CrossMark

ORIGINAL RESEARCH ARTICLE

Applying a Multicriteria Decision Analysis (MCDA) Approach to Elicit Stakeholders' Preferences in Italy: The Case of Obinutuzumab for Rituximab-Refractory Indolent Non-Hodgkin Lymphoma (iNHL)

Martina Garau[1] · Grace Hampson[1] · Nancy Devlin[1] · Nicola Amedeo Mazzanti[2] · Antonio Profico[2]

Abstract

Background Healthcare decision makers need to make trade-offs between different elements of value of new treatments. Multicriteria decision analysis (MCDA) provides a framework that can help decision makers to understand stakeholders' preferences and be explicit about the trade-offs that are being made.

Objective The objective of this study was to use MCDA to obtain preferences and views on decision criteria across three stakeholder groups (patients, clinicians and payers) in Italy and to use these to assess the performance of obinutuzumab for rituximab-refractory indolent non-Hodgkin lymphoma (iNHL).

Methods We used EVIDEM V3.0, an MCDA framework, and collected participants' preferences via an online survey and structured meetings.

Results Patients and clinicians expressed a preference for interventions targeting severe conditions. Payers expressed preference for treatments targeting areas with an unmet need, which are cheaper than the comparator, and with high-quality evidence. Obinutuzumab in combination with bendamustine, compared with bendamustine alone, received high positive scores for the criteria 'disease severity' and 'type of therapeutic benefit' by all three groups, and negative scores on the economic-related criteria, according to all stakeholder groups.

Conclusions MCDA can be used to elicit the views of different stakeholder groups and has the potential to structure and inform reimbursement decisions.

Key Points for Decision Makers

Multicriteria decision analysis (MCDA) provides a framework that can help decision makers to understand stakeholders' preferences and be explicit about the trade-offs that are being made between different elements of value.

Based on a convenience sample, we found that patients and clinicians give greater weight to targeting severe conditions; payers are most concerned with unmet need, comparative costs, and high-quality evidence.

The overall value score of obinutuzumab based on all stakeholder groups' responses combined was mainly driven by the criteria disease severity, type of therapeutic benefit and unmet needs.

✉ Martina Garau
  mgarau@ohe.org

[1] Office of Health Economics, London, UK

[2] Roche S.P.A, Milan, Italy

1 Background

Stakeholders such as patients can be involved in medicine reimbursement decisions and Health Technology Assessment (HTA) in different ways, from contributions to evidence submissions to participation in advisory committees.

In some HTA processes, such as that of the National Institute for Health and Care Excellence (NICE) in

England and Wales, and the Canadian Agency for Drugs and Technologies in Health (CADTH) in Canada, patients submit evidence and participate in committee meetings [1]. The Scottish Medicines Consortium (SMC) in Scotland has introduced the option of consulting a panel of clinicians and patient representatives during the assessment of medicines for rare or end-of-life conditions [2].

HTA decision makers need to understand how stakeholders trade-off between different types of benefit potentially generated by a new treatment. Stakeholders need to understand to what extent and how their input affects final HTA decisions. A study of appraisals in five countries [3] found that patient preferences are rarely mentioned in HTA reports, and that patient participation in HTA tend to be unsystematic.

Multicriteria decision analysis (MCDA) provides a clear framework to assess the value of a treatment compared with alternative treatments or standard of care (SoC), against multiple and competing criteria [4]. It can support decision makers to be explicit about the trade-offs made between the selected criteria, and offers systematic and robust ways to elicit preferences and consider evidence from stakeholders. For example, in Germany, the Institute for Quality and Efficiency in Health Care (IQWiG) ran MCDA pilots to elicit patient preferences on clinical outcomes in depression and hepatitis C [5, 6]. In Italy, the Lombardia region uses an MCDA framework to select health technologies to reimburse [7].

The purpose of this study was to use an MCDA approach to obtain preferences and views on criteria and on performance of obinutuzumab for rituximab-refractory indolent non-Hodgkin lymphoma (iNHL) across three stakeholder groups (patients, clinicians and payers) in Italy. As far we know, this is the first study exploring stakeholders' preferences to inform national reimbursement decisions of medicines in Italy. While Castro et al. describe the use of EVIDEM in Italy in the context of a regional payer and focused on medical devices, we conducted our MCDA exercise to inform value assessment and decision making at the national level, specifically in the context of medicines [8], which is currently under the responsibility of Agenzia Italiana del Farmaco (AIFA[1]).

Our work built on the study by Wahlster et al. [9], who applied EVIDEM to capture stakeholders' preferences in Germany using an intervention case study. In addition to the online survey that was used to obtain preferences individually by Wahlster et al., we applied EVIDEM to support group discussions. This is a fundamental role that MCDA can play in HTA to structure committees' consideration of different and often conflicting perspectives.

2 Methods

There are number of approaches to eliciting preferences with different levels of complexity and theoretical bases, including swing weights, analytic hierarchy process and discrete choice experiments [6]. Recently, there has also been a proliferation of frameworks to assess the value of new interventions, developed by organisations including the American Society of Clinical Oncology (ASCO), the European Society for Medical Oncology (ESMO) and the Institute for Clinical and Economic Review (ICER) [10]. For the purposes of this study, we selected the EVIDEM approach, which combines a coherent set of criteria with tools to operationalise it in decision making (including suggested approaches for preference elicitation). In addition, EVIDEM was specifically designed for healthcare decision making. Its choice here enables results to be compared with a growing literature from its use in other healthcare decision-making settings.

EVIDEM (V3.0)[2] is an open-source framework resulting from a collaboration of experts and stakeholders, and is subject to continued testing and adaptations [11]. In the last 10 years, the EVIDEM collaboration has merged the practical aspects of developing an implementable framework with ethical foundations for its criteria and objectives. In particular, ethical justification has been provided to the criteria, including "distributive justice and fairness (prioritise those who are worst off)" [11]. EVIDEM also advocates for important procedural values such as 'participatory decision making', which aligns with the ethical framework of accountability for reasonableness developed by Daniels and Sabin for 'legitimate' priority settings [11, 12].

EVIDEM comprises a broad range of criteria, capturing elements of value relevant to patients, healthcare systems and society, with properties (including non-overlap between criteria), and with operational purposes, enabling us to inform real-life decision making. EVIDEM has been applied in HTA in a number of jurisdictions, including Canada [13] and, more recently, Spain, where the HTA body in Catalonia has explored the framework for appraising orphan drugs [14]. The Italian region Lombardia uses EVIDEM to inform local funding decisions of health interventions. Our aim was to extend the application of MCDA to support national decision making of medicine reimbursement.

EVIDEM includes a set of 13 clearly defined and measurable 'core' criteria grouped into domains. Some criteria are measured in absolute terms, not relative to other interventions, and other criteria are measured comparatively to existing interventions.

---

[1] http://www.agenziafarmaco.gov.it/en.

[2] https://www.evidem.org/.

Figure 1 presents the core EVIDEM framework, including five domains: need for intervention, comparative outcomes of intervention, type of benefit, economic consequences of intervention, and knowledge about intervention. Each domain comprises a set of criteria.

In May 2016, an online survey (using SurveyMonkey®) was sent to three groups of stakeholders involved or affected by reimbursement decisions in Italy: clinicians, patients, and payers. This step of the process was labelled as the 'survey first round'. The survey was used to elicit preferences around the relative importance of criteria (weights), and the degree of achievement of obinutuzumab against these criteria compared with the current SoC (scores).

We derived weights using the 'point allocation' approach, where we asked participants to allocate 100 points, first across criterion domains and, second, across criteria within each clusters. This method was selected because it combines simplicity with the ability of 'forcing' people to prioritise criteria. We also noted that van Til et al. [15] showed that the choice of the weight elicitation method does not affect value estimates at the group level. The description of the criteria and the survey instructions given to participants are available in the electronic supplementary material. To obtain one set of weights for each criterion, we combined domain weights with those within each domain and normalised the values (to sum up to 1).

Respondents were asked to score the performance of obinutuzumab in combination with bendamustine followed by obinutuzumab maintenance, compared with bendamustine alone (which is the only efficacious intervention in this indication [16]), in patients with rituximab-refractory iNHL. This is in line with obinutuzumab license indication and clinical evidence [16, 17].

Incremental criteria, related to the health and non-health effects of the intervention, were measured on a scale from

Fig. 1 Structure of the MCDA EVIDEM framework, including all criteria and categories
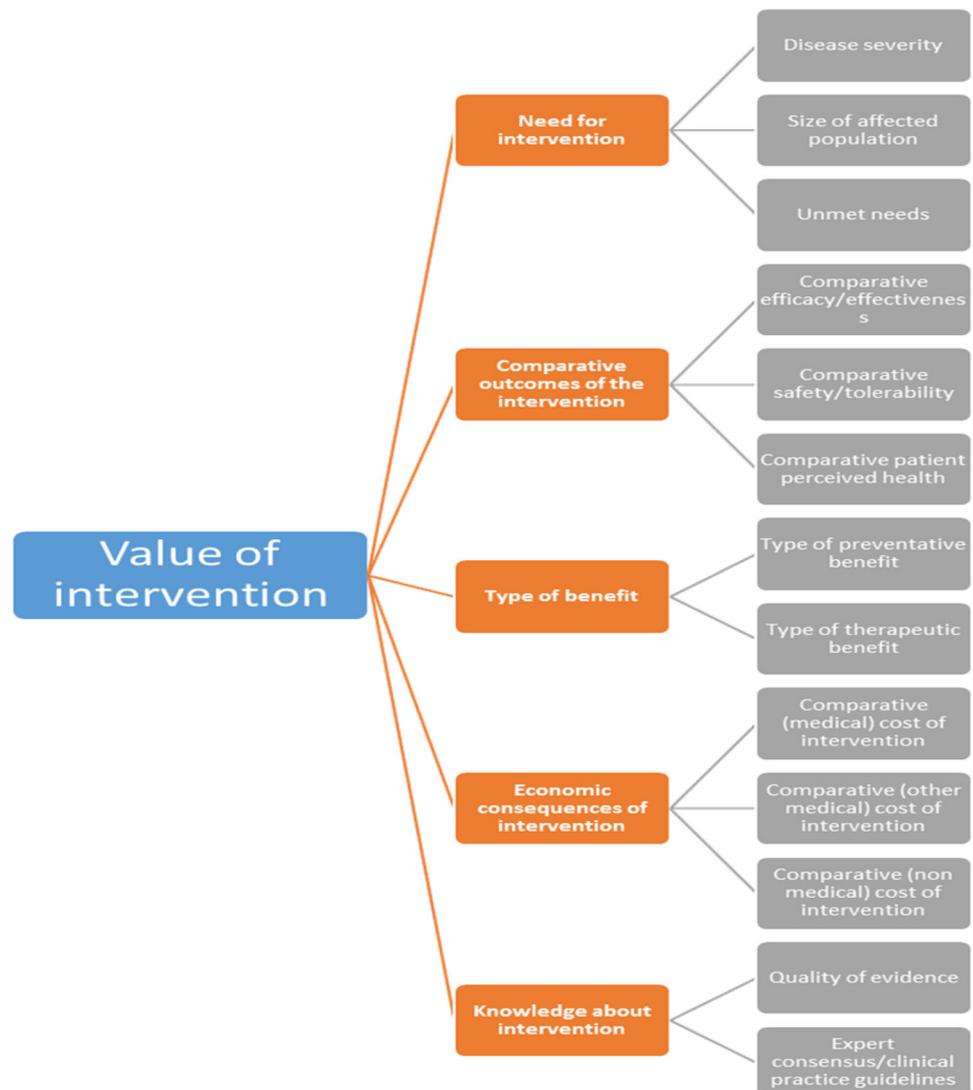
−5 to +5. Absolute criteria were measured on a scale from 0 to 5.

Evidence on obinutuzumab for iNHL was based on literature reviews included in HTA submissions and results of obinutuzumab clinical trials. Evidence was reviewed and synthesised using the EVIDEM framework. The evidence matrix provided to participants is presented in the electronic supplementary material.

Most of the evidence for the 'comparative efficacy/effectiveness' criterion referred to one of the clinical endpoints measured in the clinical trial, i.e. progression-free survival [16]. We recognise that this is a surrogate endpoint with different levels of associations with overall survival (OS), which is the primary measure of efficacy [18]. Nevertheless, there is an increasing acceptance of surrogate endpoints by regulatory agency and other healthcare decision makers, given the additional time and resources required for collecting evidence to estimate OS.

Responses were analysed in Microsoft Excel (Microsoft Corporation, Redmond, WA, USA), and key results, including average and standard deviation (SD) of weights and score values from the survey first round, were presented at structured meetings. Structured meetings, one with each stakeholder group, were run to allow participants to seek further explanation on the MCDA framework, to discuss and review weights and scores obtained, identify areas of agreement and disagreement among participants, and, where possible, reach a consensus on the weight and score values that could best represent the group's perspective.

To minimise the participants' cognitive burden and have a manageable number of criteria, the meeting discussions concentrated on the EVIDEM core criteria (presented in Fig. 1) and did not include the contextual qualitative criteria, which were omitted from the final results. The contextual criteria do not require weights and scores, therefore their exclusion do not affect the overall value score.

Participants were invited to complete the survey in light of the discussion at the meeting (survey second round). Following this, the average of the normalised weights and scores from the three groups were combined with linear aggregation to calculate the intervention value score. The literature suggests a variety of approaches to aggregating the preferences and views expressed by individuals to inform group decision making. These include agreeing the weight and score values as part of the committee discussion; aggregating by using the average of weights of scores obtained from responders; and retaining and comparing respondents' values [19]. We implemented the third approach as we observed the differences between groups' and individuals' values. We also used the second approach as an example of incorporating stakeholders' preferences in value assessment.

Participants were drawn from existing manufacturer networks. The payer group involved hospital, regional and national decision makers; patient representatives were members of Italian patient groups related to lymphomas; and clinicians were lymphoma specialists. Members of each group were distributed across Italian regions, covering the north, centre, and south areas.

3 Results

A total of 19 people were invited and completed the first round of the survey, including nine patients, five clinicians and five payers. The structured meetings were found to be useful by participants to discuss the MCDA framework and study approach with their peers. Only the clinician group reviewed their answers following the structured meetings.

3.1 Stakeholders' Weights

Weights represent trade-offs between criteria [6] and thus reveal which aspects of value matter most to each group. To obtain 'generic' weights to apply to any intervention, respondents were required to express their preferences between the EVIDEM criteria, based solely on their definition, and scale measurement, not the description of the intervention.

Figure 2 compares the sets of normalised weights from each group. According to patients, the two most important criteria were the 'type of therapeutic benefit' and 'disease severity', both with weights of 11% (SD 0.07 and 0.10, respectively). These are both absolute criteria (not relative to comparative interventions). The three least important criteria were the three economic indicators: 'comparative non-medical costs', 'comparative other medical costs', and 'comparative cost of intervention', with weights of 3% (SD 0.03), 4% (SD 0.04) and 5% (SD 0.04), respectively.

According to clinicians, the two most important criteria were 'disease severity' and 'comparative efficacy/effectiveness', with weights of 15% (SD 0.10) and 12% (SD 0.05), respectively. The three least important criteria were two economic indicators ('comparative non-medical costs' and 'comparative other medical costs'), and 'type of preventative benefit', with weights of 3% (SD 0.02), 4% (SD 0.02) and 4% (SD 0.03), respectively.

Finally, payers indicated that the three most important criteria were 'unmet needs', 'comparative cost of intervention', and 'quality of evidence', with weights of 11% (SD 0.07), 11% (SD 0.05) and 10% (SD 0.02), respectively. The three least important criteria were 'comparative non-medical costs', 'size of affected population' and 'comparative other medical costs', with scores of 4% (SD 0.01), 5% (SD 0.02) and 5% (SD 0.02), respectively. The

Fig. 2 Relative weights of individual criteria by stakeholder group
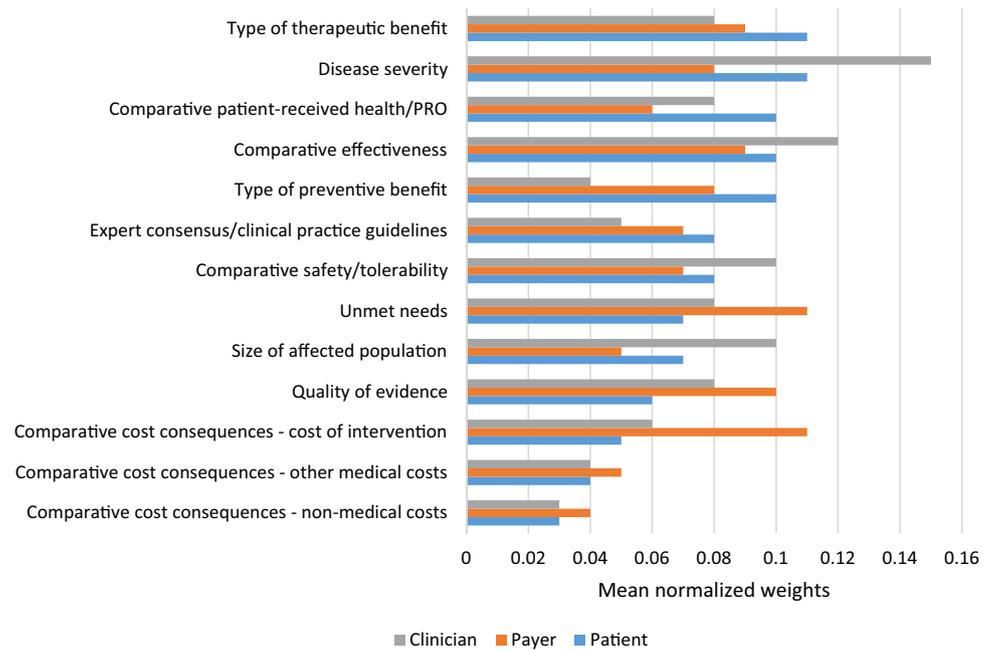
small SDs reflect a high level of agreement within this stakeholder group about the importance of each criteria, despite a small sample size.

Given the limited sample size, we did not perform any statistical comparison across stakeholder group weights; however, it is worth highlighting key differences and commonalities. Compared with patients and clinicians, payers distributed the weights more equally among the domains. Payers' weights range between 17 and 24%, compared with those given by clinicians and patients, which range between 12 and 33%. These two groups give less weight to the domains 'economic consequences of intervention' and 'knowledge about the intervention'.

Patients' and clinicians' views were aligned as they expressed preference for interventions targeting severe conditions. The highest weights in both group were observed for this criterion. Patients also believed that priority should be given to interventions that have a significant therapeutic effect (for example, they offer a cure or significantly delay progression of the disease), while clinicians indicated that one of the most important criteria is the improvement in clinical outcomes compared with SoC. Both groups ranked all the economic criteria among the five least important criteria.

Payers allocated higher weights to the economic-related criteria, with the direct (incremental) cost of the intervention being one of the most important. Their preferences were for treatments targeting populations in which there is little or no effective treatment, which are less expensive than the comparator, and which are underpinned by high-quality evidence. In contrast, the quality of evidence criterion was deemed a low priority by patients.

3.2 Stakeholders' Scores

Based on different outcome measures and types of evidence on obinutuzumab for the treatment of rituximab-refractory iNHL, participants provided scores. Unlike weights, scores are specific to the intervention under consideration. Figure 3 presents the scores allocated by the three stakeholder groups.

Patients assigned the highest scores (representing the areas where obinutuzumab performs best) to 'unmet needs', with an average score (AS) of 3.7 (SD 0.65); 'disease severity', with an AS of 3.6 (SD 0.46); and 'type of therapeutic effect', with an AS of 3.1 (SD 1.05). The criteria part of the 'comparative outcomes of intervention' domain were, on average, all positive, indicating that obinutuzumab is expected to generate incremental health gains compared with SoC. However, in the criterion 'patient-perceived health/patient-reported outcomes', there was a large variation in the assigned scores, which ranged from −2 to +5. All three criteria related to the economic impact of implementing the intervention were scored negative, on average.

Areas in which obinutuzumab was deemed to perform better than the comparator by clinicians were 'size of affected population', with an AS of 4 (SD 0.71); and 'disease severity', 'type of therapeutic effect', and 'quality of evidence', with an AS of 3.6 (SD 0.89, 0.89, 0.55, respectively). Two criteria part of the 'comparative

Fig. 3 Mean scores for obinutuzumab for each criteria by stakeholder group

Scores for obinutuzumab for the three stakeholder groups
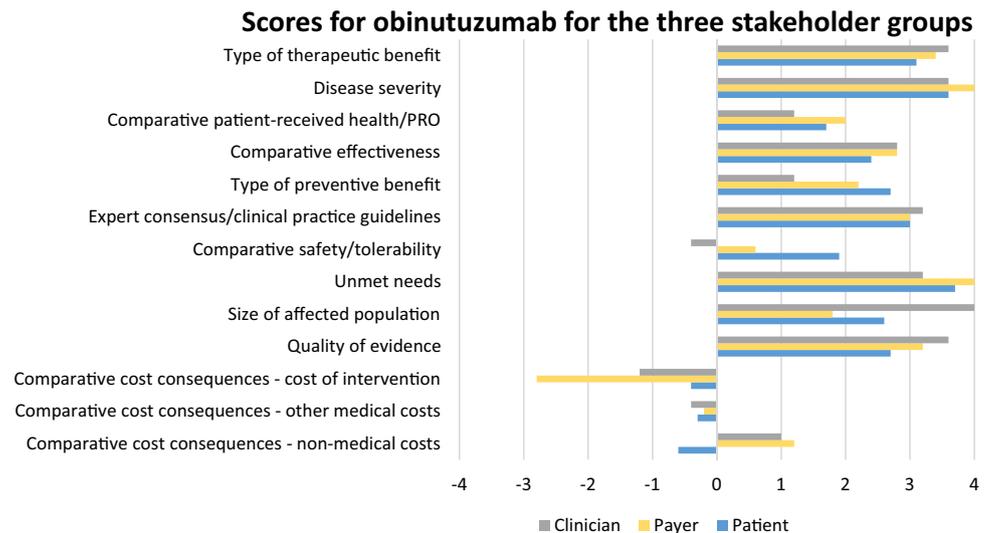
outcomes of intervention' domain were, on average, positive. However, in terms of safety and tolerability, obinutuzumab was deemed slightly worse than its comparator [with a score of −0.4 (SD 0.89)]. Clinicians also assigned negative scores to obinutuzumab against 'comparative cost of intervention' and 'comparative other medical costs', which obtained a score of −1.2 (SD 1.10) and −0.4 (SD 0.89), respectively.

Payers gave the highest scores to 'disease severity' and 'unmet needs', with an AS of 4.0 (SD 0.71, 0.71, respectively); and 'type of therapeutic benefit', with an AS of 3.4 (SD 0.55). Two criteria received negative scores, on average: 'comparative cost of intervention' and 'comparative other medical costs'.

We observed some consistency across stakeholder groups in relation to obinutuzumab scores.

The criteria 'disease severity' and 'type of therapeutic effect' were consistently assigned the highest scores by the three groups. This means that all groups believed that iNHL is a severe condition, given the patients' life expectancy after diagnosis and possible persistence of symptoms, and that obinutuzumab could bring clinical benefits at the patient level, including moderately delaying progression and helping to control disease symptoms. Both payers and patients thought that another area where obinutuzumab could bring value is 'unmet needs'. Available interventions for iNHL have limitations (e.g. a proportion of the population does not respond to SoC), which need to be addressed. Clinicians thought that data presented was relevant to decision makers and valid with respect to scientific standard. Finally, in the economic-related criteria 'comparative cost of intervention' and 'comparative other medical costs', obinutuzumab obtained a negative score (between −2.8 and 0.4) when compared with its comparator bendamustine,

whose patent has recently expired, according to all stakeholder groups. This is because the cost of obinutuzumab in combination with bendamustine and related medical costs was higher than those of bendamustine alone.

3.3 Overall Value Score of Obinutuzumab

To develop a combined perspective on obinutuzumab value, all survey responses were included and weighted equally. As shown in Fig. 4, the value score of obinutuzumab was 0.45.

A number of MCDA best-practice articles challenge the inclusion of (incremental) costs as a separate criterion. If the overall score is a composite measure of benefit, costs are not an attribute of benefit [20]. In addition, this would not allow for an appropriate consideration of the opportunity costs of the coverage decision [21]. Instead, costs can be considered separately to explicitly trade-off (incremental) benefits generated by a new treatment against its (incremental) costs (for a discussion about this issue see Garau and Devlin [22]). When decision makers face a fixed budget constraint, an aggregate measure of benefit (similar to the score presented in Fig. 5) can be compared with an estimate of costs. This approach is presented by Golan and Hansen, who developed an MCDA framework piloted by the Israeli Advisory Committee to select new interventions to fund [23].

Figure 4 also shows that the key drivers of the obinutuzumab score value are 'disease severity' (which accounts for approximately 18% of the total value), 'type of therapeutic benefit' and 'unmet needs' (which accounts for approximately 13% of the total value).

We conducted a sensitivity analysis where we set the weight for the comparative costs criteria to zero, scaled up
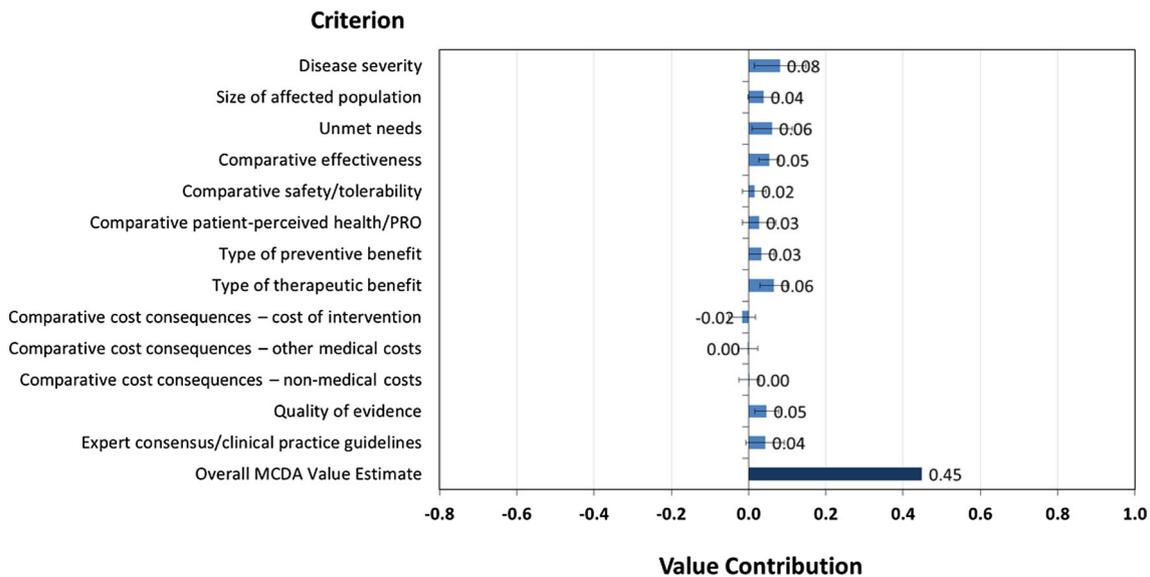
Criterion

Fig. 4 Overall value score for obinutuzumab (all stakeholder groups combined) and contribution of each criterion to the total value
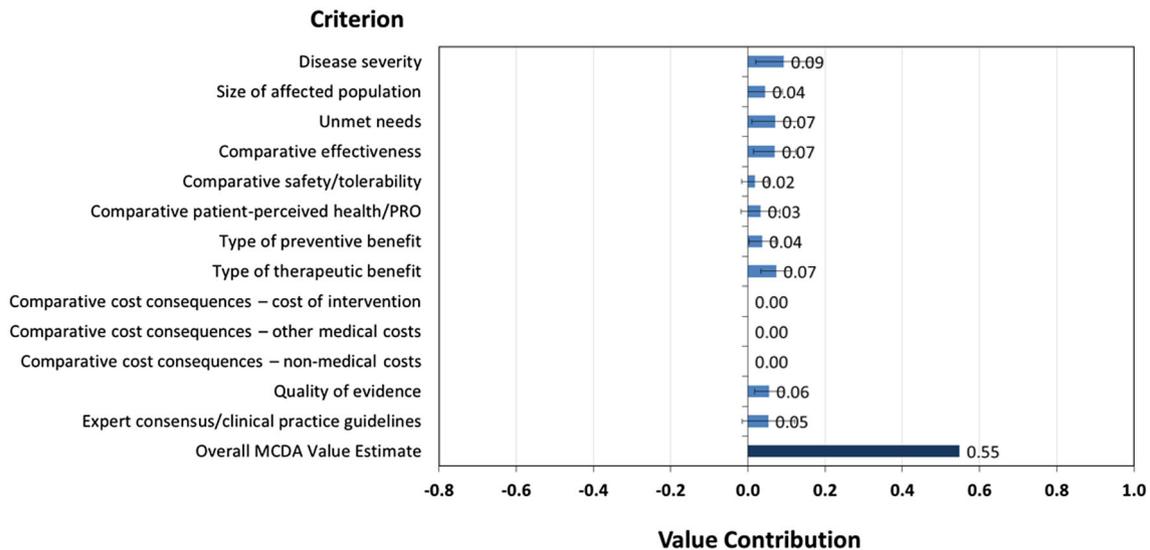
Criterion

Fig. 5 "Benefit" score for obinutuzumab (sensitivity analysis) and contribution of each criterion to total benefit. *Note* The benefit score is the overall value score excluding the criteria relating to cost

the weights for the remaining criteria clusters, and recalculated the overall score from a combined perspective. The result is shown in Fig. 5.

Removing the cost criteria increases the obinutuzumab value score from 0.45 to 0.55. If this version of the framework was used, decision makers would need to assess and consider the net economic impact alongside this benefit score. This approach might also be helpful in those systems where the price of the intervention is defined following its

benefit assessment, similar to that followed by the AIFA in Italy.

4 Discussion

In many systems, including the Italian system, the perspectives of stakeholders such as patients are not elicited or incorporated at any stage of the assessment and decision-

making process. The use of an MCDA framework such as EVIDEM could enable the collection of stakeholders' preferences (via weights allocation) and help to ensure that they are taken into account more systematically in decision making (via determination of the value score and its consideration in decision making).

The value score can help identify the key criteria impacting the intervention's value, and lead to an in-depth discussion within the decision-making committee around the evidence presented on those criteria and the level of consensus that was obtained across participants when assigning weights and scores. It can also inform sensitivity analyses evaluating the robustness of the decision outcome.

The value score has limited use in absolute terms if there is no specification on how it should inform coverage decisions. If used to compare and rank competing technologies, or across successive decisions about different technologies, score values might be useful. An example of this approach is the MCDA framework for health technology prioritisation developed for the Israeli Advisory Committee [23]. This approach used the benefit score and the net total cost to draw efficiency frontiers and, based on budget constraints, selected the technologies to be funded.

On the other hand, for repeated reimbursement or HTA decisions affecting a fixed budget, there is a need to define the 'hurdle for adoption' [24]; in other words, the incremental cost per value score to compare against the cost per value score of individual interventions to understand whether they are good value for money. However, given the methodological issues in defining and estimating the opportunity cost of HTA decisions [25–28] and the role of regional (as opposed to national) jurisdictions in the management of the health budget, there might be a need to develop new approaches to ensure efficient decision making [22].

In the context of the Italian National Health Service (NHS), EVIDEM is implemented in the Lombardia region to make listing and de-listing decisions on medical devices [7]. More than 20 interventions have been appraised and have obtained value scores, including economic criteria, between 0.22 and 0.72 [9]. Therefore, obinutuzumab fits in the middle of this range. However, the version of EVIDEM used in Lombardia is slightly different to that used in our study; it is applied in a regional context rather than a national context, and information on which scores, on average, allowed interventions to be approved for reimbursement does not appear to be available.

We should also highlight that, consistent with the purposes of MCDA, value scores are intended to inform and support decision making and not to be used as a prescriptive rule in place of deliberations. A deliberative component is seen as necessary in all decision-making processes [29].

Currently, the AIFA does consider some of the criteria included in the EVIDEM framework; however, this is not done systematically and it remains unclear how evidence on those criteria is developed and to what extent it influences decision making. An MCDA process such as the one applied in this study can make both aspects more explicit and lead to more consistent consideration of multiple criteria in decision making. To implement MCDA in practice at national level, broader and larger groups of stakeholders embracing different disease areas would need to be consulted. Alternatively, the decision-making committee (either the Technical Scientific Committee or the Prices and Reimbursement Committee within the AIFA) can act as the agent and represent different stakeholders (such as the local NHS payers, general public and patients—the principals) in determining the relative importance of criteria.

5 Limitations

This study was exploratory as it applied MCDA in the context of medicine reimbursement decision making in Italy in a convenience sample. In future uses, improvements can be made to increase the validity of results and their applicability in formal decision-making processes.

Our convenience sample ensured a high rate of response but its size and clinical areas covered could be expanded, e.g. involving patients and clinicians of other non-oncology conditions. Survey instructions and synthesis of evidence on obinutuzumab were provided in English, while during the meetings, the participants and the moderator spoke in Italian, which helped in the interpretation of the scientific evidence and the instructions. If the exercise is conducted on a larger scale, it would benefit from the translation of all the material into the relevant language to increase understanding and rate of responses.

Patient representatives also raised the need to simplify the language used to explain the framework and make it more accessible to lay persons. This shows that the interactive component is not only needed at the end of the process, to consolidate survey responses, but also at the start, to ensure full understanding around the criteria and their definitions. Challenges in communicating the elicitation exercise to patients was also highlighted by Marsh et al. [30]. Therefore, validation with stakeholders should be part of each step of an MCDA exercise [21].

Given the importance of the interaction part of an MCDA process, future exercise should include qualitative analysis on methods used to structure group discussions and the type of interaction. Exploring how structured group discussions change (or not) when participants meet in person compared with when they interact in a virtual space

(e.g. webinar) might help to establish efficient and effective ways to support group dynamics.

In addition, as discussed in Mühlbacher and Kaczynski [31], there are underlying difficulties in understanding the data provided and interpreting the score scales. For example, three patients reported that obinutuzumab was cheaper than the comparator (in the cost of intervention criterion), which is not in line with what we would have expected from the data presented to them. This may reveal that they did not fully understand the task. In contrast, all patients reported that obinutuzumab was more clinically effective than the comparator, which is in line with what we would have expected.

The interval scale incorporated into EVIDEM has been tested and validated in several applications [31] and refined accordingly over time. For example, scores were initially measured on a scale from 0 to 3 [13] and subsequently on scale from 0 to 5, probably to increase its discriminatory power. Others have suggested the use of a scale from 0 to 100 [32]. Further validation exercises might be required to assess consistency and interval properties of the scoring scale.

We observe that more guidance needs to be provided in order to score some of the criteria, particularly the economic-related criteria. What constitutes 'substantial additional expenditures' (corresponding to a score of −5) for one respondent might be different to that of another respondent. We included information about the national pharmaceutical expenditure to provide some context; however, clear ranges or cut-off values should be included for each score to ensure consistency in responses.

The validation process should also ensure that criteria are 'preference independent', meaning that it is possible to judge how well one criterion is achieved without knowing how well any of the other criteria are achieved. One participant pointed out that the criterion 'unmet needs' should be considered in conjunction with 'disease severity' as the lack of alternative interventions is meaningless to decision makers if it is not referred to a serious condition. This issue can be addressed with alternative aggregation approaches, such as multiplicative instead of additive methods [21].

To limit the cognitive burden of participants, we focused the meeting discussions on a restricted number of criteria (the 13 EVIDEM 'core criteria'). There is no rule on the optimal number of criteria to include in an MCDA framework; however, it is important to consider the trade-off between breath/inclusiveness of a framework and resources needed to develop relevant evidence and analyse it for decision making [21].

On the method to obtain weights, we selected one of the methods recommended by the EVIDEM collaboration: the point allocation approach. However, because of the top-down approach used to allocate points, firstly among the domains and secondly among criteria within each domain, we noted some distortions for the values obtained by criteria in domains with two criteria, which tended to have higher weights, versus criteria in domains with three criteria. One way to avoid this is to assign points directly to the criteria rather than splitting the task into two stages. For larger-scale applications of MCDA, alternative instruments to elicit preferences that have strong theoretical foundations, and that have been used in other types of healthcare assessments, can be considered, such as discrete choice experiments and PAPRIKA [33].

6 Conclusions

To our knowledge, this is the first time that an MCDA approach has been used to inform reimbursement decision making of medicines at the national level in Italy.

This study showed that MCDA (in particular, EVIDEM) can be used to elicit the views of different stakeholder groups. We found that the views of patients and clinicians were broadly aligned as they expressed preference for interventions targeting severe conditions and they ranked economic criteria as the least important interventions. Payers allocated higher weight (compared with patients and clinicians) to the economic criteria and to the quality of evidence. The key criteria driving the value of obinutuzumab accordingly to all the stakeholder groups were disease severity, type of therapeutic benefit and unmet needs.

Decision makers in Italy already consider some of the EVIDEM criteria, such as disease severity, but with no systematic approach. The perspectives of stakeholders (such as patients) are not elicited or incorporated at any stage of the assessment or decision-making process. Our MCDA study provides useful evidence to decision makers, such as the AIFA, on what health interventions attribute different stakeholders' value the most, and has tested methods to ensure that this is captured consistently across different interventions.

receives funding from a variety of sources, including the Association of the British Pharmaceutical Industry. Nicola Amedeo Mazzanti and Antonio Profico are employees of Roche S.P.A.

References

1. HTAi Patient and Citizen Involvement. Good Practice Examples of Patient and Public Involvement in Health Technology Assessment. 2015. http://www.htai.org/fileadmin/HTAi_Files/ISG/PatientInvolvement/EffectiveInvolvement/Good_Practice_Examples_Feb_2015.pdf. Accessed 14 Aug 2017.
2. SMC. PACE—patient and clinician engagement. Factsheet. 2015. https://www.scottishmedicines.org.uk/files/PACE/PACE_factsheet_0417.pdf. Accessed 14 Aug 2017.
3. Shah KK, Mestre-Ferrandiz J, Towse A, Smyth EN. A review of health technology appraisals: case studies in oncology. Int J Technol Assess Health Care. 2013;29(1):101–9.
4. Devlin N, Sussex J. Incorporating multiple criteria in HTA: methods and processes. OHE Monograph. London: Office of Health Economics; 2011.
5. Danner M, Hummel JM, Volz F, van Manen JG, Wiegard B, Dintsios CM, et al. Integrating patients' views into health technology assessment: analytic hierarchy process (AHP) as a method to elicit patient preferences. Int J Technol Assess Health Care. 2011;27(4):369–75.
6. Thokala P, Devlin N, Marsh K, Baltussen R, Boysen M, Kalo Z, et al. Multiple criteria decision analysis for health care decision making-an introduction: report 1 of the ISPOR MCDA emerging good practices task force. Value Health. 2016;19(1):1–13.
7. Radaelli G, Lettieri E, Masella C, Merlino L, Strada A, Tringali M. Implementation of EUnetHTA core Model(R) in Lombardia: the VTS framework. Int J Technol Assess Health Care. 2014;30(1):105–12.
8. Castro H, Tringali M, Cleemput I, Devriese S, Leoni O, Lettieri E. Advancing MCDA and HTA into coverage decision-making. In: Marsh K, Goetghebeur M, Thokala P, Baltussen R, editors. Multi-criteria decision analysis to support healthcare decisions. Cham: Springer International Publishing; 2017. p. 119–46.
9. Wahlster P, Goetghebeur M, Schaller S, Kriza C, Kolominsky-Rabas P. Exploring the perspectives and preferences for HTA across German healthcare stakeholders using a multi-criteria assessment of a pulmonary heart sensor as a case study. Health Res Policy Syst. 2015;13:24.
10. Bentley TG, Cohen JT, Elkin EB, Huynh J, Mukherjea A, Neville TH, et al. Validity and reliability of value assessment frameworks for new cancer drugs. Value Health. 2017;20(2):200–5.
11. Wagner M, Khoury H, Willet J, Rindress D, Goetghebeur M. Can the EVIDEM framework tackle issues raised by evaluating treatments for rare diseases: analysis of issues and policies, and context-specific adaptation. Pharmacoeconomics. 2016;34(3):285–301.
12. Daniels N, Sabin JE. Accountability for reasonableness: an update. BMJ. 2008;337:a1850.
13. Tony M, Wagner M, Khoury H, Rindress D, Papastavros T, Oh P, et al. Bridging health technology assessment (HTA) with multi-criteria 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.
14. Gilabert-Perramon A, Catalan A, Prat A, Fontanet M, Puig-Peiró R, Merino-Montero S, et al. Development of a multi-criteria decision analysis (MCDA) framework for health care decision-making in Catalonia (Spain): pilot study focused in orphan drugs. Vienna: ISPOR 19th Annual European Congress; 2016.
15. van Til J, Groothuis-Oudshoorn C, Lieferink M, Dolan J, Goetghebeur M. Does technique matter: a pilot study exploring weighting techniques for a multi-criteria decision support framework. Cost Eff Resour Alloc. 2014;12:22.
16. Sehn LH, Chua N, Mayer J, Dueck G, Trneny M, Bouabdallah K, et al. Obinutuzumab plus bendamustine versus bendamustine monotherapy in patients with rituximab-refractory indolent non-Hodgkin lymphoma (GADOLIN): a randomised, controlled, open-label, multicentre, phase 3 trial. Lancet Oncol. 2016;17(8):1081–93.
17. European Medicines Agency. Assessment report Gazyvaro Procedure No. EMEA/H/C/002799/II/0007. 2016. http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Assessment_Report_-_Variation/human/002799/WC500211313.pdf. Accessed 14 Aug 2017.
18. Lee L, Wang L, Crump M. Identification of potential surrogate end points in randomized clinical trials of aggressive and indolent non-Hodgkin's lymphoma: correlation of complete response, time-to-event and overall survival end points. Ann Oncol. 2011;22(6):1392–403.
19. Belton V, Pictet J. A framework for group decision using a MCDA model: sharing, aggregating or comparing individual information? J Decis Syst. 1997;6(3):283–303.
20. Claxton K. Three questions to ask when examining MCDA. Value Outcomes Spotlight. 2015;1:10–8.
21. Marsh K, IJmerzan M, Thokala P, Baltussen R, Boysen M, Kalo Z, et al. Multiple criteria decision analysis for health care decision making—emerging good practices: report 2 of the ISPOR MCDA emerging good practices task force. Value Health. 2016;19(2):125–37.
22. Garau M, Devlin NJ. Using MCDA as a decision aid in health technology appraisal for coverage decisions: opportunities, challenges and unresolved questions. In: Marsh K, Goetghebeur M, Thokala P, Baltussen R, editors. Multi-criteria decision analysis to support healthcare decisions. Cham: Springer International Publishing; 2017. p. 277–98.
23. Golan O, Hansen P. Which health technologies should be funded? A prioritization framework based explicitly on value for money. Isr J Health Policy Res. 2012;1(1):44.
24. Sussex J, Towse A, Devlin N. Operationalizing value-based pricing of medicines: a taxonomy of approaches. Pharmacoeconomics. 2013;31(1):1–10.
25. Culyer A, McCabe C, Briggs A, Claxton K, Buxton M, Akehurst R, et al. Searching for a threshold, not setting one: the role of the National Institute for Health and Clinical Excellence. J Health Serv Res Policy. 2007;12(1):56–8.
26. Karlsberg Schaffer S, Sussex J, Devlin N, Walker A. Local health care expenditure plans and their opportunity costs. Health Policy. 2015;119(9):1237–44.
27. Appleby J, Devlin N, Parkin D, Buxton M, Chalkidou K. Searching for cost effectiveness thresholds in the NHS. Health Policy. 2009;91(3):239–45.

28. Claxton K, Martin S, Soares M, Rice N, Spackman E, Hinde S, et al. Methods for the estimation of the National Institute for Health and Care Excellence cost-effectiveness threshold. Health Technol Assess. 2015;19(14):1–503 (v–vi).

29. Baltussen R. Question is not whether but how to use MCDA. Value Outcomes Spotlight. 2015;1(1):14–6.

30. Marsh K, Caro JJ, Hamed A, Zaiser E. Amplifying each patient's voice: a systematic review of multi-criteria decision analyses involving patients. Appl Health Econ Health Policy. 2017;15(2):155–62.

31. Mühlbacher AC, Kaczynski A. Making good decisions in healthcare with multi-criteria decision analysis: the use, current research and future development of MCDA. Appl Health Econ Health Policy. 2016;14(1):29–40.

32. Phillips LD. Best Practice for MCDA in Healthcare. In: Marsh K, Goetghebeur M, Thokala P, Baltussen R, editors. Multi-criteria decision analysis to support healthcare decisions. Cham: Springer International Publishing; 2017. p. 311–29.

33. 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.