PharmacoEconomics (2014) 32:345–365
DOI 10.1007/s40273-014-0135-0

REVIEW ARTICLE

Assessing the Value of Healthcare Interventions Using Multi-Criteria Decision Analysis: A Review of the Literature

Kevin Marsh · Tereza Lanitis · David Neasham · Panagiotis Orfanos · Jaime Caro

Published online: 7 February 2014

Abstract The objective of this study is to support those undertaking a multi-criteria decision analysis (MCDA) by reviewing the approaches adopted in healthcare MCDAs to date, how these varied with the objective of the study, and the lessons learned from this experience. Searches of EMBASE and MEDLINE identified 40 studies that provided 41 examples of MCDA in healthcare. Data were extracted on the objective of the study, methods employed, and decision makers' and study authors' reflections on the advantages and disadvantages of the methods. The recent interest in MCDA in healthcare is mirrored in an increase in the application of MCDA to evaluate healthcare interventions. Of the studies identified, the first was published in 1990, but more than half were published since 2011. They were undertaken in 18 different countries, and were designed to support investment (coverage and reimbursement), authorization, prescription, and research funding allocation decisions. Many intervention types were assessed: pharmaceuticals, public health interventions, screening, surgical interventions, and devices. Most used the value measurement approach and scored performance using predefined scales. Beyond these similarities, a diversity of different approaches were adopted, with only limited correspondence between the approach and the type of decision or product. Decision makers consulted as part of these studies, as well as the authors of the studies are positive about the potential of MCDA to improve decision making. Further work is required, however, to develop guidance for those undertaking MCDA.

Key Points for Decision Makers

Decision makers who participated in healthcare MCDAs were positive about their ability to improve decision making through the facilitation of knowledge transfer, ensuring all factors relevant to a decision are considered, and improving the communication of the rationale for decisions.

It is important that the design of the MCDA considers the objective of the analysis, the validity and reliability of the numerous MCDA approaches, and the resources available. Further guidelines are required to help achieve this balance.

Further research on MCDA techniques should include the development of guidelines on the assumptions underlying different approaches and their practical implications for decision makers, and further testing of the impact of different MCDA approaches on decision making.

1 Background

The valuation of healthcare interventions is relevant to several decisions—investment, authorization, reimbursement, resource allocation, and prescription. This valuation is challenging given the different effects, both beneficial

K. Marsh (✉) · T. Lanitis · D. Neasham · P. Orfanos
Evidera, 26–26 Hammersmith Grove, London W6 7HA, UK
e-mail: kevin.marsh@evidera.com

J. Caro
Evidera, 430 Bedford Street, Suite 300, Lexington Office Park, Lexington, MA 02420, USA

J. Caro
Faculty of Medicine, McGill University, Montreal, Canada

and adverse, and the fact that the same effects may be given greater value in a severe disease, or where no other intervention exists. Further difficulty arises from disagreement amongst diverse stakeholders. How can the value be captured and the implications of diverse opinions understood in a transparent, accountable, consistent, and reliable manner?

There is dissatisfaction with current approaches to this challenge. Given prominent drug withdrawals over the past decade, regulatory agencies are searching for more rigorous, consistent, and transparent approaches to benefit-risk assessment (BRA) [1, 2]. In 2006, the Institute for Medicine Report on Drug Safety recommended that the Centre for Drug Evaluation and Research at the US Food and Drug Administration develop a systematic approach to BRA [3]. As a consequence, enhancing BRA in regulatory decision making is now one of the Prescription Drug User Fee Act's Reauthorisation Performance Goals [4]. Similarly, the European Medicines Agency (EMA)'s Committee for Medicinal Products for Human Use set up a working group that recommended further research to develop BRA methods [5]. Accordingly, EMA initiated five work packages to develop tools and processes for BRA to support informed, science-based, regulatory decision making about medicinal products [6]. In addition, the Pharmacoepidemiological Research on Outcomes of Therapeutics by a European Consortium (PROTECT) program has also developed methodologies for BRA.[1]

Health authorities, patients, and clinicians are concerned about the narrowness of commonly used measures of the value of interventions. For instance, organizations such as the National Institute for Health and Care Excellence (NICE) have emphasized cost-effectiveness analysis to rule on market access, using the quality-adjusted life-year (QALY) to measure benefits. The QALY focuses on survival gained by an intervention, with some adjustment for changes in quality of life. It does not capture the full impact of an intervention and even more factors are considered in health technology assessments (HTAs), including: disease severity; treatment access; target population size; whether an intervention is curative or preventative; budgetary and other practical constraints, such as staff availability; economic impact; evidence quality; and/or political factors [7–10]. HTA bodies such as NICE point out that they do consider a wider set of outcomes through committee deliberation [11]. This approach has, however, been criticised for not being transparent or consistent [12].

These deficiencies have prompted proposals to use MCDA to support decisions because it has the potential to consider whatever criteria a decision maker judges relevant and, if done well, can support transparent and consistent decision making [1, 9, 12–14]. Though MCDA sometimes refers to structured group decision making [15], the definition adopted here is any method that establishes criteria, weights them in terms of importance, and scores each alternative on each criterion to create an overall assessment of value [13]. MCDA methods have been applied in many settings, including transport, environmental protection, construction, defence, and finance [16]. In healthcare, however, their application has been limited [14].

Several decision makers have shown interest in MCDA. EMA, for instance, has proposed MCDA as an approach to support BRA [5, 6]. The US Institute of Medicine has recently proposed an MCDA-based prioritization framework for vaccines [17]. The Institute for Quality and Efficiency in Health Care (IQWiG) in Germany has explored a form of MCDA—conjoint analysis and the analytical hierarchy process (AHP)—to weight the multiple endpoints considered in its assessments [18]. The Advisory Group for National Specialised Services (AGNSS) in the UK developed a MCDA-based decision-making framework to support reimbursement decisions for orphan drugs [19]. NICE, taking over responsibility for the assessment of orphan drugs, has continued AGNSS' emphasis on MCDA [20], also consulted on the role of MCDA in HTA [13]. Finally, it has been suggested that MCDA would be a pragmatic way to implement value-based pricing in the UK [21].

Despite the interest in MCDA, there is little guidance on its conduct in healthcare. Those performing MCDA to support healthcare decision making face a number of choices. There are different MCDA approaches—value measurement, outranking, and goal programming [14]—and the appropriate approach depends on the specific problem and user demands [16, 22].

Beyond selecting the type of MCDA, researchers face methodological choices at each step. First, which criteria to include? There is no 'rule' about the number to include, but it should be kept as low as is consistent with a well-founded decision [16]. A related concern is how the metrics currently used to assess health interventions, such as cost effectiveness, should be incorporated in MCDA? Second, how should interventions be scored on these criteria? This is of particular concern when the MCDA assesses multiple interventions across multiple criteria. How should all these data be collected in an efficient and yet unbiased manner? Third, which weighting techniques are most appropriate? Fourth, whose values should be used to weight criteria? Fifth, what is the best way to summarize the MCDA results to support decision making? Sixth, how should the impact of uncertainty be assessed?

The objective of this paper is to support those undertaking MCDA by reviewing existing MCDAs of healthcare interventions, summarizing the approaches adopted, how these varied with the objective of the study, and the lessons learned from this experience.

[1] See http://www.imi-protect.eu/benefitsRep.shtml.

2 Method

EMBASE and MEDLINE were searched on 19 August 2013 for English-language papers with no date restriction. The following search terms were used: 'multi-criteria decision*', 'multiple criteria decision*', 'MCDA', 'benefit risk assessment*', 'risk benefit assessment*', 'multicriteri* decision*', 'MCDM', 'multi-criteri* decision*'.[2] After duplicate removal, 1,865 titles were left. Titles and abstracts were reviewed independently by two reviewers. Studies were included in the review if they reported the application of MCDA to assess healthcare interventions, where MCDA was defined as above. Studies were excluded if they did not apply MCDA, such as discussion of how MCDA could be used; or did not evaluate healthcare interventions, such as MCDAs to assess level of health need in a locality. Disagreements between reviewers were resolved in a meeting with the rest of the research team, yielding 106 studies for full text review.

During the full text review, studies were excluded if they did not meet the criteria, were not available in full text format, did not complete all the steps in the MCDA definition adopted, including quantitative weighting of criteria, or did not report all stages of the MCDA. Studies were not excluded on the basis of methodological quality. The final sample size for the analysis was 41 MCDAs. These came from 40 studies, two of which [23, 24] each included two MCDAs, and two of which [25, 26] reported the same MCDA so this was only included once. See the PRISMA flowchart in Appendix 1.

There were fewer studies than identified by a recent bibliometric review of healthcare MCDAs [27]. This can be explained by three characteristics of the review undertaken here. First, only studies assessing specific healthcare interventions were included, whereas the bibliometric review had a broader scope, including, for instance, studies that evaluated healthcare disparities between regions or assessed the performance of hospitals. Second, MCDA is defined as including explicit quantitative weighting of criteria. Third, studies were only included in this review if they were titled MCDA or BRA, or variations upon these. That is, for instance, the searches were not defined to identify discrete choice experiments (DCE) where they were not described as MCDAs.

Data were extracted on: date of study; location (either jurisdiction for the research or, location of lead author if no particular jurisdiction); decision to inform; number of interventions assessed; type of MCDA; criteria included; approach used to arrive at these criteria; sources of data to measure criteria; whether criteria were scored and by who; weighting method employed; method used to assess uncertainty; decision makers' opinions on the advantages and disadvantages of MCDA; and comments from the authors on advantages and disadvantages of MCDA. Appendix 2 summarizes the key characteristics of the studies included in the review.

Two sets of analyses were undertaken. Descriptive statistics were generated on the frequency of different approaches, both across the whole sample and by location, type of decision being informed, and type of intervention assessed. Second, qualitative data on decision maker and author opinions were coded by: impact and usefulness of MCDA; criteria selection; criteria measurement and scoring; weighting criteria; and assessing uncertainty.

3 Results

3.1 Where and When Were MCDAs Undertaken?

The interest in healthcare MCDA is very recent (Fig. 1). While the first study was published in 1990, more than half were published since 2011. The MCDAs were undertaken in various locations (Fig. 2a). The 40 studies were undertaken in 18 locations, the most prevalent being: The Netherlands (7), USA (7), and the UK (4).

3.2 What Decisions Were MCDAs Informing?

The published MCDAs informed four decision types: investment—prioritization of interventions for coverage or reimbursement; prescription—selection of intervention; authorization—assessment for licensing; and research funding—allocation of research funds. The majority were concerned with investment decisions (Fig. 2b). This trend is observed in all geographic regions, except North America. In USA, only one of seven studies was investment focused.

Given their proprietary nature, it is not surprising that the review did not identify any examples of MCDA to support manufacturers' internal investment decisions. However,

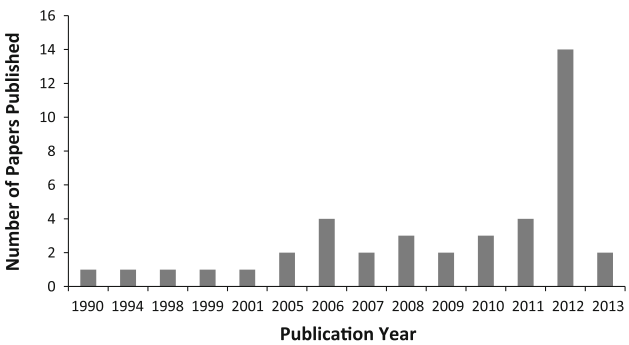

Fig. 1 Date of publication of the multi-criteria decision analysis

[2] The asterisk refers to any ending to the search term being acceptable.

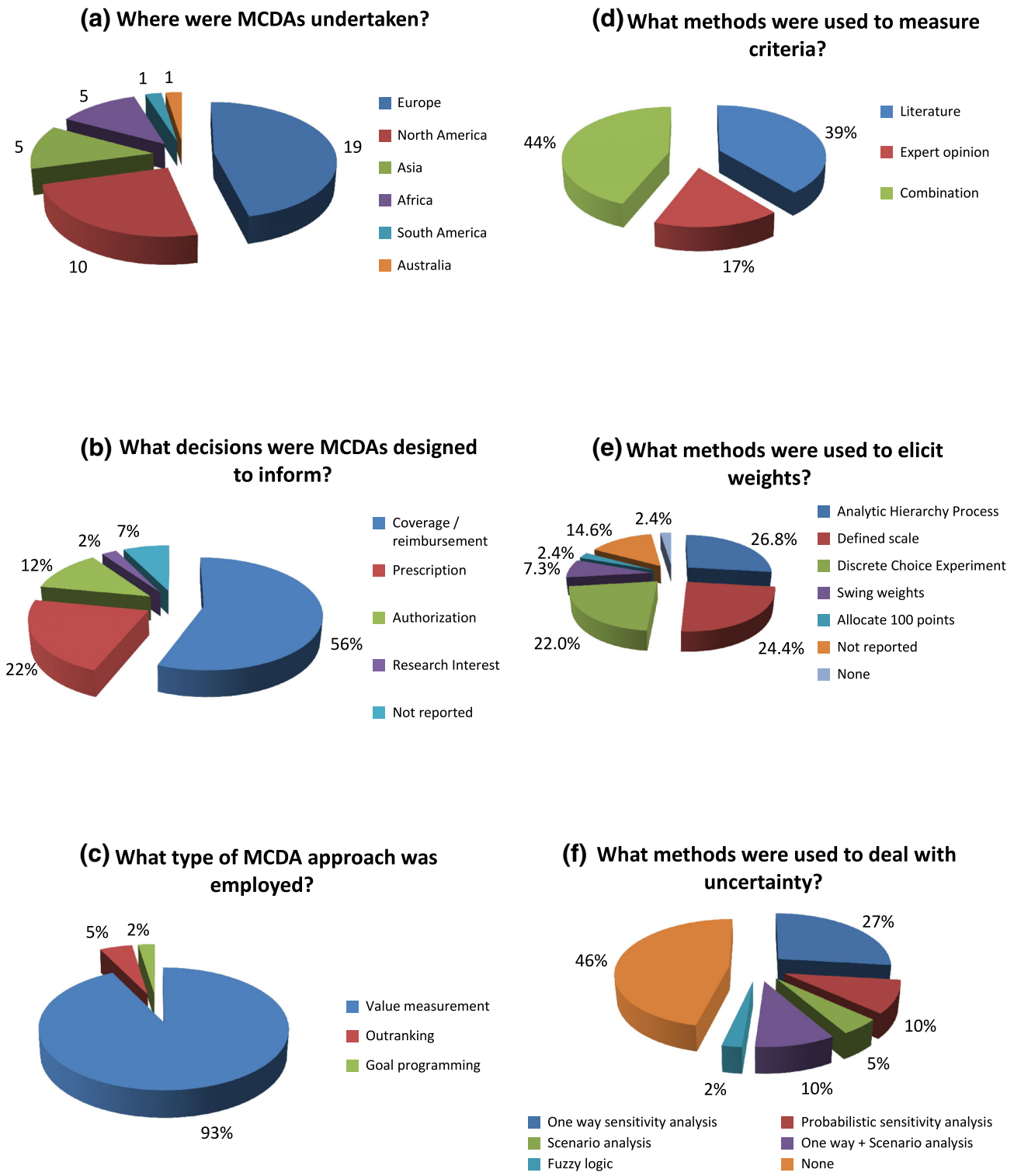

Fig. 2 Characteristics of multi-criteria decision analyses (MCDAs) in healthcare

two studies noted that MCDA could allow manufacturers to understand the benefits most preferred by patients and regulators, and ensure consistency in evidence collection across the development continuum [28, 29].

The MCDAs evaluated pharmaceuticals ($n = 17$), public health interventions (9), screening (7), surgical interventions (4), and devices (1). Three studies included interventions from more than one category [7, 30, 31].

The MCDAs evaluated an average of 13 interventions, with the minimum being one and the maximum 100. Investment-focused studies evaluated an average of 20 interventions; more than studies concerned with authorization (4), prescription (4), and research funding (2). More interventions were evaluated in studies of public health interventions (an average of 18) than screening (15), pharmaceuticals (11), devices (7), and surgical (3). This may be partly explained by the type of decision being informed. A larger number of interventions were considered in investment-focused studies in Africa, Asia, and Europe (34, 28, and 18, respectively) than in North America (4).

Two studies included interventions that were not being assessed but acted as 'anchors' [30, 32]. These were interventions that were already accepted as good practice and provided a benchmark, or measure of the decision's opportunity.

3.3 What Type of MCDA Approach Was Employed?

Three different types of MCDA are commonly distinguished: value measurement; outranking; and goal programming [14]. Value measurement models evaluation interventions based on an overall benefit score estimated as the weighted average of the criteria. Outranking models compare the alternatives pair wise, initially in terms of each criterion, to assert the extent of preference for one over the other for that criterion. The preference information across all criteria is aggregated to establish the strength of evidence favoring selection of one alternative over another. Goal programming involves derivation of the alternative(s) that are closest to achieving the pre-defined desirable (or satisfactory) levels of achievement for each criterion.

Most adopted the value measurement approach (Fig. 2c), with the output generated being primarily an overall weighted benefit score ($n = 30$), though other outputs were also used: ratio of weighted benefits to weighted risks (2); ratio of weighted benefits to costs (3); the probability that an intervention had the highest weighted benefit (1); and break-even points (1) (the value of each criterion that would result in an intervention being preferred to doing nothing). Most studies stopped at reporting these outputs, though some ($n = 10$) argued that they should be just the starting point for deliberation, as decision making should not be undertaken formulaically.

3.4 What Criteria Were Included?

An average of 8.2 criteria was used to assess interventions (range 3–19). There was little difference by type of decision: authorization (9.2), investment (8.7), and prescription (7.9). Although MCDA was designed to inform the allocation of research funds adopting a structure with 19 sub-criteria, several studies ($n = 19$) organized criteria into a hierarchy, splitting some criteria into sub-criteria.

Some studies ($n = 10$) adopted existing MCDA criteria sets, such as the EVIDEM framework (4). Other sources of criteria included the literature (23), stakeholders (20), or researchers themselves (5).

Authorization-focused studies measured clinical benefits and risks, as did prescription-focused studies, but a few also estimated intervention complexity and patient comfort [33, 34] and drug cost [34, 35]. The greatest range of criteria was adopted in investment-focused studies (Fig. 3).

Less than half ($n = 10$) of investment-focused studies included cost as a criterion, though a further three captured cost by including cost effectiveness as a criterion. Including cost in an MCDA implies that the criteria have a monetary value, though none of the authors discussed this. Study authors did identify challenges with incorporating cost effectiveness as a criterion. First, there is a risk that including both cost effectiveness and cost or effect "double counts", in the sense that the same feature of an intervention is credited more than once [36]. Four studies included cost and cost effectiveness, and six included effectiveness and cost effectiveness. Second, a cost-effectiveness criterion cannot capture differences in stakeholders' preferences for costs versus health effects—the same cost-effectiveness number can be generated by many combinations of costs and health effects [37].

Several criteria not conventionally included in economic analysis were considered by investment-focused studies (Fig. 3). These varied with the type of intervention. Studies of public health interventions were more likely to include productivity effects (25 % vs. 0 % for pharmaceuticals) and intervention complexity (57 % vs. 33 %), and less likely to include the disease severity (28 % vs. 82 %) and treatment access (29 % vs. 50 %). One study included the opportunity of the investment decision (the benefits that

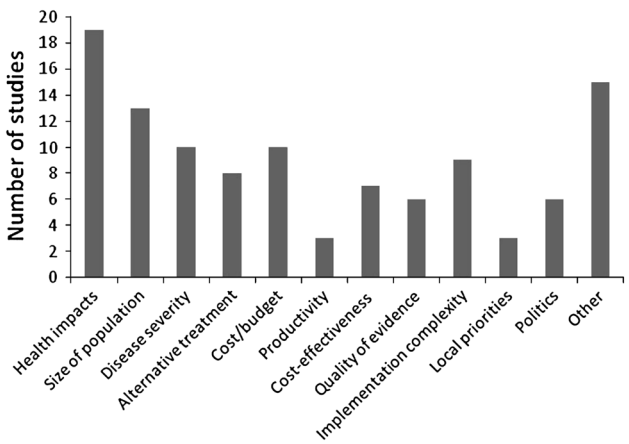

Fig. 3 Criteria included in multi-criteria decision analyses designed to support investment decisions ($n = 23$)

could be generated by allocating resources to a different intervention) as a criterion [38].

All investment-focused studies undertaken in North America ($n = 3$) adopted the EVIDEM framework and included disease severity, treatment access, costs, and cost effectiveness. Elsewhere, studies were less likely to include these criteria: disease severity (Europe 10 %, Asia 67 %, Africa 80 %), treatment access (Europe 18 %, Asia 33 %, Africa 20 %), costs (Europe 36 %, Asia 33 %, Africa 40 %), and cost effectiveness (Europe 18 %, Asia 33 %, Africa 20 %). Papers included little discussion on the process of defining criteria. Only one study explained that criteria were defined to meet MCDA requirements, such as avoiding double counting [39]. There was clearly a challenge in defining criteria because of the lack of established measures of factors such as health inequalities and socioeconomic groups [31, 40, 41]. Expert opinion is important in responding to this challenge, but further guidelines should also be developed [31].

3.5 How was Intervention Performance Measured and Scored?

Expert opinion, research evidence or both were used to measure intervention performance (Fig. 2d). Investment-focused studies (70 %) used expert opinion more frequently than authorization- (20 %) or prescription decision-focused (11 %) studies. Studies in Africa used expert opinion (80 %) more often than Europe (53 %), North America (40 %), and Asia (20 %).

Numerous sources of evidence were used: the literature ($n = 32$), analysis of trial data (3), economic modeling (3), and analysis of administrative data (1).

A key challenge with performance measurement identified in the studies was a lack of relevant data [9, 30, 31, 42], especially on disease severity, longer term economic impact, and implementation feasibility and acceptability. Expert opinion was recommended to fill these gaps [30, 31, 41, 43]. For instance, one study evaluating 100 different interventions in the Ivory Coast was unable to identify cost-effectiveness studies on all interventions [41]. Expert opinion was used, as undertaking an economic evaluation of each intervention would have been too time consuming.

There are several potential sources of bias in the evidence identified by the studies. Lack of resources required to gather data may lead to bias because of the selection of data based on convenience [44]. There is also lack of comparability because of the use of trials with different controls [39]. Mixed and indirect treatment comparisons would help, but were only used by one study [45].

Most studies ($n = 27$) used a scale to score intervention performance on criteria on equivalent scales. Most studies used a scale from 0 to 100, or another predefined scale, but six used the AHP, which uses a statistical analysis to derive the scores [33, 34, 46].The remainder (12) used the natural units without applying an additional scale. This was done when data were available using the same scale (e.g., rates per 100 patient-years [23] or probability of avoiding events [47]) or where DCE were used to generate weights [30, 42]. Even in these circumstances, scoring was still sometimes adopted when data were unavailable and interventions were measured against criteria using scores derived from experts.

There was disagreement on appropriate scoring methods. One study adopted a score between 0 and 100, rejecting a 0–3 scale using a qualitative description of levels as too laborious and difficult for participants to understand, and a simple 'high', 'medium', 'low' scoring as not providing sufficient discrimination [40]. Another study adopted a scale of 0–3, concluded that it had a "fair to good reproducibility' with 62.5 % of scores being identical in test–retest comparisons [36]. The authors noted, however, concern that the assumption of linearity is invalid, but warned against more complex approaches that raters may struggle to apply [36].

Study authors also expressed concern that raters may not understand some of the measures being used [31]. This was particularly the case where raters were from the general public and not familiar with measures. A related concern was that the packages of evidence provided to raters need to take into account the different demands for information from different types of raters [9, 36]. For instance, electronic tools can allow raters to access more evidence if desired, perhaps even reviewing source documents. Study authors also identified the concern that raters may also interpret scores differently, highlighting the importance of providing clear guidance on the meaning of each score [32]. Group discussions were identified as improving assessments of interventions by allowing participants to share information and expertise [28, 33, 43]. It was observed, however, that participants had no standard reference point to guide scoring, so the result has no comparative meaning outside the exercise [44].

3.6 How Were Criteria Weights Generated?

Numerous weighting techniques were employed (Fig. 2e). One study did not formally elicit weights from stakeholders, preferring instead to assume criteria had equal weights [31].

The large majority of prescription-focused studies adopted AHP (Fig. 4). Two studies designed to inform authorization decisions used weights to estimate a benefit score and a risk score, but did not combine these scores into an overall assessment of the benefit-risk balance [28, 48]. One proposed that a heuristic be developed to support

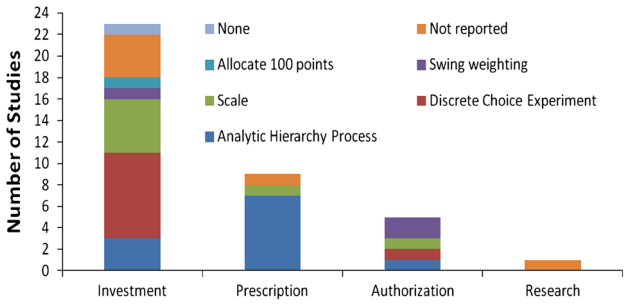

Fig. 4 Weighting method by decision type

manufacturers and regulators in judging whether a benefit-risk ratio is acceptable [28]. The other assumed that benefits and risks were equally important and undertook a sensitivity analysis to test the implications of this assumption [48].

In North America, the majority used either AHP (40 %) or defined scales (40 %), and none used DCE. This was not the case in other regions, with DCE being used by 60 % in Africa, 40 % in Asia, and 21 % in Europe.

Scoring the importance of criteria on defined scales, for instance, 0 for not important to 3 for very important, was the preferred technique for pharmaceuticals (35 % of studies) and screening interventions (43 % of studies), while most public health studies used DCE (37.5 %), and all surgical intervention studies used AHP.

Weights were elicited from many stakeholders, and some studies drew on more than one type of stakeholder: medical professionals ($n = 16$), payers and managers (16), patients (9), and researchers (8), and the public (2). Most studies using DCE (78 %) elicited weights from decision makers, for AHP medical professionals (55 %) or patients (45 %) were used, and for defined scales medical professionals (55 %) and decision makers were used (36 %).

Most investment-focused studies elicited weights from decision makers (50 %) or medical experts (25 %). None of the prescription- or authorization-focused studies elicited weights from decisions makers, preferring instead medical experts (50 % of prescription, 20 % of authorization) or patients (30 % of prescription, 40 % of authorization). These studies were also more likely to draw on researchers' preferences (40 % of authorization, 20 % of prescription, vs. 6.3 % of investment).

Probably reflecting the objectives of the study, the source of weights also varied with the type of health intervention. Medical experts were the preferred source for screening (50 %) and surgical (60 %) interventions; and decision makers for public health interventions (58 %). Weights for pharmaceuticals were elicited from decision makers (14 %), medical experts (33 %), patients (24 %), and researchers (29 %).

The preferences of the general public were only used in studies of public health investments that employed DCEs, and tended to be in Asia.

The studies identified a need for further work to develop weighting methods, including: the impact of different techniques [36, 49]; whose weights should be used [30, 42]; accommodating divergent weights [31, 39]; understanding of the assumptions underlying techniques, such as the independence of criteria weights [23, 36]; variation in weights between stakeholders [11, 12]; and whether interactions between weights need to be considered [37].

Two studies assessed the test–retest reliability of using expert opinion to score criteria on linear scales [9, 36]: 50.8 % and 65.1 % of weights retested identically. The authors speculate that variation in weights between tests may be the result of participants struggling to make implicit perceptions explicit [36]. They noted that this relatively simple weighting method may not capture weights accurately, but cautioned against adopting more complex methods. Other studies also noted challenges with more complex methods, such as the cognitive burden caused of a high number of attributes in DCEs [24, 37, 41]. Others emphasized the advantage of AHP over DCE in this regard, as it was not as limited in the number of attributes [25], but still others noted that AHP may lead to inconsistent rankings [48].

A number of challenges were observed with workshops, such as variations in weights as a result of the dominance of a particular viewpoint [32], and it being difficult to achieve weights that are representative of the populations whose preferences are of interest [39].

3.7 How was Uncertainty Dealt With?

Almost half the studies did not report an assessment of uncertainty (Fig. 2f); though about a quarter of these ($n = 5$) incorporated uncertainty as a criterion. Those studies that assessed uncertainty focused on the data inputs and did not consider structural uncertainty—for instance, in the criteria included. Most studies that addressed uncertainty used a one-way deterministic analysis ($n = 11$). Other approaches were: probabilistic sensitivity analysis (PSA) (4), a combination of one-way sensitivity and scenario analysis (4), scenario analysis alone (2), and fuzzy logic (1). One study called for Bayesian methods for PSA, drawing on expert beliefs [43], and another observed that uncertainty may be counted twice as raters already gave lower scores where there was more uncertainty [36].

3.8 How Useful is MCDA?

Where their opinions were elicited, decision makers were positive about MCDA, emphasizing its role in: providing a

systematic approach to decision making [9, 49]; facilitating knowledge transfer and improving decision makers' understanding of interventions [36, 50]; identifying data gaps [9, 36]; forcing decision makers to think through all relevant factors [9]; improving the transparency of decisions; and communicating their rationale [9, 36, 38, 50]. Most participants found the process of the MCDA acceptable, though a minority were concerned about the validity of using expert opinion and the challenge of conceiving of an "average" patient [40]. More concerning were: the time and resources required to implement the MCDA [9, 47, 50]; the challenge of absorbing a lot of information [38]; the need to participate in the process several times before participants were familiar with the process and able to contribute properly to capture its full utility [9, 36]; and challenges interpreting the MCDA outputs [9].

Decision makers also raised concerns about the generalizability of the results beyond the time and place of the specific decision being informed [33, 40]. This was mirrored by some of the study authors who argued that weights and scores were committee specific and an MCDA that attempts to be generalizable may ignore context-specific factors [29, 36, 51]. As a result, they felt that the MCDA should not be used formulaically but as the basis for deliberation that can ensure that context-specific factors are incorporated into an assessment [9, 24, 31, 32, 36, 42].

Study authors also identified the potential of the MCDA, noting that it is an improvement on current often ad hoc approaches to decision making that are often captured by interest groups [31]; it improves transparency and accountability of decision making [29, 37, 41, 49, 52]; it can simultaneously account for the trade-offs among multiple factors [37, 42, 51]; it packages a large amount of data in a way that is digestible, enabling manufacturers to communicate complex information [28, 29, 36]; it forces decision makers to think through all relevant issues [36, 52]; it reveals different perspectives, facilitating discussion and consensus generation [36, 41] including between manufacturers and regulators [28]; it helps to identify gaps in the data and prioritize data generation [28]; and it is readily updatable [29, 52].

3.9 What Impact Do MCDA Methods Have?

Two studies assessed the impact of MCDA methods. The first compared supporting patients' decisions with MCDA rather than with a simple educational intervention [26]. Measures of patients' decisional conflict—uncertainty about a course of action—suggested that MCDA improved patients' decisions about colorectal cancer screening. It did not, however, affect decision implementation, measured as adherence to a chosen treatment.

The second study assessed whether different MCDA methods produced different results [53]. Simple additive weighting (SAW—a value measurement approach) and technique for order preference by similarity to an ideal solution (TOPSIS—a goal programming approach) were found to produce the same conclusions. The study authors concluded that SAW may be better as it is easier to understand and perform, but noted the advantage of TOPSIS is that it allows an assessment of the extent to which a health intervention approaches an ideal.

4 Discussion

The objective of this paper was to summarize existing approaches for MCDA of healthcare interventions and to identify the lessons from this experience to support its ongoing application. For two reasons, the existing literature only provides limited guidance. First, the approaches and methods are very diverse, with limited correspondence between approaches employed and the type of healthcare decision or product evaluated. Second, there was little justification for the methods chosen, or discussion of the method's performance. There are, however, some trends in the literature and challenges that should be considered by those applying MCDA in healthcare.

First, most of the studies adopted the value measurement approach, but the reasons were unclear. Most advocates of MCDA in healthcare have recommended this approach [14] because most outcomes relevant to public decisions can be traded off [7], a requirement of the value measurement approach. The output from the value measurement approach is an overall benefit measure estimated as the weighted average of performance on the criteria. This approach assumes that a poor performance on one criterion can be compensated for by a good performance on another. There are, however, circumstances when this is not the case, such as where there are ethical issues making trade-offs unacceptable [7]. Precisely when this is the case is not easy to say. For instance, does a person's right to access to treatment constitute an ethical issue that undermines the compensatory assumption underlying the value measurement approach? Several studies included treatment access as a criterion, but did not address whether the compensatory assumption held. Further work is required to ensure that the assumptions underlying the different approaches are transparent, including guidance on the practical implications of these assumptions.

Second, there is a risk of double counting criteria. For instance, studies that applied the EVIDEM framework identified that costs and health effects were double counted as they are also included in cost effectiveness. Researchers should avoid this and other double counting when designing studies. We would support one study authors' observation that cost effectiveness should not be a criterion, as a single cost-effectiveness number can be generated by many combinations of costs and health effects, limiting its ability to capture stakeholders' preferences for costs and health effects. Further discussion is required to determine the appropriate way to include cost in the analysis and whether a decision should be expressed in net benefit terms when cost is a criterion. Some investment-focused studies included cost as a criterion. Other studies did not or compared the MCDA benefit score against cost in a production possibility frontier.

Third, assessing the interventions against the criteria is made difficult by gaps in the evidence, and using expert opinion to fill these is tenuous at best. Decisions makers are rightly concerned about the lack of evidence for scoring many interventions. Even where evidence is available, there is the potential for bias in the selection of data. More systematic reviews and a meta-analysis, including mixed treatment comparisons could overcome this concern. Although additional analyses to understand the data inputs may run up against another decision makers' concern, the time and resources required, this should not be a factor in any serious assessment of the healthcare decisions involved, given their potentially enormous impact.

Fourth, most studies scored interventions against the criteria, rather than adopting a weighting approach that is applied to criteria measured in natural units, such as DCE. The review identified a number of challenges with scoring. First, raters have different levels of understanding of the data and interpret scales differently. Group discussion can help participants to understand the scoring exercise and to ensure consistent interpretation of scales. Second, the complexity of scoring scales varies with, for instance, the number of points on the scale and whether they allow for non-linear preferences. It is recommended that the least complex scale is used that can capture preferences and discriminate between interventions.

Fifth, further debate and guidance are required on which weighting techniques are appropriate under which circumstances. Some studies discussed the advantages and disadvantages, but invariably focused on only one or two techniques, and addressed only a few of the relevant issues. Further work is required to ensure the reliability of approaches, including: whether the method elicits trade-offs between criteria or elicits opinions about the relative importance of criteria, whether the technique considers the magnitude of changes in criteria scores, whether the technique allows for non-linear preferences, and the cognitive feasibility of the elicitation exercise.

Sixth, more attention should be given to quantifying the impact of uncertainty. Almost half of the studies did not assess the impact of uncertainty. This is particularly important given the challenges with scoring interventions against criteria and consequent reliance on expert opinion. Further work is required to develop methods for assessing the impact of uncertainty in MCDAs. For instance, is one-way uncertainty analysis sufficient or should PSA be adopted? Is it appropriate to include uncertainty itself as a criterion in the MCDA? Should it be reflected in the scoring?

Seventh, decision makers expressed difficulty interpreting the MCDA output. The weighted average number generated by many MCDA approaches is "meaningless" aside from its use to rank the interventions [22]. Researchers should consider which of the types of output generated by different types of MCDA is most appropriate to their needs. For instance, where the objective is to communicate value to an audience beyond participants, it is recommend that a DCE might be the most appropriate weighting technique, as the output from this approach does have meaning outside of the analysis—for instance, the probability that a stakeholder group would prefer one intervention over another. Further work is required to provide guidance to researchers on this issue.

Finally, only two studies were identified that assessed the impact of MCDA methods. This effort should be broadened to understand which variations in MCDA approaches will have an impact on decisions.

It is important to acknowledge the limitations with this review. There were only a small number of relevant studies, reducing the ability to identify methodological trends. Restricting the review to English-language papers may have affected the types of studies identified, as schools of MCDA approaches differ across countries. For example, Lootsma [54] draws a distinction between French and American schools in MCDA. Unpublished studies are not captured; for instance, those employed by local decision makers or by manufacturers internally. The choice of search terms may have overlooked relevant applications that did not explicitly use the term MCDA, such as programme budgeting marginal analysis, DCEs, multi-objective optimization, or other non-compensatory approaches. Finally, the definition of MCDA adopted limited the review to those studies that explicitly weighted criteria.

5 Conclusion

The recent interest in MCDA in healthcare is mirrored in an increase in the application of MCDA to evaluate healthcare interventions. Decision makers consulted as part of these studies, as well as the authors of the studies are positive about the potential of MCDA to improve decision making. Current practice does not, however, provide sufficient guidance on which techniques should be adopted and whether this choice should depend on the type of healthcare interventions, the context, and other factors. Further work is required to develop this guidance.

Acknowledgments This study was conducted without external funding. All authors contributed to all aspects of the paper. Kevin Marsh is the guarantor for the overall content of this paper. The authors have no conflicts of interest that are directly relevant to the contents of this study.

Appendix 1

See Fig. 5.

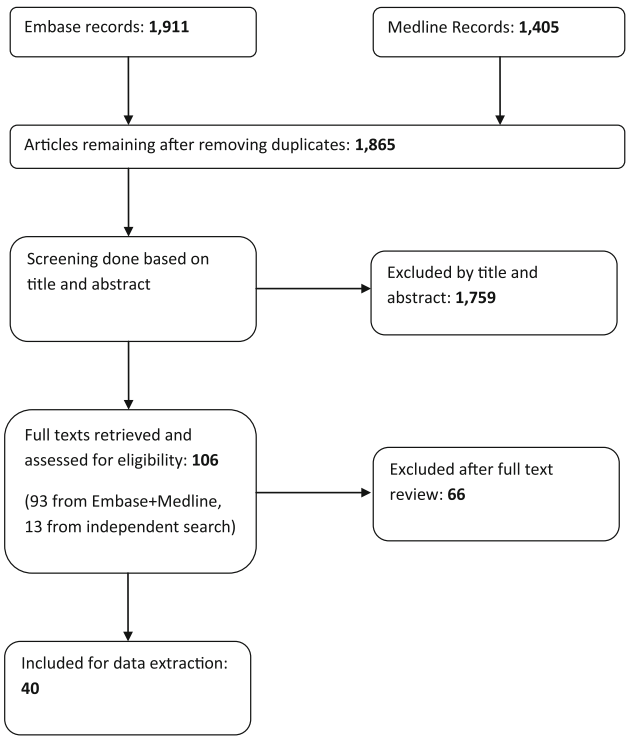

Fig. 5 PRISMA flowchart describing study selection

Appendix 2

See Table 1.

Table 1 Description of studies included in the review

References	Author/Year	Location	Decision = informed	MCDA type	Value measure	Criteria description	Scoring (source)	Weighting method (source)	Number of criteria/sub criteria (measurement source—identification source)[a]	Number of interventions (type)	Uncertainty assessment (type)
[40]	Airoldi et al. (2011)	UK	Investment			Health benefit to local population (number of patients who benefit, and the quality of life gained), health inequalities (reduction in geographically related, gender-related, access- and equality-related differences), Feasibility (probability of achieving what is meant to be achieved)	Yes (policy makers)	Swing weighing (patients, clinicians, policy makers, public)	4/5 (literature and expert opinion—workshop and published sources)	21 (public health)	Yes (deterministic)

Table 1 continued

References	Author/Year	Location	Decision = informed	MCDA type	Criteria description	Scoring (source)	Weighting method (source)	Number of criteria/sub criteria (measurement source—identification source)[a]	Number of interventions (type)	Uncertainty assessment (type)
[55]	Baeten et al. (2010)	Netherlands	Investment	Value measure	Cost effectiveness, severity of the disease, average age of targeted population, number of potential beneficiaries, individual health benefits, poverty reduction	No	DCE (policy makers)	6 (literature and expert opinion—published sources)	6 (screening)	No
[42]	Baltussen et al. (2007)	Nepal	Investment	Value measure	Severity of disease, number of potential beneficiaries, age of target group (proxy for economic benefit), individual health benefits, poverty reduction, cost effectiveness	No	DCE (clinician, policy makers)	6 (literature and modeling—workshop)	34 (public health)	No
[7]	Baltussen et al. (2006)	Ghana	Investment	Value measure	Cost-effectiveness (combined criterion of the following overlapping criteria: cost of treatment, effectiveness of treatment and cost effectiveness), poverty reduction, age of target group, severity of disease, health effect, total budget impact	No	DCE (policy makers)	6 (literature—workshop and published sources)	32 (pharmaceutical, screening)	No
[28]	Brass et al. (2013)	USA	Authorization	Value measure	Smoking cessation, smoking cessation attempts, smoking rates, increased adaptation of healthy lifestyles, Adverse events, combined use with smoking, use during pregnancy, abuse by adolescents, decreased success rates vs. healthcare provider supervised attempts	Yes (health researchers)	Scale (researchers)	2/9 (literature—published sources)	1 (public health)	No

Table 1 continued

References	Author/Year	Location	Decision = informed	MCDA type	Criteria description	Scoring (source)	Weighting method (source)	Number of criteria/sub criteria (measurement source—identification source)[a]	Number of interventions (type)	Uncertainty assessment (type)
[23]	Broekhuizen Henk (2012)	Netherlands	Authorization	Value measure	Response, remission, adverse events	No	AHP (patients)	3 (literature—published sources)	3 (pharmaceutical)	Yes (PSA)
					Progression-free survival, fatigue, diarrhea, hand-foot syndrome, mouth sores, liver failure, blood clot		DCE (patients)	7/9 (literature—published sources)	3 (pharmaceutical)	Yes (PSA)
[47]	Cunich et al. 2011	Australia	Prescription	Value measure	Relative risk, unnecessary biopsy, over diagnosis, bowel problems	No	Scale (clinicians)	4 (literature—published sources)	2 (screening)	No
[56]	Dalalah et al. (2008)	Jordan	Prescription		Duration of illness, infectious complications, antibiotic side effects, and antibiotics use	Yes (health researchers)	Not reported (not reported)	4/7 (literature—authors and published sources)	5 (pharmaceutical)	Yes (fuzzy logic)
[57]	Defechereux et al. (2012)	Norway	Investment	Value measure	Severity of the disease, number of potential beneficiaries, age of target group, health benefits, willingness to subsidise and cost-effectiveness	Yes (policy makers)	DCE (policy makers)	6 (expert opinion—published sources)	34 (public health)	No
[41]	Diaby et al. (2011)	Cote d'Ivoire	Prescription	Value measure	Cost effectiveness, severity of disease, social class affected	No	DCE (not reported)	4 (literature and expert opinion—workshop and published sources)	100 (pharmaceutical)	Yes (deterministic)
[35]	Diaz-Ledezma et al. (2013)	USA	Prescription	Value measure	Successful treatment of labral tears, of chondral lesions or of cam deformity, complications related to the surgical technique, cost, modified Harris hip score	Yes (health researchers)	AHP (researchers)	4/6 (literature—authors)	3 (surgery)	Yes (sensitivity analysis)

Table 1 continued

References	Author/Year	Location	Decision = informed	MCDA type	Criteria description	Scoring (source)	Weighting method (source)	Number of criteria/sub criteria (measurement source—identification source)[a]	Number of interventions (type)	Uncertainty assessment (type)
[25, 26]	Dolan et al. (2002, 2005)	USA	Prescription	Value measure	Colorectal cancer, major side effects, false positives, frequency, test preparations, test procedure	Yes (patients, family)	AHP (patients)	4/7 (literature-workshop)	6 (screening)	No
[48]	Dolan et al. (1994)	USA	Prescription	Value measure	Death and disability due to pulmonary and extra pulmonary infections, fatal side effect, nonfatal hepatitis, other physical side effects associated with isoniazid, and the inconveniences associated with isoniazid therapy	Yes (health researchers)	AHP (researchers)	2/4 (literature—not reported)	2 (pharmaceutical)	Yes (deterministic)
[34]	Erjaee et al. (2012)	Iran	Prescription	Value measure	Comfort in drug consumption, side effects of the drug, cost and efficacy	Yes (patients/family, clinicians)	AHP (patients, clinicians)	4 (trial and literature-authors)	3 (pharmaceutical)	No
[29]	Felli et al. (2009)	USA	Authorization	Value measure	Risks (carbohydrate metabolism, cancer, thyroid, skeletal, hepatic, other), tolerability and improper use, bBenefits (adult height gain, height velocity increase, skeletal maturation), life effects (juvenilisation, teasing and bullying, adult effects) and convenience (treatment administration, treatment burden, monitoring, acquisition and storage)	Yes (experts)	Swing weighting (researchers)	6/19 (literature and expert opinion—workshop and previous framework)	10 (pharmaceutical)	Yes (PSA)
[36]	Goetghebeur et al. (2012)	Canada	Investment	Value measure	EVIDEM framework criteria (Disease impact, context of intervention, intervention outcomes, type of benefit, economics, quality of evidence)	Yes (experts)	Scale (clinicians, policy makers, researchers)	6/15 (literature—EVIDEM framework)	10 (pharmaceutical)	No

Table 1 continued

References	Author/Year	Location	Decision = informed	MCDA type	Criteria description	Scoring (source)	Weighting method (source)	Number of criteria/sub criteria (measurement source—identification source)[a]	Number of interventions (type)	Uncertainty assessment (type)
[38]	Goetghebeur et al. (2010)	Canada	Investment	Value measure	EVIDEM framework criteria (quality of evidence, disease impact, intervention, economics, extrinsic value components, ethical framework, other components)	Yes (experts)	Scale (patients, clinicians, researchers)	6/15 (literature and expert opinion—EVIDEM framework)	1 (pharmaceutical)	No
[58]	Gonzalez-Zapata. (2009)	Europe	Investment	Value measure	Obesity, public sector costs, societal benefits, practical feasibility, social acceptability, health benefits, economic impact on individuals, economic impact unspecified, economic impact on commercial sector, others	Yes (policy makers)	Not reported (policymakers)	10 (expert opinion-workshop)	20 (public health)	Yes (deterministic)
[52]	Guest et al. (2012)	USA	Research	Value measure	Preclinical criteria, clinical criteria, criteria related to stakeholder	Yes (policy makers, experts)	Not reported (researchers)	3/19 (expert opinion—workshop and published sources)	2 (pharmaceutical)	No
[59]	Holdsworth et al. (2012)	Morocco and Tunisia	Investment	Value measure	Obesity, other health benefits, feasibility, social acceptability, cost, benefits for society	Yes (policy makers)	Not reported (policymakers)	6 (expert opinion—workshop)	12 (public health)	Yes (deterministic)
[43]	Hummel (2012)	Netherlands	Investment	Value measure	HRQoL (pain, spinal function and self-esteem), complication (probability and severity of medical and technical), cost	No	AHP (clinicians)	3/7 (literature-workshop)	1 (surgery)	Yes (deterministic)

Table 1 continued

References	Author/Year	Location	Decision = informed	MCDA type	Criteria description	Scoring (source)	Weighting method (source)	Number of criteria/sub criteria (measurement source—identification source)[a]	Number of interventions (type)	Uncertainty assessment (type)
[39]	Hummel et al. (2012)	Germany	Investment	Value measure	Efficacy (response, remission, no relapse), serious adverse events (suicide and attempted suicide, other), adverse events (sexual dysfunction, other), effects on quality of life (social function, no anxiety, no pain, cognitive function)	No	AHP (patients, clinicians)	4/11 (literature—workshop and published sources)	3 (pharmaceutical)	Yes (deterministic)
[60]	Hummel et al. (2005)	Netherlands	Prescription	Value measure	Ease of use (hindrance, time investment, maintenance, assistive devices), social acceptance (support dependence, aesthetics, unrevealing, burden social system), arm-hand function (function, quality grip, stability arm, availability), risks (reliability, certain improved results, minimal complications, safety), load of treatment (duration complete treatment, duration clinical intake, waiting period)	Yes (experts)	AHP (patients, clinicians)	5/19 (expert opinion-workshop)	2 (surgery)	Yes (deterministic)
[24]	Jehu-Appiah et al. (20080	Ghana	Investment	Value measure	Number of potential beneficiaries, severity of disease, cost-effectiveness, poverty reduction, vulnerable population	No	DCE (policymakers)	5 (literature and expert opinion-workshop)	26 (public health)	No
[44]	Kroese et al. (2010)	United Kingdom	Investment	Value measure	Morbidity and mortality, reproductive choice, process of care, deliverability, providing additional information	Yes (experts)	Allocation of percentage points (policy makers)	5 (expert opinion—published sources)	10 (screening)	No

Table 1 continued

References	Author/Year	Location	Decision = informed	MCDA type	Criteria description	Scoring (source)	Weighting method (source)	Number of criteria/sub criteria (measurement source—identification source)[a]	Number of interventions (type)	Uncertainty assessment (type)
[51]	Le Gales et al. (1990)	France	Investment	Outranking	Effectiveness, costs, technical feasibility, feasibility, ethical acceptability, information follow up, education about risks of transmission of hemoglobinopathies and on adhension to prevention in the general public at risk groups	Yes (experts)	Not reported (not reported)	7 (literature, expert opinion and database—workshop)	74 (screening)	Yes (deterministic)
[30]	Marsh et al. (2012)	United Kingdom	Investment	Value measure	Cost effectiveness, proportion of population eligible, equity: distribution of benefit	No	DCE (policy makers)	3 (literature and modelling—workshop and published sources)	14 (public health, pharmaceuticals)	No
[50]	Miot et al. (2012)	South Africa	Investment	Value measure	Quality of evidence (completeness and consistency of reporting evidence, relevance and validity of evidence), disease impact (disease severity, size of population affected by disease), intervention (guidelines, improvement of efficacy, safety, tolerability, public health interest, type of medical service), economics (budget impact model, cost effectiveness, other spending), other (appropriate use, organizational structure, political context, population priorities, regulatory status)	Yes (clinicians)	Scale (clinicians)	5/14 (literature and expert opinion—EVIDEM framework)	1 (screening)	No

Table 1 continued

References	Author/Year	Location	Decision = informed	MCDA type	Criteria description	Scoring (source)	Weighting method (source)	Number of criteria/sub criteria (measurement source—identification source)[a]	Number of interventions (type)	Uncertainty assessment (type)
[61]	Nobre et al. (1999)	Brazil	Investment	Outranking	Benefit population, social impact, availability of human resources, dependence of facilities, dependence on maintenance, community and professional demand, importance for patient condition, health outcomes	Yes (clinicians)	Not reported (clinicians)	8 (literature and expert opinion—not reported)	7 (devices)	Yes (Sensitivity analysis)
[62]	Perez – Encinas et al. (1998)	Spain	Not reported	Value measure	Efficacy (claudication, chest pain, hemodynamic parameters, number of patients in controlled clinical trials and The US Food and Drug Administration (FDA) approval), safety (hematologic disturbances), patient acceptance (dosage frequency, dose units, dose formulations), patient acceptance and cost	Yes (experts)	Scale (clinicians)	4/10 (literature and expert opinion—workshop and published sources)	4 (pharmaceutical)	Yes (deterministic)
[46]	Singh et al. (2006)	United States of America	Prescription	Value measure	Duration of illness, infectious complications (local complications and systemic complications), antibiotic side effect (minor and major), antibiotic use (undertreatment and over treatment)	Yes (clinicians)	AHP (clinicians)	4/7 (literature—published sources)	5 (screening)	Yes (deterministic)

Table 1 continued

References	Author/Year	Location	Decision = informed	MCDA type	Criteria description	Scoring (source)	Weighting method (source)	Number of criteria/sub criteria (measurement source—identification source)[a]	Number of interventions (type)	Uncertainty assessment (type)
[9]	Tony et al. (2011)	Canada	Investment	Value measure	Disease severity, population affected by disease, clinical guidelines, comparative interventions limitations, efficacy/effectiveness, patient-reported outcomes, public health interest, type of medical service, budget impact, cost effectiveness, other spending, completeness and consistency of evidence, relevance and validity of evidence	Yes (policy makers)	Scale (policy makers)	6/14 (trial, literature and expert opinion—workshop and published sources)	1 (pharmaceutical)	No
[45]	Valkenhoef et al. (2012)	Non-geographic specific	Authorization	Value measure	Hamilton rating scale of depression, diarrhea, dizziness, headache, insomnia, nausea, common risks	No	Ordinal swing weighing (partial) (clinicians)	6 (literature, trial and modelling—published sources)	5 (pharmaceutical)	Yes (PSA)
[33]	Van Til et al. (2008)	Netherlands	Prescription	Value measure	Outcome (4 sub-criteria), comfort, risk and side effect (short-term, long-term and reliability), cosmetics, effect (2 sub criteria), impact (duration and complexity)	Yes (clinicians)	AHP (clinicians)	6/13 (expert opinion—workshop and published sources)	6 (surgery)	Yes (deterministic)
[53]	Van-Wijk et al. (2006)	Netherlands	Not reported	Value measure	Effectiveness, persistence with treatment as a measure of tolerability, cost, and clinical experience	Yes (experts)	Scale (clinicians)	4 (literature—authors and published sources)	5 (pharmaceutical)	Yes (deterministic)
[53]	Van-Wijk et al. (2006)	Netherlands	Not reported	Goal program	Effectiveness, persistence with treatment as a measure of tolerability, cost, and clinical experience	Yes (experts)	Scale (researchers)	4 (literature—authors and published sources)	5 (pharmaceutical)	Yes (deterministic)

Table 1 continued

References	Author/Year	Location	Decision = informed	MCDA type	Criteria description	Scoring (source)	Weighting method (source)	Number of criteria/sub criteria (measurement source—identification source)[a]	Number of interventions (type)	Uncertainty assessment (type)
[63]	Wenstop et al. (2001)	Norway	Investment	Value measure	HIV cases, privacy, general anxiety, personal anxiety, stigmatization of individuals, stigmatization of population sub groups, cost, early treatment	No	Pairwise comparisons of interventions (policy makers)	9 (literature and expert opinion—workshop)	4 (public health)	No
[32]	Wilson et al. (2007)	United Kingdom	Investment	Value measure	Effectiveness, burden, equity and fairness, deliverability, engagement, acceptability, certainty, national standards	Yes (policy makers)	Allocation of 100 points (policy makers)	8 (literature and expert opinion-workshop)	6 (public health)	Yes (deterministic)
[49]	Youngkong et al. (2012)	Thailand	Investment	Value measure	Groups targeted (children, teenagers, adults, high-risk adults, gender), intervention type (prevention, treatment of HIV, and treatment of AIDS) effectiveness, quality of evidence	Yes (policy makers, health researchers)	DCE (patients, policy makers, public)	5 (literature and expert opinion—published sources)	40 (pharmaceutical)	No
[31]	Youngkong et al. (2012)	Thailand	Investment	Value measure	Prevalence, severity of disease, effectiveness, variation in practice, impact on household expenditure, equity/ethical and social implications	Yes (policy makers, patients/family, clinicians, public s, health researchers, industry)	Assumed equal weights	6 (literature—workshop and published sources)	1 (pharmaceutical)	No

MCDA multi-criteria decision analysis, *BRA* benefit-risk assessment, *AHP* analytic hierarchy process, *DCE* discrete choice experiment, *CE* cost effectiveness

[a] When only one number is reported, that signifies that the criteria considered in the study were not divided to sub-categories

References

1. Guo JJ, Pandey S, Doyle J, Bian B, Lis Y, Raisch DW. A review of quantitative risk-benefit methodologies for assessing drug safety and efficacy: report of the ISPOR risk-benefit management working group. Value Health J Int Soc Pharmacoecon Outcomes Res. 2010;13(5):657–66.
2. Garrison LP. Regulatory benefit-risk assessment and comparative effectiveness research: strangers, bedfellows or strange bedfellows? Pharmacoeconomics. 2010;28(10):855–65.
3. The future of drug safety: promoting and protecting the health of the public. Institute of Medicine, 2006.
4. Food and Drug Administration (FDA). PDUFA reauthorization performance goals and procedures fiscal years 2013 through 2017. http://www.fda.gov/downloads/forindustry/userfees/prescriptiondruguserfee/ucm270412.pdf. Accessed 30 April 2013.
5. European Medicines Agency (EMA). Report of the CHMP working group on benefit-risk assessment models and methods; 2007.
6. European Medicines Agency (EMA). Benefit-risk methodology project. Work package 4 report: benefit-risk tools and processes; 2012.
7. Baltussen R, Niessen L. Priority setting of health interventions: the need for multi-criteria decision analysis. Cost Eff Resour Alloc. 2006;4:14.
8. Golan O, Hansen P, Kaplan G, Tal O. Health technology prioritization: which criteria for prioritizing new technologies and what are their relative weights? Health Policy. 2011;102(2–3):126–35.
9. Tony M, Wagner M, Khoury H, Rindress D, Papastavros T, Oh P, et al. Bridging health technology assessment (HTA) with multicriteria decision analyses (MCDA): field testing of the EVIDEM framework for coverage decisions by a public payer in Canada. BMC Health Serv Res. 2011;11:329.
10. Guindo LA, Wagner M, Baltussen R, Rindress R, et al. From efficacy to equity: literature review of decision criteria for resource allocation and healthcare decision making. Cost Eff Resour Alloc. 2012;10:9. http://www.resource-allocation.com/content/10/1/9.
11. Rawlins M, Barnett D, Stevens A. Pharmacoeconomics: NICE's approach to decision-making. Br J Clin Pharmacol. 2010;70(3):346–9.
12. Devlin NJ, Sussex J. Incorporating multiple criteria in HTA. Methods Processes. London: Office of Health Economics; 2011. http://www.ohe.org/publications/article/incorporating-multiple-criteria-in-htamethods-and-processes-8.cfm.
13. National Institute for Health and Care Excellence (NICE). Briefing paper for methods review workshop on structured decision making. 2012. http://www.nice.org.uk/media/C67/40/TAMethodsGuideReviewSupportingDocuments.pdf. Accessed 30 April 2013.
14. Thokala P, Duenas A. Multiple criteria decision analysis for health technology assessment. Value Health J Int Soc Pharmacoecon Outcomes Res. 2012;15(8):1172–81.
15. Cross JT, Garrison LP. Challenges and opportunities for improving benefit-risk assessment of pharmaceuticals from an economic perspective. London: Office of Health Economics; 2008.
16. Communities and Local Government (CLG). Multi-criteria analysis: a manual; 2009.
17. Institute of Medicine. Ranking vaccines a prioritization framework phase I: demonstration of concept and a software blueprint. In: National Academy of Sciences, editor. Washington DC: National Academies Press; 2012.
18. Drummond M, Bridges JFP, Muhlbacher A, Ijzerman MJ. Identification, weighting and prioritization of multiple endpoints for comparative effectiveness research: what have we learned from Germany? ISPOR Panel 2011; 2011.
19. National Health Service (NHS). Decision-making framework for making recommendations on national commissioning. http://www.specialisedservices.nhs.uk/document/decision-making-framework-making-recommendations-on-national-commissioning. Accessed 30 April 2013.
20. National Institute for Health and Care Excellence (NICE). Interim Process and Methods of the Highly Specialised Technology Programme. 2013. http://www.nice.org.uk/media/188/49/HST_combined_Interim_Process_and_Methods_FINAL_31_May_2013.pdf. Accessed August 2013.
21. Sussex J, Towse A, Devlin N. Operationalizing value-based pricing of medicines: a taxonomy of approaches. Pharmacoeconomics. 2013;31(1):1–10.
22. De Montis A, De Toro P, Droste-Franke B, et al. Criteria for quality assessment of MCDA methods. In: Getzener M, Spash C, Stagl S, editors. Alternatives for valuing nature: London: Routledge; 2005: 99–133.
23. Broekhuizen H, Ijzerman MJ, Groothuis-Oudshoorn KG, Hauber AB. Integrating patient preferences and clinical trial data in a Bayesian model for quantitative risk-benefit assessment. Value Health. 2012;15(7):A474.
24. 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 J Int Soc Pharmacoecon Outcomes Res. 2008;11(7):1081–7.
25. Dolan JG. Patient priorities in colorectal cancer screening decisions. Health Expect Int J Public Particip Health Care Health Policy. 2005;8(4):334–44.
26. Dolan JG, Frisina S. Randomized controlled trial of a patient decision aid for colorectal cancer screening. Med Decis Mak. 2002;22(2):125–39.
27. Diaby V, Campbell K, Goeree R. Multi-criteria decision analysis (MCDA) in health care: a bibliometric analysis. Oper Res Health Care. 2013;2:20–4.
28. Brass EP, Lofstedt R, Renn O. A decision-analysis tool for benefit-risk assessment of nonprescription drugs. J Clin Pharmacol. 2013;53(5):475–82.
29. Felli JC, Noel RA, Cavazzoni PA. A multiattribute model for evaluating the benefit-risk profiles of treatment alternatives. Med Decis Mak. 2009;29(1):104–15.
30. Marsh K, Dolan P, Kempster J, Lugon M. Prioritizing investments in public health: a multi-criteria decision analysis. J Public Health (Oxf). 2012
31. 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):961–70.
32. Wilson E, Sussex J, Macleod C, Fordham R. Prioritizing health technologies in a Primary Care Trust. J Health Serv Res Policy. 2007;12(2):80–5.
33. van Til JA, Renzenbrink GJ, Dolan JG, Ijzerman MJ. The use of the analytic hierarchy process to aid decision making in acquired equinovarus deformity. Arch Phys Med Rehabil. 2008;89(3):457–62.
34. Erjaee A, Bagherpour M, Razeghi S, Dehghani SM, Imanieh MH, Haghighat M. A multi-criteria decision making model for treatment of *Helicobacter pylori* infection in children. Hong Kong J Paediatr. 2012;17(4):237–42.
35. Diaz-Ledezma C, Parvizi J. Surgical approaches for cam femoroacetabular impingement: the use of multicriteria decision analysis. Clin Orthop Relat Res. 2013;471:2509–16.
36. Goetghebeur MM, Wagner M, Khoury H, Levitt RJ, Erickson LJ, Rindress D. Bridging health technology assessment (HTA) and

efficient health care decision making with multicriteria decision analysis (MCDA): applying the EVIDEM framework to medicines appraisal. Med Decis Mak. 2012;32(2):376–88.
37. Baltussena R, Stolka E, Chisholmc D, Aikinsd M. Towards a multi-criteria approach for priority setting: an application to Ghana. Health Econ. 2006;15:689–96.
38. Goetghebeur MM, Wagner M, Khoury H, et al. Combining multicriteria decision analysis, ethics and health technology assessment: applying the EVIDEM decision making framework to growth hormone for Turner syndrome patients. Cost Eff Resour Alloc. 2010;8.
39. Hummel MJM, Volz F, Van Manen JG, Danner M, Dintsios CM, Ijzerman MJ, et al. Using the analytic hierarchy process to elicit patient preferences: prioritizing multiple outcome measures of antidepressant drug treatment. Patient. 2012;5(4):225–37.
40. Airoldi M, Morton A, Smith J, Bevan G. Healthcare prioritisation at the local level: a socio-technical approach. Priority Setting for Population Health; Working paper series; London School of Economics and Political Science. 2011; Working paper no.7.
41. Diaby V, Lachaine J. An application of a proposed framework for formulary listing in low-income countries: the case of Cote d'Ivoire. Appl Health Econ Health Policy. 2011;9(6):389–402.
42. Baltussen R, Ten Asbroek AHA, Koolman X, Shrestha N, Bhattarai P, Niessen LW. Priority setting using multiple criteria: should a lung health programme be implemented in Nepal? Health Policy Plan. 2007;22(3):178–85.
43. Hummel JM, Boomkamp ISM, Steuten LMG, Verkerke BGJ, Ijzerman MJ. Predicting the health economic performance of new non-fusion surgery in adolescent idiopathic scoliosis. J Orthop Res. 2012;30(9):1453–8.
44. 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(7–8):538–43.
45. Van Valkenhoef G, Tervonen T, Zhao J, De Brock B, Hillege HL, Postmus D. Multicriteria benefit-risk assessment using network meta-analysis. J Clin Epidemiol. 2012;65(4):394–403.
46. Singh S, Dolan JG, Centor RM. Optimal management of adults with pharyngitis: a multi-criteria decision analysis. BMC Med Inform Decis Mak. 2006;6:14.
47. Cunich M, Salkeld G, Dowie J, Henderson J, Bayram C, Britt H, et al. Integrating evidence and individual preferences using a web-based multi-criteria decision analytic tool: an application to prostate cancer screening. Patient. 2011;4(3):153–62.
48. Dolan JG, Bordley DR. Isoniazid prophylaxis: the importance of individual values. Med Decis Mak. 1994;14(1):1–8.
49. Youngkong S, Teerawattananon Y, Tantivess S, Baltussen R. Multi-criteria decision analysis for setting priorities on HIV/AIDS interventions in Thailand. Health Res Policy Syst BioMed Cent. 2012;10:6.
50. Miot J, Wagner M, Khoury H, et al. Field testing of a multicriteria decision analysis (MCDA) framework for coverage of a screening test for cervical cancer in South Africa. Cost Eff Resour Alloc. 2012;10.
51. Le Gales C, Moatti JP. Searching for consensus through multicriteria decision analysis: assessment of screening strategies for hemoglobinopathies in southeastern France. Int J Technol Assess Health Care. 1990;6(3):430–49.
52. Guest J, Harrop JS, Aarabi B, Grossman RG, Fawcett JW, Fehlings MG, et al. Optimization of the decision-making process for the selection of therapeutics to undergo clinical testing for spinal cord injury in the North American Clinical Trials Network. J Neurosurg Spine. 2012;17(1 Suppl):94–101.
53. Van Wijk BL, Klungel OH, Heerdink ER, de Boer A. A comparison of two multiple-characteristic decision-making models for the comparison of antihypertensive drug classes: simple additive weighting (SAW) and technique for order preference by similarity to an ideal solution (TOPSIS). Am J Cardiovasc Drugs Drugs Devices Interv. 2006;6(4):251–8.
54. Lootsma FA. The French and the American school in multi-criteria decision analysis. Revue française d'automatique, d'informatique et de recherche opérationnelle Recherche opérationnelle. 1990;24(3):263–85.
55. Baeten SA, Baltussen RMPM, Uyl-de Groot CA, Bridges J, Louis WN. Incorporating equity-efficiency interactions in cost-effectiveness analysis: three approaches applied to breast cancer control. Value Health. 2010;13(5):573–9.
56. Dalalah D, Magableh S. A remote fuzzy multicriteria diagnosis of sore throat. Telemed e-Health. 2008;14(7):656–65.
57. Defechereux T, Paolucci F, Mirelman A, Youngkong S, Botten G, Hagen TP, et al. Health care priority setting in Norway: a multicriteria decision analysis. BMC Health Serv Res. 2012;12:39.
58. Gonza lez-Zapata LI, Alvarez-Dardet C, Ortiz-Moncada R, et al. Policy options for obesity in Europe: a comparison of public health specialists with other stakeholders. Public Health Nutr. 2009;12(7):896–908.
59. Holdsworth M, El Ati J, Bour A, et al. Developing national obesity policy in middle-income countries: a case study from North Africa. Health Policy Plann. 2012;1–13.
60. Marjanhummel JM, Snoek GJ, van Til JA, van Rossum W, IJzerman MJ. A multicriteria decision analysis of augmentative treatment of upper limbs in persons with tetraplegia. J Rehabil Res Dev. 2005;42(5):635–44.
61. Nobre FF, Trotta LTF, Gomes LFAM. Multi-criteria decision making: an approach to setting priorities in health care. Stat Med. 1999;18(23):3345–54.
62. Pérez Encinas M, Fernández MA, Martín ML, et al. Multicriteria decision analysis for determining drug therapy for intermittent claudication. Meth Find Exp Clin Pharmacol 1998;20(5):425–431.
63. Wenstop F, Magnus P. Value focused rationality in AIDS policy. Health Policy. 2001;57(1):57–72.