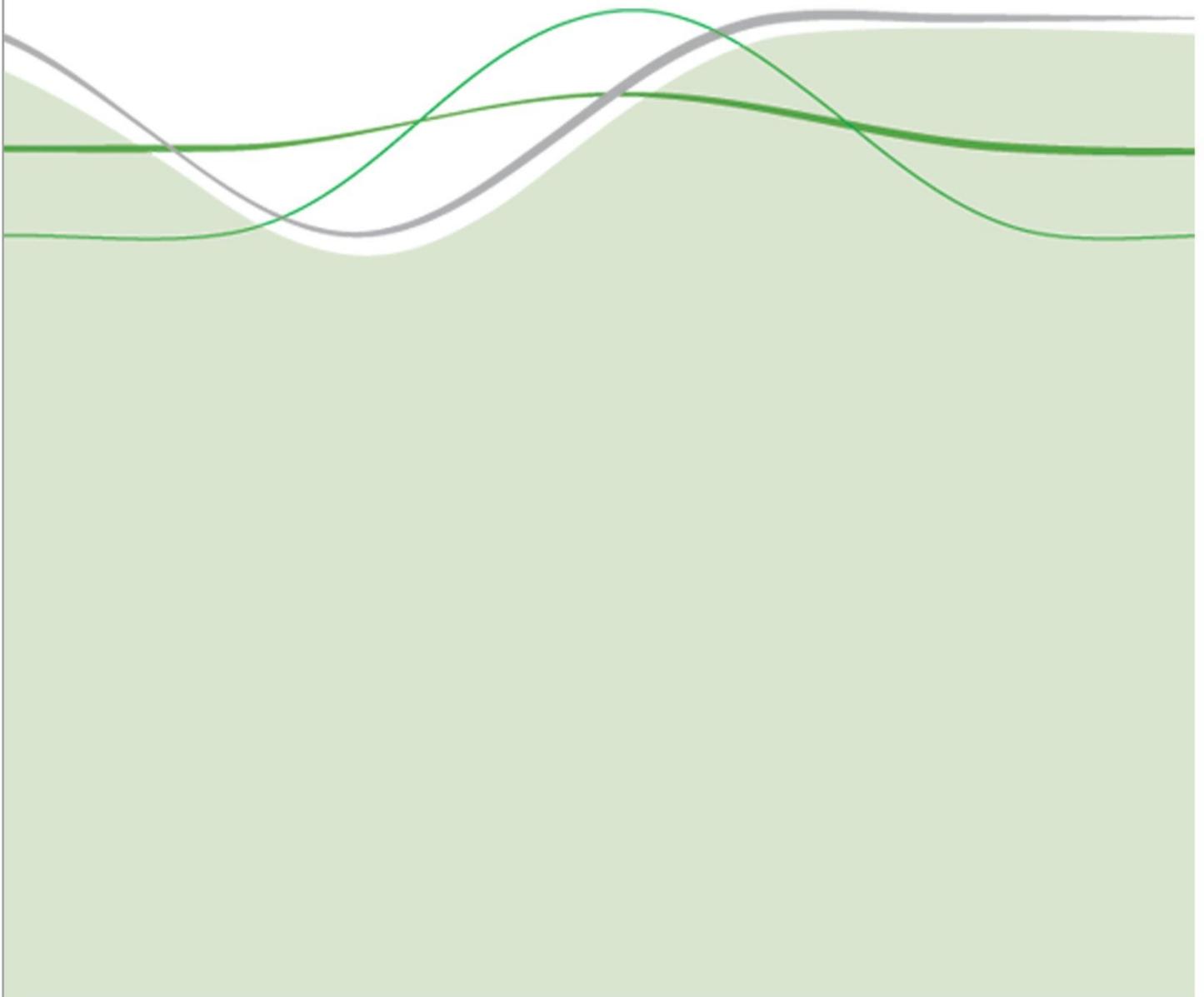


Applying a Multi-criteria Decision Analysis
(MCDA) Approach to Elicit Stakeholders'
Preferences in Italy. The Case of Obinutuzumab
for Rituximab-Refractory Indolent Non-Hodgkin
Lymphoma (iNHL)

December 2016

Martina Garau, Grace Marsden, Nancy Devlin, Nicola Amedeo
Mazzanti, and Antonio Profico



Applying a Multi-criteria Decision Analysis (MCDA)
Approach to Elicit stakeholders' Preferences in Italy.
The Case of Obinutuzumab for Rituximab-Refractory
Indolent Non-Hodgkin Lymphoma (iNHL)

Martina Garau^a, Grace Marsden^a, Nancy Devlin^a, Nicola
Amedeo Mazzanti^b and Antonio Profico^b

^aOffice of Health Economics ^b Roche S.p.A.

December 2016

For further information please contact:

Nancy Devlin

ndevlin@ohe.org

The Office of Health Economics (a company
limited by guarantee of registered number
09848965)

Southside, 105 Victoria Street

London SW1E 6QT

United Kingdom

Tel: +44 207 747 8858

©Office of Health Economics

About OHE Research Papers

OHE Research Papers are intended to provide information on and encourage discussion about a topic in advance of formal publication. Any views expressed are those of the authors and do not necessarily reflect the views or approval of OHE, its Editorial Board, or its sponsors. Once a version of the Research Paper's content is published in a peer-reviewed journal, this supersedes the Research Paper and readers are invited to cite the published version in preference to the original version.

Acknowledgements

The work on this paper was commissioned from OHE Consulting by Roche S.P.A. as a consulting project, with the intention to publish results from the outset.

Table of Contents

Abstract	1
1. Background	2
2. Methods	3
3. Results	6
3.1. Stakeholders' preferences between criteria - weights.....	6
3.2. Stakeholders' scores of obinutuzumab for iNHL.....	8
3.3. Overall value score of obinutuzumab	10
4. Discussion.....	12
5. Limitations	13
6. Conclusions	14
References	15

ABSTRACT

Background: Health technology assessment (HTA) bodies and decision makers need to understand what matters most to patients and other stakeholders, in order to allocate resources across the health system. Similarly, stakeholders in the process need to know to what extent and how their input affects final HTA decisions. The purpose of this study was to use multi-criteria decision analysis (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: An MCDA framework (EVIDEM V3.0) was used to elicit stakeholders' preferences about the relative importance of decision criteria (weights) and to assess the degree of achievement of obinutuzumab for rituximab-refractory iNHL in each criteria (scores) via an online survey and structured meetings with each stakeholder group. The normalised weights and scores from each of the groups were combined with a linear function to calculate the intervention value score.

Results: The patients and clinicians both expressed a preference for interventions targeting severe conditions and ranked the economic criteria among the five least important criteria. Payers expressed preference for treatments targeting populations in which there is currently little or no effective treatment, which are less expensive than the comparator, and which are underpinned by high quality evidence.

Obinutuzumab received high scores for the criteria "disease severity" and "type of therapeutic benefit" by all three groups. Against the economic-related criteria ("comparative cost consequences – cost of intervention" and "comparative cost consequences – other medical costs") obinutuzumab obtained a negative score compared to its comparator bendamustine, whose patent has recently expired, according to all stakeholder groups.

Conclusions: This is the first time that an MCDA approach has been used to inform reimbursement decisions in Italy at the national level. This study shows that MCDA (and in particular EVIDEM) can be used to elicit the views of different stakeholder groups.

Decision makers in Italy already consider some of the EVIDEM criteria, such as disease severity, but with no systematic approach. Perspectives of stakeholders (such as patients) are not elicited or incorporated at any stage of the assessment or decision making process. We conclude that MCDA studies provide useful evidence to decision makers on what constitutes value of health interventions according to different stakeholder perspectives, and can ensure that this is captured consistently across different decisions.

Keywords

Multi-criteria decision analysis, stakeholder perspective, stakeholder involvement, decision-making, Italian medicine reimbursement.

1. BACKGROUND

Various stakeholders such as patients can be involved in medicine reimbursement decisions and Health Technology Assessment (HTA), with their involvement taking different forms, from contributions to evidence submissions, to participation in advisory or decision making committees.

Key stakeholders can help identify which factors should be considered in decision making, which outcomes are the most important to them and which additional impacts the new treatment could have on their life, beyond those captured in standard health-related quality of life measures. For some HTA bodies, such as 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, patient involvement has taken the form of evidence submissions by and participation in committee meetings of patient representatives (HTAi Patient and Citizen Involvement, 2015).

In the case of orphan drugs, the preferences of people directly affected by rare and complex conditions can play an important role in treatment assessment. Clinicians who deal with these conditions on a daily basis can also provide valuable insights (Sussex et al., 2013) (Paulden et al., 2015). In this spirit, the Scottish Medicines Consortium (SMC) has introduced the option of forming and consulting an external panel of clinicians and patient representatives during the assessment of medicines for rare or end of life conditions (SMC, 2015).

HTA decision makers need to understand what matters most to these groups: how do stakeholders trade-off between different types of benefit potentially generated by a new treatment? Stakeholders, on the other hand, need to know to what extent and how their input affect final HTA decisions. In a study of health technology appraisals in five countries (Shah et al., 2013) found that patient preferences are rarely mentioned in publicly accessible HTA reports, and that processes for patient participation in HTA tend to be unsystematic.

Multi-criteria decision analysis (MCDA) provides a clear framework to assess the value of a new treatment compared to alternative treatment options (or standard of care), against multiple and competing criteria (Devlin and Sussex, 2011). It can support decision makers to be explicit and transparent about the trade-offs made between the selected criteria. It can also offer more 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 identify patient-relevant outcomes in depression and hepatitis C, and elicit patient preferences on the selected outcomes (Danner et al., 2011) (Thokala et al., 2016). In Italy, the Lombardia region uses an MCDA framework (EVIDEM) as a decision-making aid to select health technologies to reimburse (Radaelli et al., 2014).

The purpose of this study was to use an MCDA approach to obtain preferences and views on decision 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 with the aim of informing reimbursement decision making at the national level in Italy.

Our work built on the study by Wahlster et al. (Wahlster et al., 2015) who applied the EVIDEM framework to capture the views and preferences of stakeholders in Germany

using a heart pulmonary sensor as a case study. In addition to the online survey that was used to elicit preferences of stakeholders individually by Wahlster et al., we applied the EVIDEM framework to inform and support group discussion within each stakeholder group. This is a fundamental role that MCDA can play in HTA and other healthcare decision making processes, to structure committees' consideration of different and often conflicting perspectives (Garau & Devlin, forthcoming).

2. METHODS

There are number of approaches to elicit preferences of patients and other stakeholders with different level of complexity and theoretical bases, including swing weights, analytic hierarchy process (AHP) and discrete choice experiments (Thokala et al., 2016). For the purposes of this study, we selected a framework which has been developed and used specifically for application of MCDA to health care decision making – EVIDEM. To date EVIDEM is the only MCDA framework specifically designed for health care decision making, and its choice here enables results to be compared with a growing literature from its use in other health care decision making settings.

EVIDEM (V3.0¹) is an open-source framework resulting from a collaboration of experts and stakeholders, and is subject to continued testing and adaptations. It comprises a broad range of decision criteria included, allowing us to capture all elements of value relevant to patients, health care systems and society; their properties (including non-overlap between criteria) meeting MCDA good practice guidelines (Marsh et al., 2016); and their operational nature, enabling us to inform real-life decision making. Because of these characteristics, EVIDEM has been applied and tested in the context of HTA decision making in a number of jurisdictions, including Canada (Tony et al., 2011) and, more recently, Spain, where the HTA body in Catalonia has been exploring the framework for the appraisal of orphan drugs (Gilbert-Perramon A, 2016). The Italian region Lombardia uses EVIDEM to inform local funding decisions of health interventions, including medical devices and medicines, so we know it is feasible in the Italian context. Our aim was to extend the application of an MCDA framework to support national decision making of medicine reimbursement, which is currently under the responsibility of *Agenzia Italiana del Farmaco* (AIFA²).

The objective of the EVIDEM framework is to make priorities and perspectives explicit and to support systematic consideration of decision criteria when appraising a healthcare intervention. The framework includes a set of 13 clearly defined and measurable “core” criteria grouped into domains. Criteria can be weighted to reflect their relative importance from different perspectives, be used for scoring the intervention, and combined to derive a composite measure of the intervention's value. The core criteria includes criteria which are measured in absolute terms, not relative to other interventions, and criteria which are measured comparatively to existing interventions.

The EVIDEM framework also comprises 7 qualitative criteria to capture contextual factors, including: mandate and scope of healthcare system; population priorities and access of healthcare system/plan; common goal and specific interests; environmental

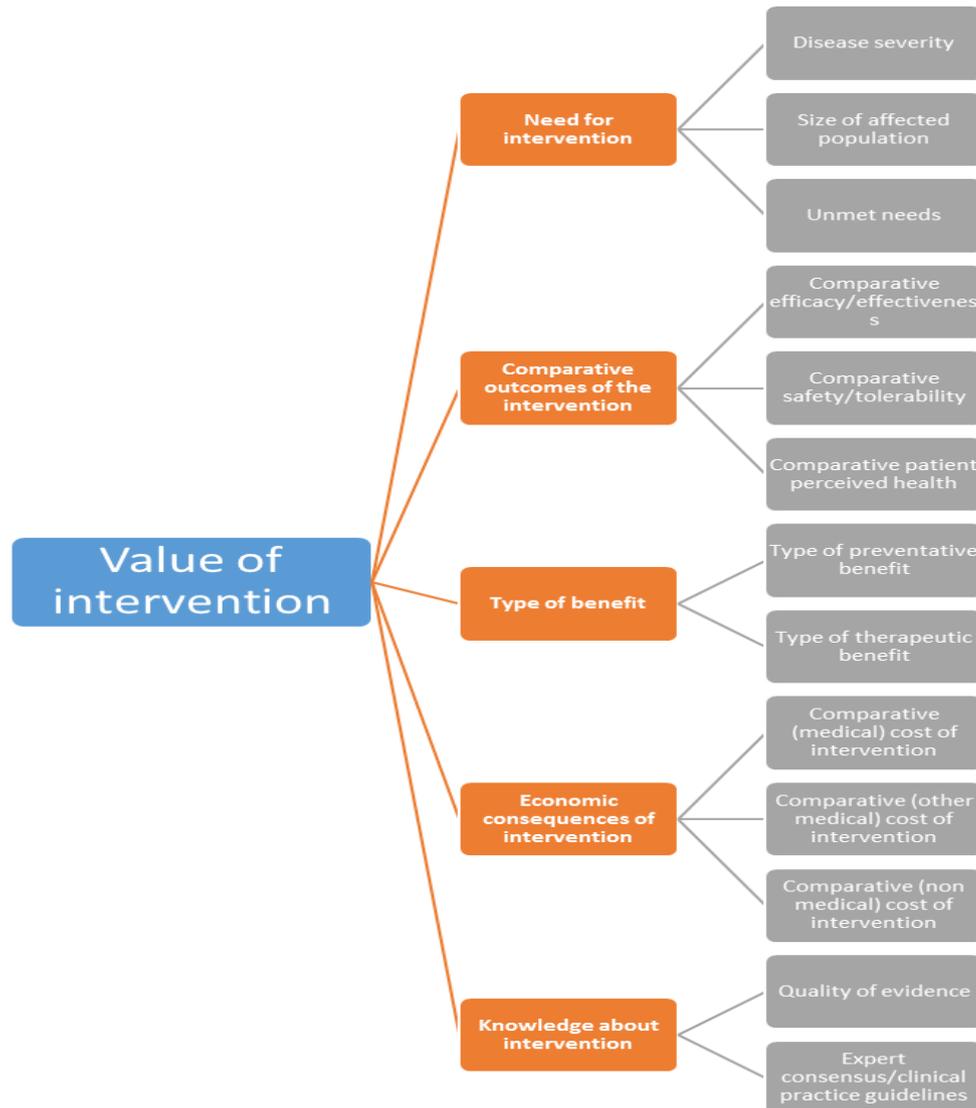
¹ <https://www.evidem.org/>

² <http://www.agenziafarmaco.gov.it/en>

impact; opportunity costs and affordability; system capacity and appropriate use of intervention; political, historical and cultural context.

Figure 1 presents the core EVIDEM framework including its five domains: need for intervention, comparative outcomes of intervention, type of benefit, economic consequences of intervention, and knowledge about intervention (indicated in orange). Each domain comprises a set of criteria (indicated in grey, 13 in total).

Figure 1: MCDA EVIDEM framework



The study entailed the following steps:

- a) First round of the survey in which an online survey was sent to three groups of stakeholders involved or affected by reimbursement decisions in Italy: clinicians, patients, and payers. The survey was used to elicit preferences around the relative importance of decision criteria of the EVIDEM framework (to establish the *weights* attached to the criteria), and the degree of achievement of obinutuzumab against the identified decision criteria as compared to the current standard of care (to determine the *scores* of the intervention).

- b) Structured meetings, one with each stakeholder group (one face-to-face and two via webinar), to allow survey participants to discuss and review weights and scores obtained through the online survey; identify areas of agreement and disagreement among participants; and, where possible, reach a consensus on the weights and scores values that could best represent the stakeholder group's perspective.
- c) Second round of the survey, to give respondents the opportunity to complete the survey in light of the discussion at the meeting;
- d) Calculation of obinutuzumab value score based on the collected stakeholders' weights and scores, and analysis of the results.

We developed an online survey, using SurveyMonkey®, to elicit the weights of the EVIDEM decision criteria and the scores of obinutuzumab for each criterion. We derived weights using the "point allocation" approach, where we asked participants to allocate 100 points first across criterion domains (indicated in orange in Figure 1) and second across criteria within each clusters (indicated in grey in Figure 1). The description of the EVIDEM criteria and instructions of the survey given to participants are available in the supplementary material 1 available on request. 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).

Out of the six methods suggested by the EVIDEM collaboration, the "point allocation" was selected because it combines simplicity with the ability of "forcing" people to priorities the criteria. We also noted that van Til et al. (van Til et al., 2014) showed that the choice of the weight elicitation method does not affect value estimates at the group level.

Respondents were asked to score the performance of obinutuzumab (called Product X in our survey) in combination with bendamustine followed by obinutuzumab maintenance, compared with bendamustine alone, in patients with rituximab-refractory iNHL. This is in line with obinutuzumab license indication and the clinical evidence (EMA, 2016) (Sehn et al., 2016). Bendamustine was chosen as the comparator as it is the only efficacious intervention in this indication (Sehn et al., 2016).

Incremental criteria, related to the health and non-health effects of the intervention, were measured on a scale from -5 to +5 to capture worse and better impact than the comparator. Absolute criteria, related to characteristics of the disease before the use of the new intervention such as "disease severity", were measure on a scale from 0 to 5.

Evidence on obinutuzumab for iNHL was based on literature reviews conducted to develop HTA submissions and results of obinutuzumab clinical trials. Sources of the literature searches were: Pubmed, Embase, Centers for Review and Dissemination, trial registries and Cochrane reviews. Relevant evidence was reviewed by one of the authors and synthesized using the EVIDEM framework. The evidence matrix provided to the survey participants to undertake the scoring part is presented in the supplementary material 2 and is available on request.

Most of the evidence for the "comparative efficacy/effectiveness" criterion referred to the clinical endpoint measured in the clinical trial, progression-free-survival (PFS) (Sehn et al., 2016). We recognise that this is a surrogate endpoint that could have different levels of associations with overall survival (OS), which is the primary measure of efficacy (Lee et al., 2011). 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 relevant evidence supporting OS estimation.

Responses were analysed in Excel and key results, including average, minimum and maximum, and standard deviation of weights and score values obtained from the first round of surveys were presented at the structured meetings.

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

Following the second round of the survey, the average of the normalised weights and scores from the three stakeholder 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. They include: agreeing the weights and scores values as part of the committee discussion; aggregating by, for example, using the average of weights of scores obtained from responders; retaining and comparing respondents' values (Belton and Pictet, 1997). We implemented the third approach, in the sense that we observed the differences between groups' and individuals' values. This was important given the diverse perspectives that the three groups considered brings in. However, we also used the second approach as an example of incorporating stakeholders' preferences in value assessment. In a more deliberative fashion and in line with the first approach suggested by Belton and Pictet (1997), we could have presented and compared each group weights and scores values, and reach a consensus on the values for the intervention value score, where possible, as a result of a group discussion.

Participants were drawn from the manufacturer existing networks. The Payers group involved hospital, regional and national decision makers. Patient representatives were members of Italian patient groups related to lymphomas. Clinicians were lymphoma specialists. Members of each group were distributed across different 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 9 patients, five clinicians and five payers. All participants in the structured meetings found it useful to discuss the MCDA framework in detail, provide feedback on the study approach, and interact with peers to identify the survey responses that better represented their perspective. One group of stakeholders, the clinicians, reviewed their answers following the structured meetings and took part in the survey second round. The other groups, partly because of participants' time constraints and partly because they believed the discussion did not affect significantly their survey responses, did not complete the survey the second time.

3.1. Stakeholders' preferences between criteria - weights

Criteria weights represent trade-offs between decision criteria (Thokala et al., 2016) and thus reveal which aspects of value matter most to each stakeholder group. To obtain "generic" weights which can be applied to any intervention, respondents were required to express their preferences between the EVIDEM criteria solely based on their definition

and scale measurement (see supplementary material 1) and not the description of the intervention.

Figure 2 compares the sets of normalised weights from each stakeholder group. According to patients, the two most important criteria (out of the 13 EVIDEM core criteria) were the "type of therapeutic benefit" and "disease severity", both with weights of 11% (Standard Deviation [SD] 0.07 and 0.1, 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 0.1) 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.

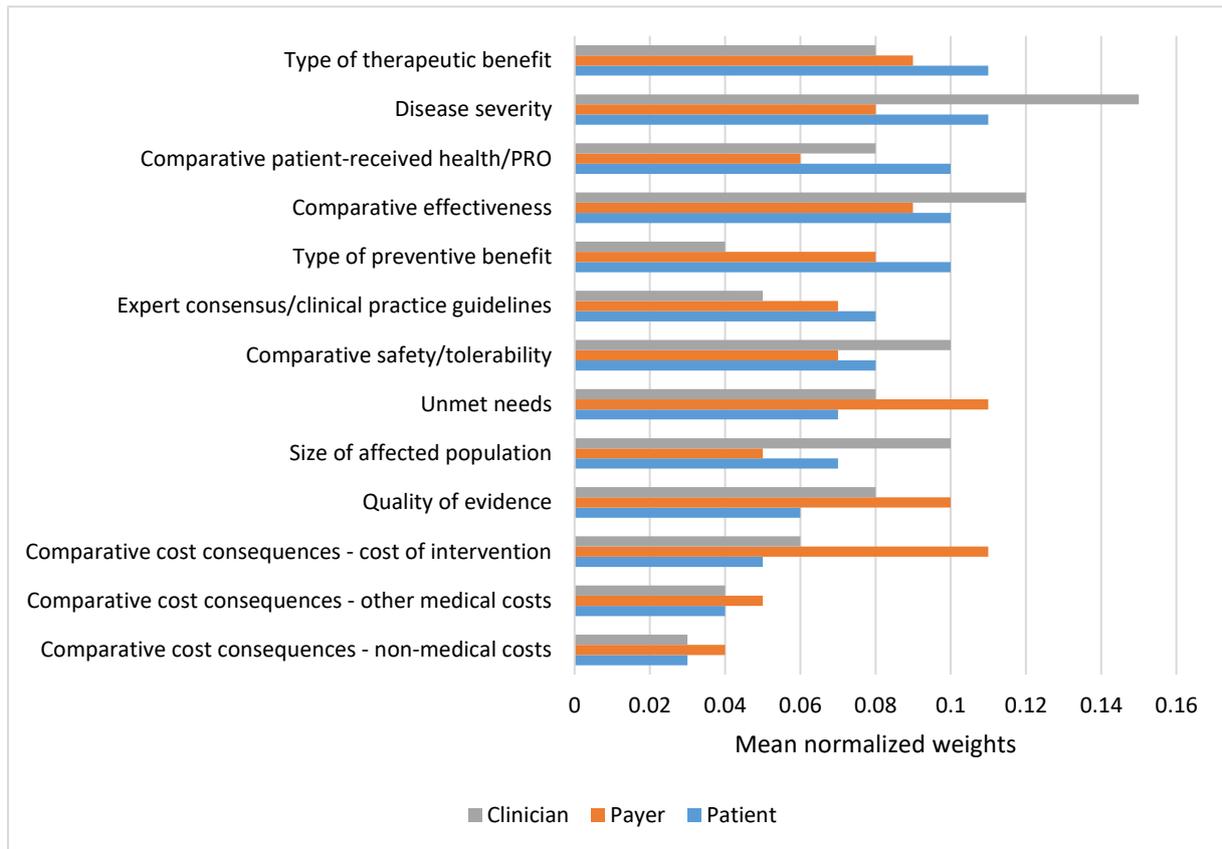
Finally, payers indicated that the three most important criteria were "unmet needs", "comparative cost of intervention", and "quality of evidence" with weights of 11% (SD 0.07), 11% (SD 0.05) and 10% (SD 0.02) respectively. The three least important criteria were "comparative non-medical costs", "size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

Given the limited sample size, we did not perform any statistical comparison across stakeholder group weights. However, it is worth highlighting some key differences and commonalities. Compared to patients and clinicians, payers distributed the weights more equally among the domains. Payers' weights range between 17% and 24% compared to 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".

The patients and clinicians views were aligned as they expressed preference for interventions targeting severe conditions, for example where patients have a short life expectancy. The highest weights in both group were observed for this criterion. Patients also believed that priority should be given to interventions which 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 to standard of care (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 currently 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.

Figure 2. Relative weights of individual criteria by stakeholder group



3.2. Stakeholders' scores of obinutuzumab for iNHL

The purpose of this stage was to convert the different outcome measures and types of evidence on obinutuzumab for the treatment of rituximab-refractory iNHL (see supplementary material 2) into criterion 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 of 3.7 (on a scale from 0 to 5); "disease severity" with an average score of 3.6 (on a scale from 0 to 5); and "type of therapeutic effect" with an average score of 3.1 (on a scale from 0 to 5). The criteria part of the "comparative outcomes of intervention" domain, including the "comparative efficacy/effectiveness" criterion, were on average all positive indicating that obinutuzumab is expected to generate incremental health gains compared to SoC. We note, however, that for the criteria on "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, with the majority of responders giving -3 (on a scale from -5 to +5).

Areas in which obinutuzumab was deemed to perform well by clinicians were: "size of affected population" with an average score of 4 (from a scale 0 to 5); "disease severity", "type of therapeutic effect", and "quality of evidence" with an average score of 3.6 (from

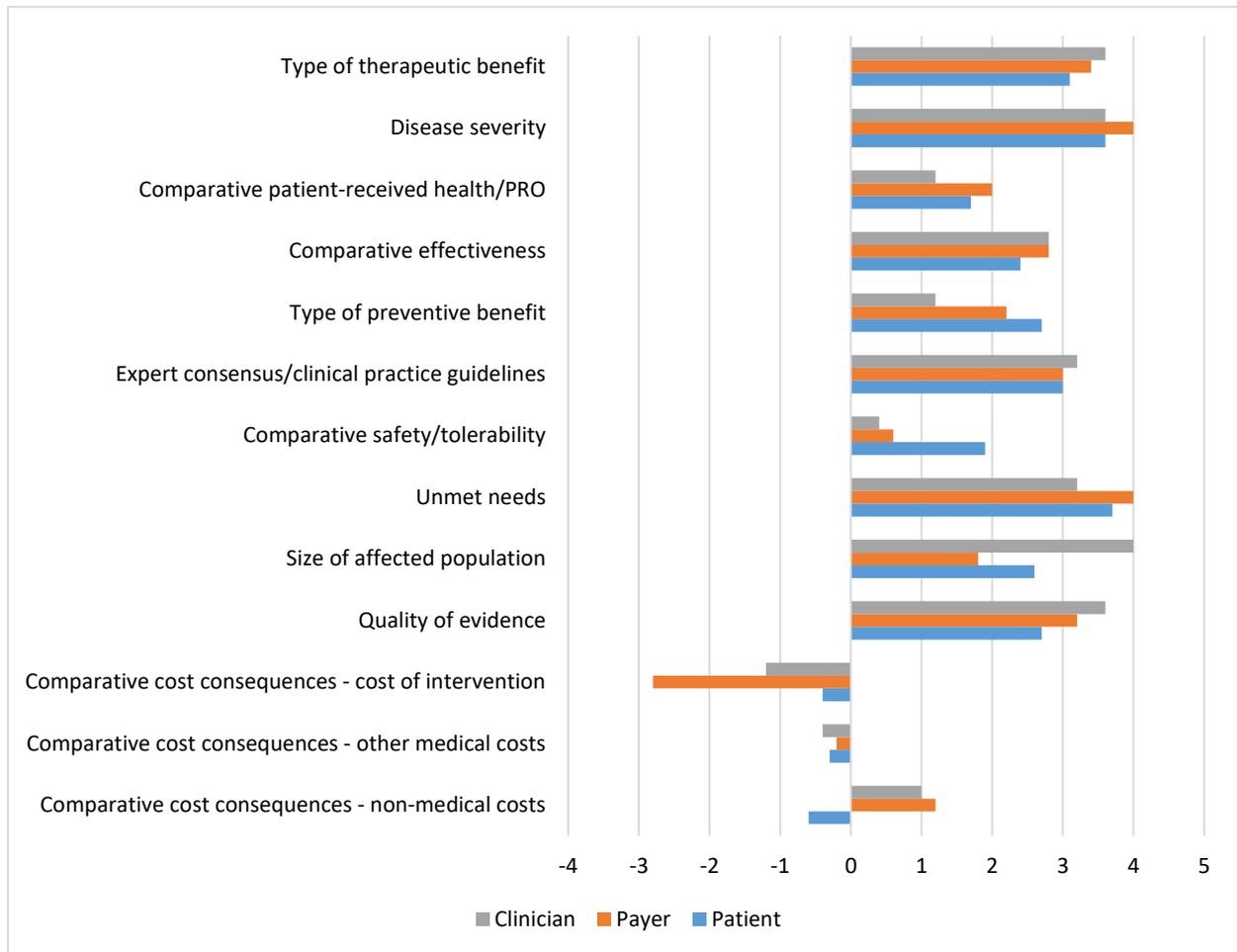
a scale 0 to 5). Two criteria part of the "comparative outcomes of intervention" domain ("comparative efficacy/effectiveness" and "patient-perceived health"), were on average positive. However, in terms of safety and tolerability, obinutuzumab was deemed slightly worse than its comparator (with a score of - 0.4). Clinicians assigned negative scores to obinutuzumab also against "comparative cost of intervention" and "comparative other medical costs", which obtained a score of -1.2 and -0.4, respectively (on a scale from -5 to +5).

Payers gave the highest scores to: "disease severity" and "unmet needs" with an average score of 4.0 (on a scale from 0 to 5); and "type of therapeutic benefit" with an average score of 3.4 (on a scale from 0 to 5). Two criteria received negative scores on average: "comparative cost of intervention" and "comparative other medical costs".

As per the group weights, we did not perform a statistical analysis but we discuss commonalities and differences across values obtained from the three groups. We observed some consistency across stakeholder groups on how well obinutuzumab performs against the EVIDEM criteria.

The criteria "disease severity", "type of therapeutic effect" were consistently assigned the highest scores by the three groups. This means that all groups believed that iNHL is a very 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 delay progression and help controlling disease symptoms. Both payers and patients thought that another area where obinutuzumab could bring value is "unmet needs": available interventions for iNHL have limitations and shortcomings (e.g. a proportion of the population does not respond to the treatment) which need to be addressed. Clinicians thought that obinutuzumab performs relatively well on "quality of evidence": 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 from a scale of -5 to +5) when compared to its comparator bendamustine, whose patent has recently expired, according to all stakeholder groups. This indicates that, based on the evidence provided, respondents thought that the cost of the obinutuzumab and related medical costs was higher than those of the comparator. This is not surprisingly given that obinutuzumab is used in combination with bendamustine and compared to bendamustine.

Figure 3: Scores for obinutuzumab for the three stakeholder groups



3.3. Overall value score of obinutuzumab

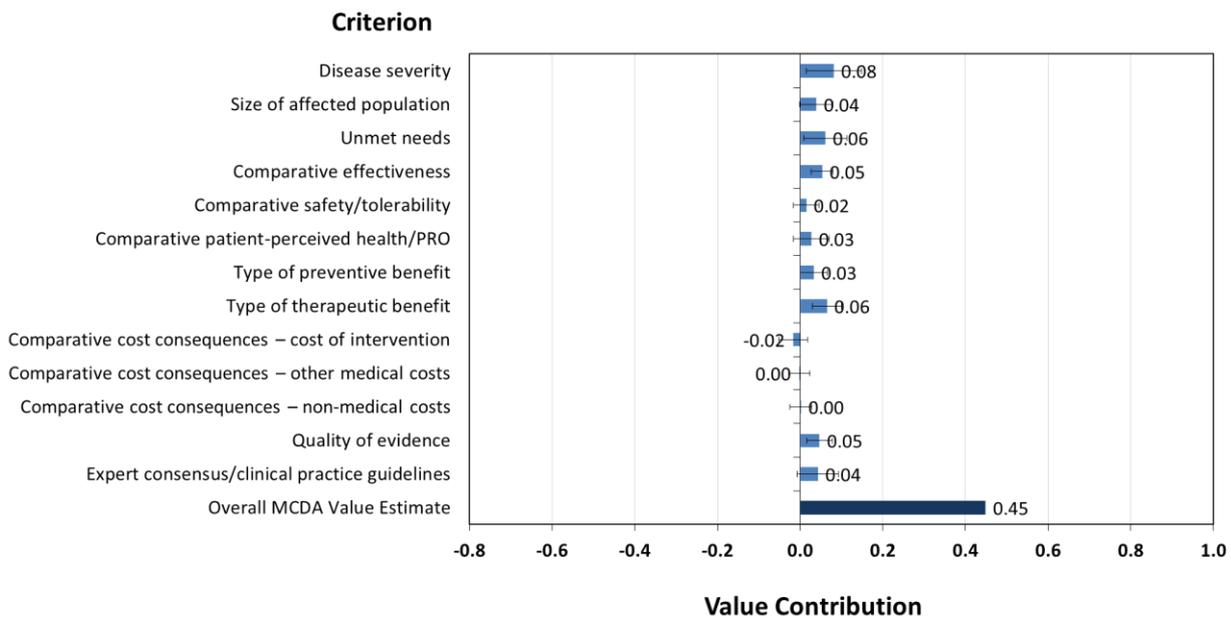
To develop a combined perspective on the value of obinutuzumab, all survey respondents were included and weighted equally. As shown in Figure 4, the value score of obinutuzumab was 0.45.

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 (Claxton, 2015). In addition, this would not allow for an appropriate consideration of the opportunity costs of the coverage decision (Marsh et al., 2016). Instead, costs can be considered separately to support decision makers 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, forthcoming)). When decision makers face a fixed budget constraint, a possible way to ensure value is maximized is to develop an aggregate measure of benefit (similar to the score presented in Figure 5) and compare that to an estimate of costs. This approach is presented by Golan and Hansen, who developed an MCDA framework which was piloted by the Israeli Advisory Committee to select new interventions to be included in the list of interventions available to all Israelis (Golan and Hansen, 2012).

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

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 (Claxton, 2015). In addition, this would not allow for an appropriate consideration of the opportunity costs of the coverage decision (Marsh et al., 2016). Instead, costs can be considered separately to support decision makers 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, forthcoming)). When decision makers face a fixed budget constraint, a possible way to ensure value is maximized is to develop an aggregate measure of benefit (similar to the score presented in Figure 5) and compare that to an estimate of costs. This approach is presented by Golan and Hansen, who developed an MCDA framework which was piloted by the Israeli Advisory Committee to select new interventions to be included in the list of interventions available to all Israelis (Golan and Hansen, 2012).

Figure 4. Value score of obinutuzumab (combined perspectives)



We conducted a sensitivity analysis in which we set the weight for the comparative costs criteria to zero, scaled up the weights for the remaining criteria clusters so that they still summed to one, and recalculated the overall score from a combined perspective. The result is shown in Figure 5.

Figure 5: "Benefit" score for obinutuzumab (sensitivity analysis)

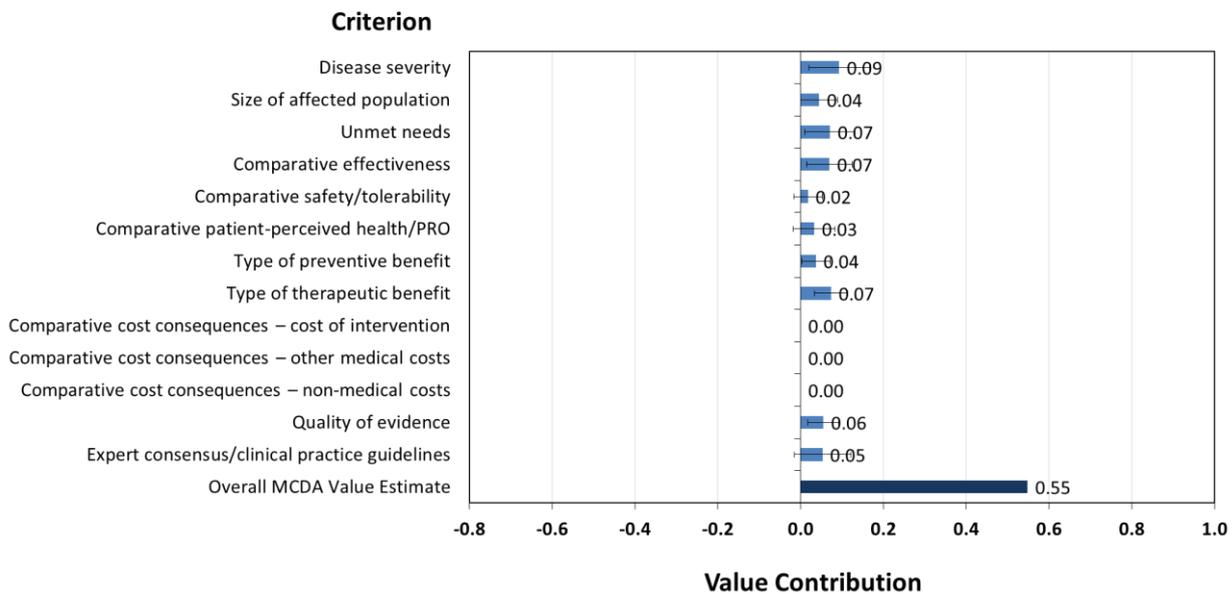


Figure 5 shows that removing the cost criteria from the calculations increases the overall value score for obinutuzumab (from all stakeholders' perspective) from 0.45 to 0.55. If this version of the framework was used, decision makers would need to assess and consider the net economic impact alongside this benefit score. This approach might also be helpful in those systems, like those followed by AIFA in Italy, where the price of the intervention is defined following its benefit assessment.

4. DISCUSSION

In many decision making systems, including the Italian one, 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 MCDAs framework such as EVIDEM can enable the collection of stakeholders' preferences (mainly via weights allocation) and ensure that they are taken into account more systematically in decision making (via determination of the value score based on these preferences 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 which 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 the value between competing technologies and rank them, or across successive decisions about different technologies by an HTA committee, score values might be useful. An example of this approach is the MCDAs framework for health technology prioritisation developed for the Israeli Advisory Committee (Golan and Hansen, 2012). 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.

In the context of the Italian NHS, EVIDEM has been implemented by the Lombardia region to make listing and de-listing decisions on new interventions including medical devices (Radaelli et al., 2014). More than 20 interventions have been appraised and have obtained value scores, including economic criteria, between 0.22 and 0.72 (Wahlster et al., 2015). 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, 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, the EVIDEM framework and 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 (Baltussen, 2015).

Currently 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 an MCDA process in practice for national level decision making, broader and larger groups of stakeholders embracing different disease areas would need to be consulted to derive the weights. Alternatively, the decision making committee (either Technical Scientific Committee or the Prices and Reimbursement Committee within 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 (weights).

5. LIMITATIONS

This study was exploratory as it applied an MCDA approach for the first time in the context of medicine reimbursement decision making in Italy in a convenience sample. In future applications of EVIDEM and potentially other MCDA frameworks, some improvements can be made to increase the validity of the results and the applicability in a formal decision making process.

Our convenience sample ensured high rate of response but it could be expanded in terms of size and clinical areas covered, for example by involving patients and clinicians of other non-oncology conditions. Survey instructions and synthesis of evidence on obinutuzumab were provided in English whilst during the meetings the participants and the moderator spoke in Italian, which helped in the interpretation of the scientific evidence and of 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 that there is a need to simplify the language used to explain the MCDA framework and make it more accessible to lay persons. This shows that the interactive component is needed not only at the end of the process, to consolidate survey responses, but also at the start, to ensure full understanding around the criteria included and their definitions. Validation with stakeholders should be part of each step of an MCDA exercise (Marsh et al., 2016).

The validation process should also ensure that criteria are "preference independent", which means 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 (Marsh et al., 2016).

To limit the cognitive burden of participants, we focused the meeting discussions on a restricted number of criteria (the 13 "core criteria" of the EVIDEM framework). 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 breadth/inclusiveness of a framework and resources needed to develop relevant evidence and analyse it for decision making (Marsh et al., 2016).

On the method to obtain weights, we selected one of the recommended methods 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 preference that have strong theoretical foundations and have been used in other types of health care assessments can be considered, such as discrete choice experiments and PAPRIKA (Golan et al., 2011).

On the criteria scoring scale, we observe that more guidance needs to be provided in order to score 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.

6. CONCLUSIONS

This is the first time that an MCDA approach has been used to assess the value of medicines and inform reimbursement decision making in Italy.

This study showed that MCDA (and 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 ones. Payers allocated higher weight (compared to patients and clinicians) to the economic criteria and to the quality of evidence. The key criteria driving the value of ominutuzumab accordingly to all the stakeholders' 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 of the decision making process. Our MCDA study provides useful evidence to decision makers, such as AIFA, on what health interventions attributes different stakeholders value the

most, and has tested methods to ensure that this is captured consistently across different interventions.

REFERENCES

- Baltussen, R. 2015. Question is not whether but how to use MCDA. *Value and Outcomes Spotlight*.
- Belton, V. and Pictet, J. 1997. A Framework for Group Decision Using a MCDA Model: Sharing, Aggregating or Comparing Individual Information? *Journal of Decision Systems*, 6, 283-303.
- Claxton, K. 2015. Three questions to ask when examining MCDA. *Value and Outcomes Spotlight*.
- Danner, M., Hummel, J. M., Volz, F., Van Manen, J. G., Wiegard, B., Dintsios, C. M., Bastian, H., Gerber, A. and Ijzerman, M. J. 2011. Integrating patients' views into health technology assessment: Analytic hierarchy process (AHP) as a method to elicit patient preferences. *Int J Technol Assess Health Care*, 27, 369-75.
- Devlin, N. and Sussex, J. 2011. Incorporating Multiple Criteria in HTA: Methods and Processes. In: OHE (ed.) *OHE Monograph*.
- EMA. 2016. *Assessment report Gazyvaro Procedure No. EMEA/H/C/002799/II/0007* [Online]. Available: http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Assessment_Report_-_Variation/human/002799/WC500211313.pdf.
- Garau, M. and Devlin, N. forthcoming. Using MCDA as a decision aid in Health Technology Appraisal for coverage decisions: opportunities, challenges and unresolved questions. In: MARSH, K. (ed.) *MCDA in healthcare decision making*. Springer.
- Gilabert-Perramon A., T.-F. J., Catalan A., Prat A., Fontanet M., Puig-peiró R., Merino-Montero S., Khoury H., Goetghebeur M., Badia X . 2016. Development of a Multi-Criteria Decision Analysis (MDDA) Framework for Health Care Decision-Making in Catalonia (Spain): Pilot Study Focused on Orphan Drugs. *ISPOR 19th Annual European Congress*. Vienna.
- Golan, O. and Hansen, P. 2012. Which health technologies should be funded? A prioritization framework based explicitly on value for money. *Isr J Health Policy Res*, 1, 44.
- Golan, O., Hansen, P., Kaplan, G. and Tal, O. 2011. Health technology prioritization: which criteria for prioritizing new technologies and what are their relative weights? *Health Policy*, 102, 126-35.
- HTAi Patient and Citizen Involvement. 2015. *Good Practice Examples of Patient and Public Involvement in Health Technology Assessment* [Online]. Available: http://www.htai.org/fileadmin/HTAi_Files/ISG/PatientInvolvement/EffectiveInvolvement/Good_Practice_Examples_Feb_2015.pdf.
- Lee, L., Wang, L. and Crump, M. 2011. 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*, 22, 1392-403.
- Marsh, K., M, I. J., Thokala, P., Baltussen, R., Boysen, M., Kalo, Z., Lonngren, T., Mussen, F., Peacock, S., Watkins, J. and Devlin, N. 2016. 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*, 19, 125-37.
- Paulden, M., Stafinski, T., Menon, D. and McCabe, C. 2015. Value-based reimbursement decisions for orphan drugs: a scoping review and decision framework. *Pharmacoeconomics*, 33, 255-69.
- Radaelli, G., Lettieri, E., Masella, C., Merlino, L., Strada, A. and Tringali, M. 2014. Implementation of EUnetHTA core Model(R) in Lombardia: the VTS framework. *Int J Technol Assess Health Care*, 30, 105-12.

Sehn, L. H., Chua, N., Mayer, J., Dueck, G., Trneny, M., Bouabdallah, K., Fowler, N., Delwail, V., Press, O., Salles, G., Gribben, J., Lennard, A., Lugtenburg, P. J., Dimier, N., Wassner-fritsch, E., Fingerle-rowson, G. and Cheson, B. D. 2016. 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*, 17, 1081-93.

Shah, K. K., Mestre-ferrandiz, J., Towse, A. and Smyth, E. N. 2013. A review of health technology appraisals: case studies in oncology. *Int J Technol Assess Health Care*, 29, 101-9.

SMC. 2015. *PACE – Patient and Clinician Engagement. Factsheet* [Online]. Available: https://www.scottishmedicines.org.uk/files/PACE/PACE_factsheet_FINAL.pdf.

Sussex, J., Rollet, P., Garau, M., Schmitt, C., Kent, A. and Hutchings, A. 2013. A pilot study of multicriteria decision analysis for valuing orphan medicines. *Value Health*, 16, 1163-9.

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

Tony, M., Wagner, M., Khoury, H., Rindress, D., Papastavros, T., Oh, P. and Goetghebeur, M. M. 2011. 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*, 11, 329.

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

Wahlster, P., Goetghebeur, M., Schaller, S., Kriza, C. and Kolominsky-rabas, P. 2015. 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*, 13, 24.