study (accepted). *PharmacoEconomics*. Published online 2022.

3. O'Rourke B, Oortwijn W, Schuller T, International Joint Task Group. The new definition of health technology assessment: A milestone in international collaboration. *Int J Technol Assess Health Care*. 2020;36(3):187-190. doi:10.1017/S0266462320000215

4. Institute for Clinical and Economic Review. 2020 Value Assessment Framework | Topic. ICER. Published 2020. Accessed August 20, 2020. https://icer-review.org/topic/2020-value-assessment-framework/

5. Lakdawalla DN, Doshi JA, Garrison LP, Phelps CE, Basu A, Danzon PM. Defining Elements of Value in Health Care-A Health Economics Approach: An ISPOR Special Task Force Report [3]. *Value Health*. 2018;21(2):131-139. doi:10.1016/j.jval.2017.12.007

6. Nord E, Daniels N, Kamlet M. QALYs: some challenges. *Value Health*. 2009;12 Suppl 1:S10-15. doi:10.1111/j.1524-4733.2009.00516.x

7. Pettitt D, Raza S, Naughton B, et al. The limitations of QALY: a literature review. *Journal of Stem Cell Research and Therapy*. 2016;6(4). Accessed December 11, 2022. https://ora.ox.ac.uk/objects/uuid:d22a9b1b-c2f0-4f6c-83ee-fdc320e4af61

8. National Institute for Health and Clinical Excellence. Appraising life-extending, end of life treatments. National Institute for Health and Care Excellence. Published 2009. https://www.nice.org.uk/guidance/gid-tag387/documents/appraising-life-extending-end-of-life-treatments-paper2

9. Persson U. Value Based Pricing in Sweden: Lessons for Design? *Seminar Briefings*. Published online November 1, 2012. Accessed December 11, 2022. https://ideas.repec.org/p/ohe/sembri/000141.html

10. Garrison LP, Neumann PJ, Willke RJ, et al. A Health Economics Approach to US Value Assessment Frameworks-Summary and Recommendations of the ISPOR Special Task Force Report [7]. *Value Health*. 2018;21(2):161-165. doi:10.1016/j.jval.2017.12.009

11. Krahn M, Miller F, Bayoumi A, et al. DEVELOPMENT OF THE ONTARIO DECISION FRAMEWORK: A VALUES BASED FRAMEWORK FOR HEALTH TECHNOLOGY ASSESSMENT. *Int J Technol Assess Health Care*. 2018;34(3):290-299. doi:10.1017/S0266462318000235

12. National Healthcare Security Administration (NHSA). Work Plan for the Adjustment of the National Medical Insurance Drug List in 2022. Published 2022. Accessed January 20, 2023. http://www.nhsa.gov.cn/art/2021/6/30/art_62_5386.html

13. Guarga L, Badia X, Obach M, et al. Implementing reflective multicriteria decision analysis (MCDA) to assess orphan drugs value in the Catalan Health Service (CatSalut). *Orphanet J Rare Dis*. 2019;14(1):157. doi:10.1186/s13023-019-1121-6

14. EVIDEM. Evidence and Value: Impact on DEcision Making (EVIDEM) 10th Edition. Published 2017. Accessed February 11, 2020. https://www.evidem.org/

15. 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

16. Marsh KD, Sculpher M, Caro JJ, Tervonen T. The Use of MCDA in HTA: Great Potential, but More Effort Needed. *Value in Health*. 2018;21(4):394-397. doi:10.1016/j.jval.2017.10.001

17. Nord E. Concerns for the worse off: fair innings versus severity. *Social Science & Medicine*. 2005;60(2):257-263. doi:10.1016/j.socscimed.2004.05.003

18. Pichon-Riviere A, Garcia-Marti S, Oortwijn W, Augustovski F, Sampietro-Colom L. Defining the Value of Health Technologies in Latin America: Developments in Value Frameworks to Inform the Allocation of Healthcare Resources. *Int J Technol Assess Health Care*. 2019;35(1):64-68. doi:10.1017/S0266462319000072

19. Marsh K, IJzerman M, Thokala P, 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

20. National Health Commission of the People's Republic of China. Announcement on the Public Solicitation of Opinions on the Guidelines for the Management of Comprehensive Clinical Evaluation of Drugs. Published 2020. Accessed September 2, 2021. http://www.nhc.gov.cn/yaozs/s7656/202011/d11ddc32fae84121a0dfca36b015a31d.shtml

21. National Health Commission of the People's Republic of China. Management Guidelines for Comprehensive Clinical Evaluation of Drugs. Published 2021. Accessed December 30, 2021. http://www.nhc.gov.cn/cms-search/xxgk/getManuscriptXxgk.htm?id=532e20800a47415d84adf3797b0f4869

22. United Nations Children's Fund. Population demographics. Published 2019. Accessed January 20, 2023. https://www.unicef.cn/sites/unicef.org.china/files/2019-06/01EN-Population%20demographics%20Atlas%202018.pdf

23. OECD. *Education at a Glance 2022: OECD Indicators*. OECD; 2022. doi:10.1787/3197152b-en

24. Ming J, He Y, Yang Y, et al. Health technology assessment of medical devices: current landscape, challenges, and a way forward. *Cost Effectiveness and Resource Allocation*. 2022;20(1):54. doi:10.1186/s12962-022-00389-6

25. National Health Security Administration. Letter from the National Health Security Administration on the Proposal No. 02866 (Social Management No. 263) of the Fifth Session of the 13th National Committee of the Chinese People's Political Consultative Conference. Published 2022. Accessed January 27, 2023. http://www.nhsa.gov.cn/art/2022/10/12/art_110_9220.html

26. Sorenson C, Drummond M. Improving medical device regulation: the United States and Europe in perspective. *Milbank Q*. 2014;92(1):114-150. doi:10.1111/1468-0009.12043

27. Dubois RW, Westrich K. As Value Assessment Frameworks Evolve, Are They Finally Ready for Prime Time? *Value Health*. 2019;22(9):977-980. doi:10.1016/j.jval.2019.06.002

28. Perfetto EM, Oehrlein EM, Boutin M, Reid S, Gascho E. Value to Whom? The Patient Voice in the Value Discussion. *Value in Health*. 2017;20(2):286-291. doi:10.1016/j.jval.2016.11.014

29. *CADTH Issues in Emerging Health Technologies*. Canadian Agency for Drugs and Technologies in Health; 2016.

30. Avalere Health. Avalere and FasterCures Release Patient-Perspective Value Framework to Incorporate Patient Preferences into Healthcare Treatment Decisions. Published 2017. Accessed March 19, 2021. https://avalere.com/press-releases/avalere-and-fastercures-release-patient-perspective-value-framework-to-incorporate-patient-preferences-into-healthcare-treatment-decisions

31. Jehu-Appiah C, Baltussen R, Acquah C, et al. Balancing equity and efficiency in health priorities in Ghana: the use of multicriteria decision analysis. *Value Health*. 2008;11(7):1081-1087. doi:10.1111/j.1524-4733.2008.00392.x

32. Bao Y, Gao B, Meng M, et al. Impact on decision making framework for medicine purchasing in Chinese public hospital decision-making: determining the value of five dipeptidyl peptidase 4 (DPP-4) inhibitors. *BMC Health Serv Res*. 2021;21(1):1-7. doi:10.1186/s12913-021-06827-0

33. Geng J, Chen X, Yu X, Huang J, He D, Dong J. Study on the evidence-based decision-making framework for reimbursement technologies in view of EVIDEM. *Chinese Journal of Health Policy*. 2018;11(04):50-54.

34. Hogervorst MA, Vreman RA, Mantel-Teeuwisse AK, Goettsch WG. Reported Challenges in Health Technology Assessment of Complex Health Technologies. *Value in Health*. 2022;25(6):992-1001. doi:10.1016/j.jval.2021.11.1356

35. Mason J, Visintini S, Quay T. An Overview of Clinical Applications of 3-D Printing and Bioprinting. In: *CADTH Issues in Emerging Health Technologies*. Canadian Agency for Drugs and Technologies in Health; 2016. Accessed December 9, 2022. http://www.ncbi.nlm.nih.gov/books/NBK542711/

36. Kennedy-Martin T, Mitchell BD, Boye KS, et al. The Health Technology Assessment Environment in Mainland China, Japan, South Korea, and Taiwan—Implications for the Evaluation of Diabetes Mellitus Therapies. *Value in Health Regional Issues*. 2014;3:108-116. doi:10.1016/j.vhri.2014.03.001

37. Youngkong S, Baltussen R, Tantivess S, Mohara A, Teerawattananon Y. Multicriteria decision analysis for including health interventions in the universal health coverage benefit package in Thailand. *Value Health*. 2012;15(6):961-970. doi:10.1016/j.jval.2012.06.006

38. Goetghebeur MM, Cellier MS. Can reflective multicriteria be the new paradigm for healthcare decision-making? The EVIDEM journey. *Cost Eff Resour Alloc*. 2018;16(Suppl 1):54. doi:10.1186/s12962-018-0116-9