The Use of Multicriteria Decision Analysis to Support Decision Making in Healthcare: An Updated Systematic Literature Review

Pamela Gongora-Salazar, MSc, Stephen Rocks, MSc, Patrick Fahr, DPhil, Oliver Rivero-Arias, DPhil, Apostolos Tsiachristas, PhD

ABSTRACT

Objectives: Multicriteria decision analysis (MCDA) is increasingly used for decision making in healthcare. However, its application in different decision-making contexts is still unclear. This study aimed to provide a comprehensive review of MCDA studies performed to inform decisions in healthcare and to summarize its application in different decision contexts.

Methods: We updated a systematic review conducted in 2013 by searching Embase, MEDLINE, and Google Scholar for MCDA studies in healthcare, published in English between August 2013 and November 2020. We also expanded the search by reviewing grey literature found via Trip Medical Database and Google, published between January 1990 and November 2020. A comprehensive template was developed to extract information about the decision context, criteria, methods, stakeholders involved, and sensitivity analyses conducted.

Results: From the 4295 identified studies, 473 studies were eligible for full-text review after assessing titles and abstracts. Of those, 228 studies met the inclusion criteria and underwent data extraction. The use of MCDA continues to grow in healthcare literature, with most of the studies (49%) informing priority-setting decisions. Safety, cost, and quality of care delivery are the most frequently used criteria, although there are considerable differences across decision contexts. Almost half of the MCDA studies used the linear additive model whereas scales and the analytical hierarchy process were the most used techniques for scoring and weighting, respectively. Not all studies report on each one of the MCDA steps, consider axiomatic properties, or justify the methods used.

Conclusions: A guide on how to conduct and report MCDA that acknowledges the particularities of the different decision contexts and methods needs to be developed.

Keywords: multicriteria decision analysis, healthcare, decision-making, priority-setting, systematic literature review.

Introduction

Multicriteria decision analysis (MCDA) is defined as a systematic and theory-based approach to perform a comparative analysis of several competing options (e.g., healthcare interventions) based on their performance on multiple and often conflicting criteria. The development of a quantitative MCDA involves the following steps: (1) defining the decision problem, (2) selecting the evaluation criteria, (3) assessing the performance scores of each alternative on each criterion, (4) determining the criteria weights, (5) aggregating performance scores and criteria weights in an overall value, (6) dealing with uncertainty, and (7) examining findings and deliberate.

In healthcare, MCDA has emerged as an alternative or complementary framework to address limitations of traditional health technology assessment (HTA) approaches such as cost-utility analysis. As documented previously, the growing need for MCDA in healthcare is reflected by the steep increase of empiric studies across different decision contexts. Although the application of MCDA in healthcare has been summarized in several review studies, none of them has systematically detailed how MCDA has been applied within and across different healthcare decision contexts. The most comprehensive review of MCDA studies was published 8 years ago by Marsh et al. and did not examine the methodological differences by decision context. In addition, all previous review studies did not include grey literature, such as technical documents from HTA agencies and policy reports from healthcare authorities or multilateral bodies, thereby increasing the risk of missing important applications of MCDA in healthcare. The aim of this updated systematic literature review was to provide a comprehensive overview of MCDA studies in different healthcare decision contexts.
achieve this, we provided a detailed overview of (1) the different decision contexts in healthcare and then by decision context, (2) the criteria used, (3) the data sources and methods used to derive performance scores, (4) the applied scoring and weighting techniques, (5) the approaches used to obtain overall values and rank the compared alternatives, and (6) the type of sensitivity analysis conducted.

This is the first review that includes grey literature and summarizes key methodological information by decision context. Such information could guide researchers in designing MCDA studies by supporting the selection of appropriate methods and criteria relevant to their decision context. In light of the numerous MCDA methods available and the criticism of lack of transparency, findings from this review could also be used as a basis to start developing a detailed methodological and reporting guidance by decision context.

Methods

This review was registered in PROSPERO (CRD42020219093) in November 2020, and the results were reported following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses 2020 statement, and the guidance for updated systematic reviews, and recently published systematic literature reviews in Value in Health journal.

Search Strategy

To update the review by Marsh et al (2014), we conducted a database search of scientific literature published between August 2013 and November 2020. We decided to update this review because it seems to be the most comprehensive in terms of the number of included studies, extracted information, and data synthesis. As Marsh et al (2014) included studies up to 2013, we searched for more recent MCDA studies in the reviews of Adulinn et al (2015) and Frazão et al (2018), which also reviewed the application of MCDA in different decision contexts in healthcare. We also searched grey literature published between January 1990 and November 2020 to expand the scope of review of Marsh et al. All searches were conducted by P.G. under the guidance of a librarian who advised on the search strategy and selection of databases.

Variations in search terms were added and terms were searched as free-text keywords in the title and abstract fields. The search terms were based on Marsh et al (2014), Adulinn et al (2015), and Frazão et al (2018). Scientific studies were identified from the following databases: MEDLINE, Embase, and Google Scholar (100 first hits).

It is expected that all relevant academic studies are identified from MEDLINE and Embase, because they are high-level academic databases. Nevertheless, Google Scholar was used as an additional tool to locate potentially relevant studies that might not have been included in MEDLINE or Embase. Although for primary review searches Google Scholar is unsuitable, it is considered a suitable supplementary source of evidence when conducting systematic literature reviews. Due to the vast number of results displayed and the inability to directly export results in bulk as citations, systematic reviews typically screen the title and short text under each other. Relevant titles were selected for further screening by a reviewer using the eligibility criteria. All previously defined search terms were used. We used Google to search within 53 websites of governmental entities, networks, societies, and international bodies in health economics, using the “site” option in the search box (eg, site:nhs.uk) and all previously defined search terms. For Google, the search was limited to 4 search terms given that searches cannot be limited to abstract and title. See Appendices 1 and 2 in Supplemental Materials found at https://doi.org/10.1016/j.jval.2022.11.007 for more in detail information about the search terms and websites used.

Selection Process and Eligibility Criteria

All study titles and abstracts were exported to EndNote X9 (Clarivate, Philadelphia, PA). Before the screening of titles and abstracts, all reviewers met to finalize the eligibility criteria. Titles and abstracts were doubled screened independently by 2 reviewers (P.G., A.R., and S.R.): P.G. screened titles and abstracts of all records, and the same studies were equally split and screened by A.R. and S.R. Disagreements between reviewers were resolved through discussion, and the fourth reviewer (A.T.) was consulted when necessary. All studies were assessed and included based on the following eligibility criteria:

Inclusion criteria

1. Full-text studies in English language
2. Studies that reported an empiric application of MCDA to inform decisions in healthcare and described the MCDA methods used
3. Studies that specified the MCDA method implemented

Exclusion criteria

1. Studies that did not apply MCDA
2. Studies that did not complete all the steps in the MCDA definition stated in the introduction (eg, relevant criteria were identified but weights were not elicited or defined)
3. Studies that did not aim to inform a decision in healthcare (eg, studies that used MCDA to derive health status utility values and with this calculate quality-adjusted life-years)

Similar to Marsh et al’s study (2014), studies were not excluded on the basis of methodological quality. Nevertheless, we used a broader definition of healthcare by adopting a system perspective. Therefore, we included all studies that addressed decisions made within the healthcare sector or the entire healthcare system (ie, decisions made across the production, sourcing, organization, and delivery of healthcare) and exclude studies that inform interventions delivered outside the health sector (eg, social, occupational, or environmental interventions), despite their potential health impact.

Data Extraction and Analysis

A comprehensive template for data extraction was developed and piloted to ensure high level of consistency in extracting data between reviewers (P.G., A.R., and P.F.). The data were grouped in decision contexts, criteria, weighting and scoring techniques, sources of data, stakeholders involved, and methods to address uncertainty. The categorization of decision contexts was based on Marsh et al (2014), but we provided another layer of detail by providing subcategories in each decision context. Then, we allocated the selected studies to each subcategory using the reported (often in the study aim) information on what decision each MCDA study was aiming to inform. To define the different criteria categories, we used the taxonomies from Tsiachristas et al, and
Marsh et al (2014) and used an explicit code book with definitions and examples for each category. To identify the type of uncertainty addressed by the MCDA studies, we used Broekhuizen et al’s classification and distinguish among heterogeneity, parameter uncertainty (ie, criterion weight or performance score), and structural uncertainty.

Descriptive statistics were used for analyzing data with regard to the decision context, the applied MCDA techniques, and the type of stakeholders that participated in each of the MCDA steps. We used Microsoft Excel (Microsoft, Redmond, WA) for constructing all tables, figures, and descriptive statistics and summarized the results for each of the MCDA steps.

**Reviewers’ Consistency**

For reducing potential biases during the assessment and data extraction process and following good practice guidelines, a selection of papers was discussed between reviewers during the agreement meeting. P.G. reviewed and extracted data from a sample of 10% of the studies reviewed by A.R. and P.F., respectively. A.R. and P.F. did the same with 10% of the records reviewed by P.G. The extracted data were then compared and a strong consensus was achieved between reviewers for both the assessment and extraction process (> 95% and > 85% inter-reviewer agreement, respectively).

**Results**

The Preferred Reporting Items for Systematic Reviews and Meta-Analyses flowchart of this review is presented in Figure 1. After removing duplicates, 4295 studies (3256 and 1039 studies from the database and grey literature search, respectively) were retained. After the screening of all records, 438 academic titles and 35 grey literature documents met the inclusion criteria. The full-text examination of the retrieved records excluded 230 academic studies and 15 grey literature documents. The analysis included in the end 228 eligible studies (219 MCDA studies). A summary of the results of all searches conducted, with the list of studies included in the review, can be found in Appendix 2 in Supplemental Materials found at https://doi.org/10.1016/j.jval.2022.11.007. Derived data supporting the findings of this study are available upon request.

**Study Characteristics**

Figure 2 illustrates the number of MCDA studies (ie, 219 MCDA studies in total) in healthcare between 1990 and 2020. Although the first studies were conducted in the early 90s, it is since 2010 that a sharp upward trend can be observed, with no grey literature found before 2010. Studies informing priority-setting decisions were the first and have been the most predominant, although the share of studies informing other decision contexts has increased over the last 10 years. A definition of each one of the decision contexts identified is provided in Appendix 2 in Supplemental Materials found at https://doi.org/10.1016/j.jval.2022.11.007.

In almost 90% of the MCDA studies (n = 192), academia was involved either solely (62%, n = 135) or in conjunction with the public sector (18%, n = 39), the private sector (10%, n = 22), or international organizations (3%, n = 7).

**Decision Context**

As presented in Table 1, most studies have informed priority-setting decisions (48%, n = 107). Roughly 22% of MCDA studies (n = 48) were concerned with clinical decision making, and 11% (n = 25) with guiding regulatory decisions such as marketing authorizations. Although to a lesser extent, MCDA studies have...
also informed hospitals or health systems when making capital investment decisions (4%, n = 8) or decisions on the location or reallocation of healthcare services or facilities (4%, n = 8). We also identified MCDA studies conducted to support manufacturers in the development of new medicines (4%, n = 8), to guide research decisions (2%, n = 5), to inform the regulation of addictive substances (1%, n = 3), and to support the selection of best healthcare wastes management method (1%, n = 3).

Most of the MCDA studies (63%, n = 137) have taken place in high-income countries, principally in Europe (37%, n = 82) and North America (19%, n = 41). With exception of regulatory decision contexts, where almost all studies have informed high-income countries (96%, n = 22), approximately a third of MCDA studies applied in the other decision contexts have taken place in low- and middle-income countries.

More than a third of the MCDA studies (35%, n = 76) have been developed to inform decisions at national level, followed by studies conducted to inform decisions at local level (14%, n = 31). Nevertheless, 21% of the studies (n = 47) did not report the decision level. Regulators or budget holders are the most common users of MCDA in healthcare, with 40% of the studies (n = 90) informing ministries of health, commissioners, public health institutes, or local authorities. Nevertheless, in clinical decision making, 56% of the MCDA studies (n = 27) aimed to inform physicians, 29% (n = 14) informed patients or carers, and only 10% (n = 5) informed regulators or budget holders.

Table 1. Decision contexts that MCDA studies have been applied in.

<table>
<thead>
<tr>
<th>Decision context</th>
<th>n</th>
<th>% of 219*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Priority setting</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prioritization of interventions for</td>
<td>85</td>
<td>39</td>
</tr>
<tr>
<td>coverage, reimbursement, funding or future developments</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prioritization of diseases for coverage, reimbursement</td>
<td>12</td>
<td>5</td>
</tr>
<tr>
<td>or funding</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prioritization of patients to access healthcare</td>
<td>10</td>
<td>5</td>
</tr>
<tr>
<td>Clinical decision making</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treatments or Prescription</td>
<td>34</td>
<td>16</td>
</tr>
<tr>
<td>Screening</td>
<td>14</td>
<td>6</td>
</tr>
<tr>
<td>Regulatory decisions</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BRA to issue recommendations†</td>
<td>25</td>
<td>11</td>
</tr>
<tr>
<td>Planning and R&amp;D</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Capital investment decisions</td>
<td>8</td>
<td>4</td>
</tr>
<tr>
<td>Location/reallocation decisions</td>
<td>8</td>
<td>4</td>
</tr>
<tr>
<td>Guide pharmaceutical developments</td>
<td>8</td>
<td>4</td>
</tr>
<tr>
<td>Research purposes</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>Regulation of addictive substances</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>Decisions on hospital wastes management decisions</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>Other investing planning decisions‡</td>
<td>9</td>
<td>3</td>
</tr>
</tbody>
</table>

BRA indicates benefit-risk assessment; MCDA, multicriteria decision analysis; R&D, research and development.

*5 MCDA studies inform more than one type of decision.
†For example, market authorization.
‡For example, MCDA to guide the selection of healthcare providers or to choose best length of stay policy alternatives.
MCDA studies in healthcare have evaluated a diverse types of alternatives. More than 40% of the studies (n = 93) have assessed pharmaceuticals, followed by public health (14%, n = 31) and screening interventions (12%, n = 27). Pharmaceuticals are the most evaluated alternative (41%, n = 90) across all decision contexts, a frequency that goes up to 92% (n = 23) in the regulatory decision context. The second most frequent are public health interventions among priority-setting studies (26%, n = 28) and medical screening programs in the case of clinical decision making (25%, n = 12). Nearly 16% of the eligible studies (n = 35) evaluated more than one type of alternative.

In terms of the number of alternatives assessed, on average, MCDA studies have evaluated 15 alternatives (median = 5, SD = 39.8), with a minimum of one and a maximum of 500. Priority-setting studies evaluated on average 20 alternatives (median = 8, SD = 51, min = 1, max = 500). This contrasts with the lower number of alternatives assessed when informing diagnosis or prescription decisions (mean = 5.1, median = 4, SD = 4.6, min = 1, max = 22) and regulatory decision (mean = 4.9, median = 4, SD = 3.6, min = 1, max = 15) or planning and research and development (R&D) decisions (mean = 18.7, median = 6, SD = 39.8, min = 1, max = 194). See Appendix 3 in Supplemental Materials found at https://doi.org/10.1016/j.jval.2022.11.007 for more detailed information about the decision contexts.

Criteria

Different sources have been used to identify and define the MCDA criteria. Almost half of the MCDA studies have chosen criteria based on literature (n = 103). From these, 50% (n = 52) have solely used literature as a source, whereas the rest (n = 51) have additionally used other sources (eg, specialists or experts in the field). A quarter of the studies (n = 56) have chosen the criteria based on expert opinion, of which 15% (n = 32) used literature as well. More than 20% (n = 48) have adopted an established set of criteria (published or not published), most frequently from the Evidence and Value: Impact on Decision Making framework (7%, n = 16).26

Less than 35% (n = 76) have considered stakeholders’ views in the process of defining the criteria set. This includes providers of medical care (13%, n = 28), health authorities (8%, n = 18), patients (6%, n = 13), the private sector (3%, n = 7), social care authorities (2%, n = 5), government entities for social care services (2%, n = 4), and nongovernmental organizations (0%, n = 1).

In terms of the ways to agree on the criteria, approximately 37% of the studies (n = 80) have resorted to group discussions, 18% (n = 39) have relied on experts’ opinions, and 11% (n = 23) have based the final selection of criteria on related literature. Nevertheless, a considerable number of studies have failed to report the method adopted to agree on the criteria (39%, n = 86).

When structuring the criteria, 41% of the studies (n = 89) have adopted a hierarchical structure. On average, 4.8 criteria (SD = 2.1, min = 2.0, max = 18) and 13.7 subcriteria (SD = 8.5, min = 3, max = 45) have been included in MCDA studies in healthcare that used a hierarchical structure. Studies that adopt a nonhierarchical structure have used on average 7.1 criteria (SD = 3.2, min = 2, max = 18).

As presented in Figure 3, several criteria have been used to inform decisions in healthcare, with “safety” (54%, n = 119), “cost” (49%, n = 107), “quality of care delivery” (40%, n = 88), “intermediate health outcomes” (34%, n = 74), and “feasibility or acceptability” (30%, n = 65) as the 5 most frequently used. Although in regulatory decisions > 70% of the studies have used “safety” (100%, n = 25) and “intermediate health outcomes” criteria, in priority setting a wider variety of criteria have been used, with “cost” (52%, n = 56) and “safety” (44%, n = 47) at the front. Studies informing clinical decisions have mainly used “safety” (75%, n = 36) and “quality or process of care delivery” (60%, n = 29), whereas in the context of planning and R&D, “cost” (66%, n = 29) and “other” (50%, n = 22) are the 2 most predominant criteria. If we aggregate the number of studies that have used intermediate, final, and unspecified health outcomes as criteria, “health outcomes” becomes the most common criterion in all decision contexts, but planning and R&D.

MCDA Approaches

As presented in Figure 4A, the vast majority of studies have adopted a value-measured approach22 (86%, n = 189), with linear additive models (eg, multiattribute utility theory, weighted sum) and analytical hierarchy process (AHP)28 as the most common approaches adopted. Only 21 studies (10%) have used outranking models, such as PROMETHEE,30 and 15 studies (7%) adopted a goal programming approach,2 with TOPSIS25 being the most popular one. A similar pattern is observed across studies informing prioritization decisions and treatment or screening decisions. Almost all studies that inform regulatory decisions via benefit-risk assessment (BRA) have adopted the linear additive model (91%, n = 21). These approaches were frequently applied using Microsoft Excel (29%, n = 37) and Expert Choice (15%, n = 19), although just 58% of the studies (n = 127) have reported the software used.

More than 70% of the studies retrieved (n = 156) have provided some justification for the MCDA approach applied, but few have justified the scoring and weighting technique adopted. Some of the reasons offered by these studies include the broad application of the method, the theoretical soundness, and the flexibility and relative ease to use.

Performance Scoring

Direct rating methods seem to be the preferred ones for scoring alternative’s performance, with 30% of the studies (n = 65) using direct scales (eg, Likert scale) and 5% (n = 12) using a point allocation system (see Fig. 4B). AHP (26%, n = 56) and value function (13%, n = 29) are the second and the third most frequently used methods, respectively. Scales are the preferred technique (38%, n = 41) among prioritization-focused studies, whereas in clinical decision making and planning and R&D most studies have used AHP (35%, n = 17). Value function seems to be preferred when conducting BRA (60%, n = 15).

Regarding the sources to measure the performance of alternatives, 60% of the studies (n = 131) have used evidence from the literature. Approximately 32% of the MCDA studies (n = 69) have relied on experts’ opinions, 16% (n = 35) have used trial data, and 14% (n = 30) have used routinely collected data. Approximately 20% of the studies (n = 41) have solely used research evidence, and 5% (n = 12) have relied only on expert opinions as a source. Almost half of the studies (n = 104) have used more than one source when measuring alternatives’ performance, and more than a fifth (n = 50) have failed to report information on this MCDA stage. There are no considerable differences across decision contexts.

In terms of the stakeholders that participate in the scoring of alternatives, most of the studies have involved experts (41%, n = 90), healthcare providers (40%, n = 88), representatives from health authorities (23%, n = 50), and patients or carers (15%, n = 32). Few studies involved the private sector (7%, n = 15) or the general public (6%, n = 14). In 98 studies (45%), more than one actor has participated in this MCDA step, and on average 69 individuals (median = 12.5; SD = 346.7, min = 1, max = 3914) have been involved at this stage. Nevertheless, roughly 40% of the studies (n = 87) failed to report the number of stakeholders.
involved in the scoring of alternatives (see Appendix 3 in Supplemental Materials found at https://doi.org/10.1016/j.jval.2022.11.007).

**Criteria Weighting**

To elicit stakeholders’ preferences between criteria, 33% of the studies (n = 73) have used AHP, with 76% of them (n = 56) using this technique also at the scoring stage (see Fig. 4B). Direct rating was also applied (24%, n = 52), followed by swing weighting (11%, n = 25), discrete choice experiment (6%, n = 14), and other pairwise comparison methods (6%, n = 14). A similar pattern is observed across decision contexts, except for studies informing “regulatory decisions” in which swing weighting has been used by most MCDA studies (56%, n = 14), with relatively few using AHP (16%, n = 4) (see Appendix 3 in Supplemental Materials found at https://doi.org/10.1016/j.jval.2022.11.007).

Although virtually the same type of stakeholders are involved in the weighting and scoring steps (see Appendix 3 in Supplemental Materials found at https://doi.org/10.1016/j.jval.2022.11.007), a relatively larger number of individuals is on average involved in the weighting of the criteria (mean = 118, median = 12.5, SD = 493.8, min = 1, max = 4288).
Value-measurement models include: (A) MAUT, MAVT, weighted sum, and SMART methods; (B) AHP; (c) MACBETH; (d) DCE; (e) PAPRIKA; and (f) other value-measurement models (eg, Socio-Technical Allocation of Resources). Outranking models include: (g) PROMETHEE, and (h) other outranking models (eg, elimination and choice expressing the reality). Goal programming models include: (i) TOPSIS and (j) other goal programming models (eg, VIKOR). (K) Direct rating includes scales and point allocation techniques.

AHP indicates analytical hierarchy process; BRA, benefit-risk assessment; DCE, discrete choice experiment; MACBETH, measuring attractiveness through a categorical-based evaluation technique; MAUT, multiattribute utility theory; MAVT, multiattribute value theory; MCDA, multicriteria decision analysis; PAPRIKA, potentially all pairwise rankings of all possible alternatives; PROMETHEE, preference ranking organization method for enrichment evaluations; SMART, simple multiattribute rating technique; TOPSIS, technique for order preference by similarity to ideal solution.
Uncertainty Analyses

More than 65% of the MCDA studies (n = 145) in healthcare have addressed one or more types of uncertainty.23 Half of the MCDA studies (n = 111) have addressed uncertainty in the weighting of the criteria, a fifth in the scoring step (n = 45), and nearly 15% (n = 31) have accounted for both types of parameter uncertainty (Fig. 5). Heterogeneity in criteria weights or performance scores (i.e., variability attributed to a person’s characteristics) was assessed in 18% of the MCDA studies (n = 40). Dealing with structural uncertainty (e.g., MCDA method used, criteria included) is less common, with only 9% of the studies (n = 20) reporting it.

Across decision contexts, the patterns are roughly similar. Nevertheless, studies informing planning and R&D decisions are the exception, with < 2% of these studies reporting structural uncertainty and only 5% reporting parameter uncertainty on the performance scores.

Most studies that assessed uncertainty have used scenario analysis (29%, n = 64) or a one-way deterministic analysis (23%, n = 51). Fuzzy logic and probabilistic sensitivity analysis are less...
Discussion

The application of MCDA continues growing in healthcare, with most of the studies being conducted in high-income countries to inform priority setting and regulatory decisions at national level. The fact that no MCDA has been conducted to inform regulatory decisions in low- and middle-income countries, despite the increasing use of MCDA in other decision contexts, potentially reflects the embryonic stage of Hta in these countries.39,41

As identified by Khan et al (2021),42 it is just after 2010 that a rise in MCDA studies in healthcare took off. This upward trend might have been driven by the rise and consolidation of HTA agencies across Europe, an increasing concern from national and local governments to justify investment and authorization decisions,36,37 and the use of MCDA in emerging decision contexts (eg, pharmaceutical R&D). MCDA studies informing regulatory decisions emerged between 2009 and 2012, which coincides with the publication of the first report from the European Medicines Agency in 2010, acknowledging MCDA as one of the quantitative approaches to conduct BRA for medicinal products.38 In the United States and later on across Latin America, the industry might have played a role in promoting the use of MCDA when conducting BRA of medicines.39-41

As expected, we found considerable differences across decision contexts on the criteria used and some noticeable preferences for specific criteria sets, particularly among studies that have informed regulatory decisions. Nevertheless, there is significant within-context variation in the type of criteria used, which does not facilitate study comparisons and can potentially affect the consistency of MCDA studies.7 Defining a set of top-level criteria clusters for each one of the different decision contexts could contribute to the comparability of MCDA studies and help decision makers take into account all relevant aspects when agreeing on the criteria. This could reduce subjectivity and variation in criteria sets within decision contexts and improve the MCDA’s credibility in informing healthcare decisions. An example of this is the advanced value tree with 5 key value domains proposed by Angelis and Kanavos42 for evaluating new medicines in HTA.

With regard to the number of criteria used, we found that more than a third of the studies have used > 5 criteria and in some cases > 10 criteria. Nevertheless, this appears to be not in line with good practice and MCDA guidelines that recommend the inclusion of only few criteria.4 This increases the risk of violating the non-overlap and preferential independence requirements of linear additive models undermining the usefulness of the most used MCDA method.43,44 A large number of criteria also increases data requirements, making the application of MCDA more complex and compromising the quality of the results. Although flexibility in the criteria selection step might be desirable in some settings, a uniform set of criteria by decision context that considers all axiomatic properties could serve as a starting point for future applications of MCDA in healthcare. This could potentially improve the consistency and credibility of the resulting recommendations.2,22,46

With regard to the structure of the criteria, we found that less than half of the studies have adopted a hierarchical structure. Exploring hierarchies, using for instance value trees, is a crucial step of the definition of criteria and should be reported by studies.5 Even if the subcriteria are not operationalized in the end, identifying them can help MCDA developers verify that the axiomatic requirements for linear additive models are not violated. Identifying hierarchies could be an essential part of a detailed MCDA guidance.

This review is the first that summarizes the criteria used in MCDA studies in healthcare, and such findings could be used as a basis to start developing a uniform set of top-level criteria by decision context.

Our results highlight a preference of MCDA users in applying the additive model for calculating total scores (ie, MCDA approach) and hybrid techniques (ie, scales for scoring and AHP for weights elicitation) in the scoring and weighting steps. Nevertheless, not all studies made a clear distinction between the MCDA approach used and the scoring and weighting techniques applied, and very few justified the selection of the techniques used. Most of the studies have also favored scoring and weighting techniques that are relatively easier to apply (eg, direct rating approaches and AHP) but less theoretically sound.57 This might partly reflect the little guidance on the choice and application of MCDA methods. As highlighted previously, the lack of standardization of the terms and classifications of MCDA methods may partially explain the flaws in reporting MCDA studies.10,11 Khan et al (2021)36 identified at least 4 different classifications of MCDA models or approaches. Different nomenclatures and categorizations have also been used to summarize the scoring and weighting techniques applied. For instance, the review conducted by Marsh et al (2014)7 reported scales, AHP, natural units, and discrete choice experiment as weighting techniques, whereas Oliveira et al (2019)11 reported point systems, direct rating, measuring attractiveness through a categorical-based evaluation technique, AHP, and selecting functions. This lack of consensus affects the comparability, credibility, and policy usefulness of MCDA studies in healthcare.2,41

Although many types of uncertainty might arise when conducting MCDA in healthcare23,48-50 and all guidelines recommend addressing them,44 uncertainty analysis is still not universally and consistently used across MCDA studies. This contrasts with studies using traditional type of economic evaluation, which tend to address uncertainty in a systematic way.51,52 Nevertheless, we found an increasing trend in the incorporation of uncertainty analysis in MCDA studies, suggesting that MCDA designers and analysts are responding to the voices calling to address uncertainty.3,4,23 In terms of the approaches adopted to deal with uncertainty, there seems to be a preference for one-way deterministic approach or scenario analysis, which is consistent with findings from Broekhuizen et al (2015).23 Structural uncertainty is barely explored, which might be due to the lack of methodological approaches to systematically deal with it23 or the cost of re-running analyses under different MCDA specifications (eg, another set of alternatives or criteria).

The possibility of involving all relevant parties in the decision-making process is one of the advantages that MCDA offers,12,27 and we found that different types of stakeholders have participated in the scoring and weighting steps. Nevertheless, > 50% of the studies were not explicit about the stakeholders involved in each one of the MCDA stages. Additionally, several studies have highlighted limitations in regard to stakeholders’ involvement. This includes the small number of participants,53-56 and the exclusion of relevant actors, such as patient representatives or caregivers.57-59 Given that big groups of participants might not be possible, some studies have suggested involving representatives of the relevant stakeholders when conducting an MCDA.60 Other studies also emphasize the need to include all relevant stakeholders from the beginning to ensure the relevance and usefulness of the MCDA.61,62

All flaws in the application and reporting across all MCDA steps described earlier compromise the transparency and legitimacy of
the MCDA-based recommendations. The lack of standardization of MCDA methods and their classification jeopardizes the comparability, credibility, and policy usefulness of MCDA studies in healthcare. This reinforces the need to clearly report the methods applied in each MCDA step. Even though some reports have been published to guide the application of MCDA in healthcare, the lack of a detailed methodological and reporting guidance might partly explain the questionable methodological quality of MCDA studies in healthcare. This guide could include (1) a universal classification of the MCDA methods; (2) a guidance on the application of each method, highlighting the particularities for each one of the different decision contexts in healthcare; (3) a common set of criteria or macro domains by decision context; and (4) a checklist with standards for reporting, similar to the Consolidated Health Economic Evaluation Reporting Standards (CHEERS) used in traditional economic evaluations. The development of this methodological and reporting guide should be based on empirical work and the experience accumulated by scholars and decision makers in healthcare.

Such a guidance could increase transparency and clarity in the application of MCDA and ultimately improve the credibility and systematic application of MCDA in different decision contexts in healthcare.

Limitations

First, we might have missed out non-English MCDA studies that have informed decisions in healthcare, particularly grey literature. Second, all searches were conducted by one reviewer, which might have affected the identification of relevant studies. Third, the inclusion criteria we adopted were less restrictive compared with the ones in Marsh et al (2014) with the risk of missing relevant scientific studies before 2013. To address this, we supplemented Marsh et al's studies with 2 other reviews with start review period as early as in the 90s and a similar scope to our study. Our time trend of MCDA studies published over the last years is similar to the one from Khan et al (2021), which suggests that the sharp increase observed from 2011 onward is not an artifact of our updating strategy (ie, starting academic searches in 2013).

Conclusions

Although the application of MCDA keeps increasing across different decision contexts in healthcare, the lack of quality and transparency is an obstacle to the adoption of MCDA widely in healthcare. Findings of this review could guide future MCDA applications in healthcare. Nevertheless, to support the usefulness and applicability of MCDA in healthcare, it is crucial to standardize the methods and reporting. A guide on how to conduct and report MCDA, which acknowledges the particularities of the different decision contexts and methods, needs to be developed.

Supplemental Material

Supplementary data associated with this article can be found in the online version at https://doi.org/10.1016/j.jval.2022.11.007.

REFERENCES
