What Does the Public Want? Structural Consideration of Citizen Preferences in Health Care Coverage Decisions

Irina Cleemput, Stephan Devriese, Laurence Kohn, Carl Devos, Janine van Til, Catharina G. M. Groothuis-Oudshoorn, and Carine van de Voorde

Abstract

Background. Multi-criteria decision analysis can improve the legitimacy of health care reimbursement decisions by taking societal preferences into account when weighting decision criteria. This study measures the relative importance of health care coverage criteria according to the Belgian general public and policy makers. Criteria are structured into three domains: therapeutic need, societal need, and new treatments' added value. **Methods.** A sample of 4,288 citizens and 161 policy makers performed a discrete choice experiment. Data were analyzed using multinomial logistic regression analysis. Level-independent criteria weights were determined using the log-likelihood method. **Results.** Both the general public and policy makers gave the highest weight to quality of life in the appraisal of therapeutic need (0.43 and 0.53, respectively). The general public judged life expectancy (0.14) as less important than inconvenience of current treatment (0.43), unlike decision makers (0.32 and 0.15). The general public gave more weight to "impact of a disease on public expenditures" (0.65) than to "prevalence of the disease" (0.56) when appraising societal need, whereas decision makers' weights were 0.44 and 0.56, respectively. When appraising added value, the general public gave similar weights to "impact on quality of life" and "impact on prevalence" (0.37 and 0.36), whereas decision makers judged "impact on quality of life" (0.39) more important than "impact on prevalence" (0.29). Both gave the lowest weight to impact on life expectancy (0.14 and 0.21). **Limitations.** Comparisons between the general public and policy makers should be treated with caution because the policy makers' sample size was small. **Conclusion.** Societal preferences can be measured and used as decision criteria weights in multi-criteria decision analysis. This cannot replace deliberation but can improve the transparency of health care coverage decision processes.

Coverage decisions in health care are complex. Decision making involves not only evaluation of safety, clinical effectiveness, and organizational issues related to implementation of the intervention but also tradeoffs with regard to the importance of different outcomes.[1] Making tradeoffs to reach a final decision is often referred to as "appraisal." There is a growing interest in applying multi-criteria decision analysis (MCDA) to increase consistency and transparency of appraisal processes. MCDA involves the identification of the relevant decision criteria, the determination of the relative weight of the

Belgian Health Care Knowledge Centre, Brussels, Belgium (IC, SD, LK, CD, CVV); Health Technology and Services Research, Mira Institute, University of Twente, Enschede, Netherlands (JVT, CGMG). This work was performed at the Belgian Health Care Knowledge Centre (KCE) in Belgium. The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article. No external financial support was received for this study. The study was entirely performed with resources of the Belgian Health Care Knowledge Centre, the employer of most of the authors.

Corresponding Author
Irina Cleemput, Belgian Health Care Knowledge Centre, AC Kruidtuin, Doorbuilding (10th Floor), Kruidtuinlaan 55, 1000 Brussels, Belgium; +32 1126 8714 (Irina.Cleemput@kce.fgov.be).

Table 1 Key Questions and Possible Criteria for a Health Care Coverage Appraisal Process[a]

Decision	Question	Possible Criteria
Therapeutic and societal need for a better treatment for the condition	Does the product target a therapeutic and/or societal need?[9]	*Therapeutic need*: Effectiveness of best available current treatment, inconvenience of current treatment *Societal need*: Prevalence of the disease, health inequality, baseline health level
Preparedness to pay out of public resources for *a* treatment for the condition	Are we, as a society, in principle, prepared to pay for a treatment that will improve this indication out of public resources?	Own responsibility, life-style related condition
Preparedness to pay out of public resources for the treatment under consideration targeting the condition	Are we, as a society, prepared to pay for this particular treatment, given that we in general would be prepared to pay for a treatment for this indication?	Added value of the new treatment compared with the best treatment currently available for the condition: safety, efficacy, therapeutic benefit, significance of health gains, curative, symptomatic, preventive
Preparedness to pay more for the treatment under consideration	Given that we, as a society, are prepared to pay for this treatment out of public resources, are we prepared to pay more for this treatment than for the best alternative treatment?	Added value of the new treatment, potentially induced savings elsewhere in the health care sector, quality and uncertainty of the evidence, acceptability of co-payments and/or supplements, rarity of disease
Willingness to pay (price and reimbursement basis)	How much more are we willing to pay out of public resources for this particular treatment?	Added therapeutic value, budget impact/ability to pay, cost-effectiveness ratio, medical, therapeutic and societal need, quality and uncertainty of evidence, limits to cost sharing

[a]Adapted from Cleemput et al. (2012).[9]

criteria, the scoring of options on the criteria, and the aggregation into an overall score allowing the ranking of different options.[2,3] The weights reflect the relative importance of the criteria for the decision. MCDA is currently being used by VTS-HTA, the HTA agency in Lombardy (Italy) for reimbursement decisions.[3] Colombia is also piloting the use of MCDA for coverage decision making,[4] and several other countries are exploring the possibilities of implementing MCDA for reimbursement decisions.[5–7]

In 2010, the Belgian Health Care Knowledge Centre published a multi-criteria framework to support transparent and consistent health care coverage decision making.[8] The framework consists of five key questions, splitting the complex decision problem into manageable components that each go with a set of decision criteria (Table 1).[9]

Therapeutic need refers to the need for a better intervention than the one already available from the patients' point of view.

Societal need refers to the need for a better intervention than the one already available from the society's point of view.

Added value refers to the extent to which an intervention reduces therapeutic and societal need.

The framework has a sequential logic. For example, society is unlikely to be willing to pay (more) for a treatment people do not need. Therefore, the answer to the first question, that is, whether or not patients actually have a need for a better treatment than the one that is currently available, must be positive before reimbursement should even be considered. The framework allows to go beyond cost-effectiveness analysis in a transparent manner, incorporating explicit criteria that are not covered by the cost-effectiveness ratio.

The scoring of diseases and treatments on the different criteria is done by an appraisal committee and is based on scientific evidence with respect to the effect of a disease or treatment on each criterion. However, the answer to each question in the framework also requires an appraisal that considers the relative importance of each decision criterion. Policy makers are expected to make legitimate decisions in a transparent manner and in the best interest of the population. There is a growing trend to involve citizens in decision-making processes, including coverage decision making in health care, to reach this

goal. This is exemplified by several international and national initiatives to explore better ways to include citizen's perspectives in decision-making processes, such as the Health Technology Assessment International (HTAi) Patient and Citizen Involvement Interest Group,[10] the Public Consultation (Consulta Pública) installed in Brazil,[11] the Consumers Health Forum in Australia,[12] and the Public Involvement Programmes in England[13] and Scotland.[14]

Although little is known about the necessary conditions for public involvement in decision-making processes for increasing the legitimacy of decisions, it is often considered that public involvement could increase acceptance of the decisions.[15] Public involvement is proposed either directly through public or patient representation in a decision-making body or indirectly through the incorporation of preference values in the decision-making processes of elected or appointed decision makers.[16,17] The use of public preference values in a multicriteria framework is one way to indirectly include the public perspective in coverage decisions.[15,18,19] It allows at the same time to increase the transparency of the decision processes. Public preference values thus support decision makers by giving guidance about how treatments and diseases could be prioritized, in an MCDA context that covers more than the costs and effectiveness of treatments.

If policy makers wish to take public preferences into account and use them to make transparent decisions, data are needed. Currently, little data are available about the general public's preferences, and policy makers need to rely on their own perspective of what is important for the population, possibly informed by media as a source of public opinion. Compared with coverage decisions based on cost-effectiveness ratios only, or cost-effectiveness combined with other, implicit criteria, incorporating public preferences for multiple explicit criteria might be a game changer. It might change the way decision makers deal with decision criteria, as they no longer have to rely on their personal perceptions of the relative importance of different aspects of a disease or treatment but can use the preference weights as an external source of data to inform their decision-making process. This can improve the transparency and consistency of the decision process.

The current study addresses how public preferences could be obtained if policy makers choose to take public preferences into account but have only a vague idea about these preferences. The operational aim of the current study is thus to determine the relative importance of the appraisal criteria for therapeutic need, societal need,

and added value of new treatments, from the perspective of the general public in Belgium. To study the alignment of public and policy maker preferences, a comparison is made with the criteria weights of policy makers.

Methods

The following steps were taken to obtain the criteria weights: 1) identification of the relevant decision criteria; 2) development of a questionnaire to elicit the relative weights of the criteria from the general public and from decision makers; 3) administration of this questionnaire in a sample of the general public and in a sample of the decision makers; and 4) derivation of the criteria weights by means of statistical modeling.

Identification of Relevant Decision Criteria to Be Included in the MCDA

The relevant criteria for the appraisal of therapeutic need, societal need, and added value were identified through a literature review and a workshop with six experts.

The databases Medline, Embase, and Sociological Abstracts were searched using Mesh and Embase terms for "social values" and "health priorities." Papers or documentation on existing MCDA frameworks were searched in gray literature through a hand-search, including websites of known initiatives such as EVIDEM (www.evidem.org). The criteria were classified into the domain "therapeutic need," if they related to the individual patient, into the domain "societal need" if they related to the society at large or into the domain "added value" if they related to a treatment. The three lists were considered to be the long-list of criteria.

The experts invited to the workshop had different backgrounds: sociology (with extensive expertise in survey research), public health, philosophy, communication toward lay public (journalism), biostatistics, and biomedical research (with extensive expertise in MCDA). They were selected based on their relevant experience as demonstrated by publications in peer-reviewed literature or public reports. In advance of the workshop, the experts received a document describing the background and objectives of the study, a brief description of the methods, an overview of existing MCDA frameworks, and the long-list of criteria resulting from the literature review. The long-list was used as a starting point for discussion during the workshop. Criteria were added if deemed necessary by the expert group.

MCDA requires that criteria are clearly defined and based on clearly articulated principles, operational (i.e., it must be possible to describe or measure the characteristics of the options that decision makers are considering in terms of these criteria), mutually exclusive (i.e., they should not just be alternative measures or proxies of the same underlying principle: each criterion covers one and just one dimension of potential interest), and criteria should be preferentially independent (i.e., the state of one criterion should not influence respondents' preferences for other criteria). Following the workshop, the draft list of criteria was checked for consistency with the requirements for MCDA and possible ways to operationalize and define these criteria were discussed. Criteria were only included if they could be operationalized in a way that was thought to be comprehensible for all citizens.[20]

Questionnaire Development

The survey consisted of the usual demographic questions and three different discrete choice experiments (DCE), one for each domain. In a DCE, respondents are asked to choose between two different hypothetical alternatives, where each alternative is described by a set of attributes (characteristics), which are considered to be appraisal criteria.

Attributes. The type of choice and the attributes for the scenarios differed between domains (Figure 1). For therapeutic need, the scenarios described different patient groups and respondents were asked to choose the group that, from their perspective, had the highest need for a better treatment. For societal need, the scenarios described different diseases in terms of impact on quality and quantity of life and respondents were asked to choose the disease in which the need for a better treatment was the highest. For added value, the scenarios described the characteristics of two new treatments in terms of effectiveness, safety, and ease of use and respondents were asked to choose the treatment that they would most prefer to be reimbursed, if the two treatments were for the same disease. The alternative chosen by a respondent is considered to have the highest added value according to the respondent.

The alternatives were unlabeled for two main reasons: 1) to reduce heterogeneity in tradeoffs due to the specific label given to a scenario and 2) to be able to classify diseases in generic categories, to avoid having to repeat the survey every time a disease-specific decision is made.[21]

Figure 2 shows an example of a question for each domain included in the survey.

An English translation of the original Dutch and French questionnaire is available as an online appendix (http://kce.fgov.be/sites/default/files/page_documents/KCE_234S_reimbursement_Decisions_Appendix_0.pdf).

DCE Design. The selection of combinations of levels for the different attributes to be included for each domain in the survey was made based on an analysis of what would be needed for a D-optimal design for a main effects and nonlinear two-way interactions model. As a full factorial design using each combination of the levels of all attributes was not feasible because of the relatively high number of attributes and levels included in the survey,[22] we first removed dominant choice sets (see below) in the therapeutic need and added value domains, and choice sets with more than one (societal need) or two (therapeutic need and added value) overlapping attributes. For therapeutic need, this led to 72 combinations, for societal need 24 combinations, and for added value 96 combinations.

To keep response burden acceptable but still cover all the combinations necessary for the D-optimal design, we used 24 different versions of the questionnaire. Each version had 3 choice sets for therapeutic need, 1 for societal need and 4 for added value. In this way observations were obtained for all combinations. To ensure representativeness on each of the 24 versions, people of the same sex and age category received subsequent versions in the order of logging into the web survey. The details on the software used for constructing the DCE design and survey versions are given in the section on statistical analysis.

A dominant choice set, where one of the alternatives presented is superior on all attributes, was included in the added value domain to perform a consistency check. People who did not choose the dominant alternative did not pass the consistency check and were excluded from the analysis.

The survey development process included a pretest, a pilot test, and a test-retest phase. Both the pretesting ($N = 20$) and pilot testing ($N = 219$) were meant to improve the comprehension, presentation, and feasibility of the questionnaire. The respondents were chosen among the acquaintances of employees of the Belgian Health Care Knowledge Centre (KCE) and had variable educational and socioeconomic backgrounds. A test-retest was performed in 42 KCE employees to test the reliability of the questionnaire.

Respondent Samples

The public sample consisted of a representative sample of the general Belgian public. The sample was drawn

Therapeutic need: impact of the illness on patients, given current standard treatment

Inconvenience of current treatment
-   Patients experience much discomfort from current treatment
-   Patients experience little discomfort from current treatment

Quality of life with current treatment
-   Patients currently have a quality of life of 8 out of 10
-   Patients currently have a quality of life of 5 out of 10
-   Patients currently have a quality of life of 2 out of 10

Life expectancy
-   Patients do not die earlier from the disease
-   Patients die 5 years earlier than people without the disease
-   Patients die almost immediately from the disease, despite current care

Age group
-   Patients are older than 80 years of age
-   Patients are between 65 and 80 years of age
-   Patients are between 18 and 64 years of age
-   Patients are younger than 18 years of age

Societal need: characteristics of the illness that impact the society

Prevalence
-   The disease is rare: less than 2000 people in Belgium have the disease
-   The disease is not so frequent: between 2000 and 10 000 people in Belgium have the disease
-   The disease is rather frequent: between 10 000 and 100 000 people in Belgium have the disease
-   The disease is very frequent: more than 100 000 people in Belgium have the disease

Disease-related public expenditures
-   Low public expenditures per patient
-   High public expenditures per patient

Added value of the new treatment compared to the existing alternative

Impact on the inconvenience of treatment for the patient
-   The new treatment is less inconvenient for the patient than the already available treatment
-   The new treatment is as inconvenient for the patient as the already available treatment
-   The new treatment is more inconvenient for the patient than the already available treatment

Impact on the quality of life
-   Compared to the existing treatment, the new treatment improves the quality of life of patients
-   Compared to the existing treatment, the new treatment does not change the quality of life of patients
-   Compared to the existing treatment, the new treatment reduces the quality of life of patients

Impact on life expectancy
-   Compared to the existing treatment, the new treatment does not change the life expectancy of patients
-   Compared to the existing treatment, the new treatment increases the life expectancy of patients

Impact on the prevalence of the disease
-   The new treatment cures fewer patients than the existing treatment
-   The new treatment cures as many patients as the existing treatment
-   The new treatment cures more patients than the existing treatment

Impact on public expenditures
-   The new treatment reduces the disease-related public expenditures per patient
-   The new treatment does not change the disease-related public expenditures per patient
-   The new treatment increases the disease-related public expenditures per patient

Inconvenience of treatment refers to the inconvenience caused by for instance frequency of use (e.g. taking a drug once or more times a day), the administration route (e.g. syringes, oral drugs, administration by yourself or by someone else), the place of administration (at home, in the hospital, in a doctor's cabinet). **Quality of life** was described in the survey as having five dimensions: mobility, self-care, usual activities, physical pain and mental suffering (anxiety/depression), as in the EQ-5D, a generic health-related quality of life instrument. **Age** was added as a descriptive fact to allow respondents to judge the importance of losing 5 years of life expectancy. As age is not a direct decision criterion for reimbursement *per se* (only indirectly, e.g. through the effect of age on incremental effectiveness), age can be correlated with the decision. **Frequency of disease** refers to prevalence of the disease. It was included to examine to what extent "rarity" as such is important for defining needs, independent from disease severity. **Disease-related public expenditures** were defined as the total public expenditures per patient with the disease, including health care expenditures and productivity losses leading to benefits that replace wage losses, etc. The attributes for added value correspond to the improvement on the attributes of therapeutic and societal need. The rationale behind this choice was that if criteria were considered relevant for determining therapeutic and societal need, improvements on these criteria would also be relevant for determining the added value of a new treatment.

Figure 1 Domains, attributes, and levels used in the survey.

Therapeutic need

Two patient groups are described below. Both patient groups currently receive treatment. The discomfort associated with the treatment, the quality of life and life expectancy of patients getting this treatment and the typical age of patients with this condition are as follows:

Patients of group 1	**Patients of group 2**
• have a quality of life of **8 on 10**	• have a quality of life of **5 on 10**
• experience **much** discomfort from treatment	• experience **little** discomfort from treatment
• are between **18 and 64 years of age**	• are **older than 80** years of age
• **no longer** die from the disease	• **no longer** die from the disease

For which patients do you consider it most important to develop a new and better treatment? You may define yourself what you consider to be "better".

Choose one group.

□ Patients of group 1	□ Patients of group 2

Societal need

Two diseases are described as follows:

Disease 1	**Disease 2**
is **not so** frequent: between 2000 and 10 000_people in Belgium have the disease.	The disease is **rather** frequent: between 10 000 and 100 000 people in Belgium have the disease.
Every patient costs **little** to society.	Every patient costs **much** to society.

For which disease do you consider it most important to develop a new and better treatment? You may define yourself what you consider to be "better".

Choose one disease.

□ Disease 1	□ Disease 2

Added value of treatment

Suppose two new treatments appear on the market for the same disease. There is already a treatment available for this disease, which is fully reimbursed by the health insurance.

You can decide yourself which of the two new treatments the health Insurance should reimburse.

There is only enough money to reimburse one of the two treatments. Patients who wish to receive the treatment you do not choose, will have to pay for this themselves.

New treatment 1	**New treatment 2**
The new treatment, compared to the already available treatment,	The new treatment, compared to the already available treatment,
• gives **as much** discomfort to the patient	• **increases** the discomfort of treatment for the patient
• **does not change** the quality of life of patients	• **improves** the quality of life of patients
• **reduces** the cost of each patient to society	• **increases** the cost of each patient to society
• cures an **equal** number of patients	• cures **fewer** patients
• **increases** the life expectancy of patients	• **does not change** the life expectancy of patients

In your opinion, which treatment should be reimbursed?

Choose one treatment.

□ New treatment 1	□ New treatment 2

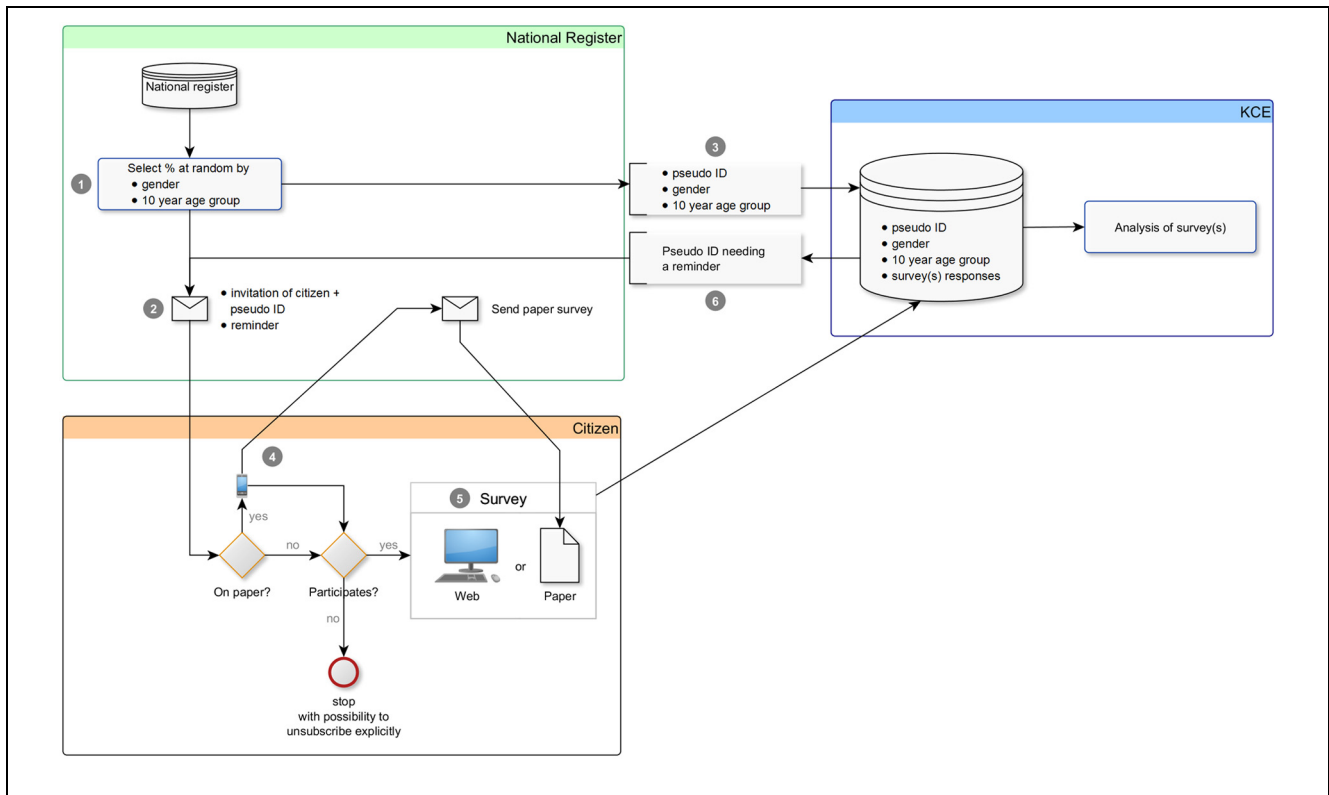**Figure 2** Example of a DCE question for each domain.

Figure 3 Survey process.

from the Belgian National Registry and consisted of a representative sample of 20,000 people between 20 and 89 years of age, stratified by age and sex. This initial sample size was determined by the DCE design: To estimate the models with sufficient power, at least 1000 valid responses were needed. A conservative response rate of 5% was assumed. The Privacy Commission gave approval for the sampling and survey process (Figure 3). The survey was anonymous. All subjects in the sample were contacted by regular mail to participate in a web or paper survey as preferred. The web survey was implemented in LimeSurvey (https://www.limesurvey.org). Three reminders were sent to nonresponders, with intervals of 2 weeks.

The decision makers' sample consisted of all members of nine public decision-making or advisory bodies embedded within the National Institute for Health and Disability Insurance, the Federal Public Service Public Health, the policy unit of the minister of Public Health, the Chamber of Representatives, and the Senate (N = 421). They were invited by e-mail and were asked to respond as representatives of the group they represent in the committee of which they are a member. Their answers were analyzed separately from those of the general public.

Statistical Analysis

Test-retest reliability of the survey was assessed by means of Cohen's Kappa.

To obtain criteria weights, a two-step procedure was used. First a main effects multinomial logit model was estimated.* Second, the log-likelihood of the estimated models was used to calculate criteria weights.

Step 1: Fitting the Multinomial Logit Regression Model. The multinomial logit regression model contained only alternative-specific variables, representing the attributes of the choice sets. All main effects of the

---

* Also other models were tested, but these could not all be reported in this article. Full information on all models tested is available from https://kce.fgov.be/publication/report/incorporating-societal-preferences-in-reimbursement-decisions-%E2%80%93-relative-importan#.WRW0BPwjHuo.

attributes were included. The model has the general form of

$$P_{ij} = \frac{e^{\left(\beta X_{ij}^T\right)}}{\sum_k e^{\left(\beta X_{ik}^T\right)}}$$

where the dependent variable of the multinomial logit model represents the probability P that respondent i chooses alternative j out of k alternatives (two in this study). β is the matrix of estimated coefficients, X_{ik}^T the transposed matrix of attribute values as presented to respondent i in alternative j. This model gives a coefficient for each level of each attribute.

Second, we used an algorithm based on differences in log-likelihood to derive the level-independent attribute-specific weights.[23]

The coefficients for the model parameters were estimated by full information maximum likelihood method using the Newton-Ralphson numerical optimization routine.[24] For the general population, each model was estimated a second time with a weight correcting for age and gender distribution to correspond with the Belgian population. If the results were very similar, the unweighted models were used.

No intercept was included in the model because the alternatives in our DCE are unlabeled.[25] Including an intercept would mean that the same attribute levels could have a different impact on the probability of choosing a disease. However, this would not make sense because the labels of the alternatives presented—"disease 1" and "disease 2"—have no meaning in themselves.

Effect coded contrasts were used for the model parameters of the attribute values.[25,26] One advantage was that coefficient estimates and standard errors could be calculated for all levels of an attribute, because in effect coding all coefficient estimates for an attribute sum to zero. For the estimation process, however, effect coding uses $n - 1$ levels per attribute (with n being the number of levels of an attribute). We estimated the coefficient and standard deviation of the omitted attribute level but did not calculate the t value as this is typically not explicitly part of the estimation process in case of effect coding.

The model fit was assessed in two ways: first, by comparing the observed proportions with the model predicted proportions of the two alternatives; second, by calculating the percentage of the choices correctly predicted by the model by comparing per choice set included in the 24 versions of the questionnaire the actual alternative chosen and the alternative with the largest probability of being chosen as predicted by the model.

Step 2: Calculating Criteria Weights. For the calculation of level-independent criteria weights, we used the log-likelihood method.[23] For this, we first calculated the log-likelihood for the full model. Then, we calculated the log-likelihood for a reduced model, that is, the model *minus* the attribute of interest. We tested if the reduced model is statistically equal to the full model with the likelihood ratio test. If the test rejects the equality hypothesis, the relative importance of the removed attribute can be considered to be different from zero. Finally, we calculated the difference in log-likelihood between the full and each reduced model as a measure of relative importance of the attribute, and converted this to a proportion.

$$W_{attribute_i} = \frac{\left|\ell_{full}\right| - \left|\ell_{A_i}\right|}{\sum_j \left(\left|\ell_{full}\right| - \left|\ell_{A_j}\right|\right)}$$

with A_i the reduced model excluding attribute i and j the number of attributes.

The relative weights of different criteria were calculated for the entire sample. Additionally, the relative weights were calculated for subgroups of respondents, defined by self-reported age category and own health status ("not in good health" and "in good health"). For each of the subgroups, the model was re-estimated and the weights were recalculated for the particular subgroup. This allowed comparisons of the weights between subgroups and between subgroups and the entire sample.

The comparison of the age and gender distributions of the sample and those of the general population was made with χ^2 tests.

All analyses were conducted in R 3.1.1, using the packages AlgDesign 1.1-7.2, car 2.0-21, lattice 0.20-29, mlogit 0.2-4, plyr 1.8.1, reshape2 1.4, sqldf 0.4-7.1, and vcd 1.3-1, in addition to the default packages.

This study was performed without external funding.

Results

Survey Reliability

The test-retest showed good overall reliability (Cohen's Kappa = 0.7, approximate 95% confidence interval: 0.62–0.77). Over all choice sets, the majority of the respondents chose the same alternative in test and retest, although the correspondence varied between questions.

Response and Sample Characteristics

Of all invited citizens, 4,810 started completing the survey and 4,485 (22.4%) answered all choice sets. Of these,

Table 2 Respondent Characteristics[a]

Item	Level	General Population Sample		Decision Makers' Sample	
		n	%	*n*	%
Response medium	Web	3,918	91.4%	160	100.0%
	On paper	370	8.6%		
Age and gender					
Female	21–30	379	8.8%		
	31–40	351	8.2%	10	6.3%
	41–50	441	10.3%	17	10.6%
	51–60	482	11.2%	19	11.9%
	61–70	375	8.7%	10	6.3%
	71–80	136	3.2%		
	81–90	68	1.6%		
Male	21–30	261	6.1%	<8	<2%
	31–40	323	7.5%	<8	<2%
	41–50	384	9.0%	13	8.1%
	51–60	467	10.9%	42	26.3%
	61–70	387	9.0%	33	20.6%
	71–80	176	4.1%	10	6.3%
	81–90	58	1.4%		
Self-reported health status	Not provided by respondent	<8	<1%		
	Very bad	<30	<1%		
	Bad	176	4.1%		
	Mediocre	785	18.3%	16	10.0%
	Good	2,241	52.3%	74	46.3%
	Very good	1,058	24.7%	70	43.8%

[a]Some cells have been obfuscated for privacy reasons.

52.1% were women (compared with 51.3% in the total population between 20 and 89 years of age). One hundred and ninety-seven respondents failed the consistency check and were hence excluded for analysis. A net sample of 4,288 respondents (21.4%) was obtained. Sample characteristics for both the general population sample and the decision makers' sample are presented in Table 2.

The proportion of male respondents was comparable to that of the Belgian population (χ^2[1 df] = 1.05, P = 0.31) but the proportions of respondents per age category differed (χ^2[6 df] = 170, $P < 0.01$; see Figure 4).

The majority of the respondents who answered all choice questions participated through the web (slightly over 91%), although a nonnegligible number of respondents asked for a paper version (almost 400 people). For comparison, according to the statistics of the Belgian Federal Public Service Economy, 87% of the Belgian citizens regularly access the Internet.[27] In the group of decision makers, 175 (41.6%) participated in the survey, of which 161 (38.2%) answered all choice sets. One respondent did not pass the consistency check. The advisory committees preparing health care coverage decisions had the highest participation rate (slightly more than 45%).

Response was lowest in the committees with a remit extending beyond health care decision making, such as the policy unit of the Minister of Public Health.

Eleven percent of the respondents in the general population sample and 5% in the decision makers' sample reported having a serious illness. None of the decision makers rated his/her health as bad or very bad. In the general population sample, a small minority rated his/her health as bad (4.1%) or very bad (0.6%). These proportions are very similar to those in the Health Interview Survey 2013, an interview survey conducted among 10,000 Belgian citizens.[28]

Modelling Results

Therapeutic Need. As expected, both groups considered the therapeutic need to be the lowest in people with a good quality of life given current treatment (i.e., a quality of life score of 8 on a scale from 0 to 10), who do not die from their disease, and with little treatment inconvenience. Both the public and the decision makers gave the highest weight to the criterion "quality of life with current treatment" (Table 3). The order of relative
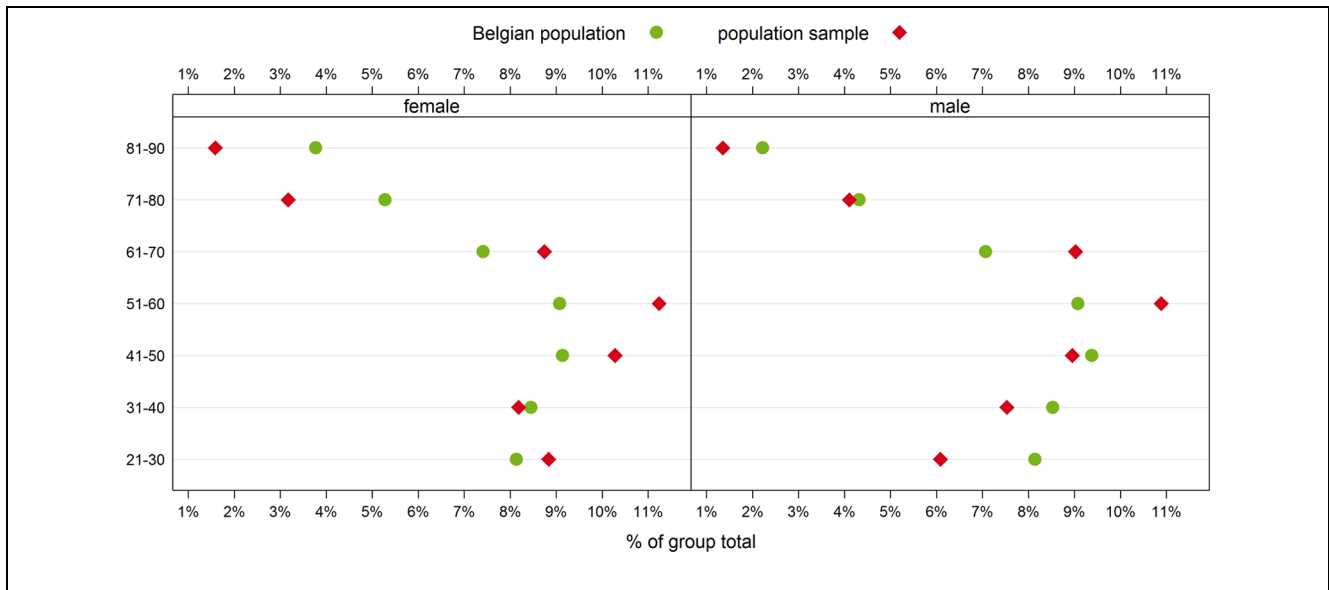
Figure 4 Age and gender distribution of the general population sample compared with the Belgian population.

Table 3 Weights (Rank) for Criteria in the Therapeutic Need, Societal Need, and Added Value Domains

	General Population	Decision Makers
Therapeutic need		
Life expectancy	0.14 (3)	0.32 (2)
Quality of life	0.43 (1)	0.53 (1)
Inconvenience current treatment	0.43 (1)	0.15 (3)
Societal need		
Public expenditure	0.65 (1)	0.44 (2)
Prevalence	0.35 (2)	0.56 (1)
Added value		
Change in quality of life	0.37 (1)	0.39 (1)
Change in prevalence	0.36 (2)	0.29 (2)
Change in life expectancy	0.14 (3)	0.21 (3)
Impact on public expenditures	0.07 (4)	0.08 (4)
Impact on inconvenience of treatment	0.06 (5)	0.03 (5)

importance of the two other attributes for therapeutic need differed between the decision makers and the general population.

The coefficients of the multinomial logit model showed that people from the general public did not seem to distinguish "dying 5 years earlier than patients without the disease" from "dying immediately from the disease" in their appraisal; both had a similar impact on therapeutic need (Table 4).

Societal Need. The weights for the societal need criteria showed that the rank order of societal need criteria

differed between decision makers and the general public. The general public attached higher importance to the impact of the disease on public expenditures per patient (weight 0.65) than to the prevalence of the disease (weight 0.35) when assessing the need for a better treatment from a societal point of view (Table 3). This means that, if two diseases are similar in all respects, except for their prevalence and disease-related public expenditures per patient, the public would consider the societal need highest for the disease with the highest cost per patient, even if the prevalence is lower.

The coefficients for rare disease and for not so frequent disease were negative, while those for rather

Table 4 Therapeutic Need: Model Summary for the General Population and Decision Maker Sample

Attribute	Level	General Population (Estimated Coefficient[a] With Confidence Interval and Significance Level)	Decision Makers (Estimated Coefficient[a] With Confidence Interval and Significance Level)
Age (years)	>80	−1.29 (CI: −1.35, −1.24)	−1.29 (CI: −1.59, −0.98)
	65–80	0.005 (CI: −0.04, 0.05)	−0.004 (CI: −0.23, 0.23)
	18–64	0.60 (CI: 0.55, 0.66)***	0.76 (CI: 0.43, 1.09)***
	<18	0.69 (CI: 0.63, 0.74)***	0.53 (CI: 0.24, 0.83)***
Quality of life given current treatment	8 out of 10	−0.31 (CI: −0.36, −0.26)	−0.47 (CI: −0.74, −0.19)
	5 out of 10	0.06 (CI: 0.02, 0.10)**	0.09 (CI: −0.11, 0.30)
	2 out of 10	0.25 (CI: 0.21, 0.29)***	0.37 (CI: 0.18, 0.57)***
Life expectancy given current treatment	Disease has no impact on life expectancy	−0.19 (CI: −0.23, −0.15)	−0.37 (CI: −0.60, −0.15)
	Patients die 5 years earlier than people without the disease	0.09 (CI: 0.05, 0.14)***	0.11 (CI: −0.12, 0.35)
	Patients die almost immediately	0.09 (CI: 0.05, 0.13)***	0.26 (CI: 0.05, 0.47)*
Inconvenience of current treatment	Little	−0.24 (CI: −0.28, −0.20)	−0.19 (CI: −0.38, −0.005)
	Much	0.24 (CI: 0.21, 0.27)***	0.19 (CI: 0.052, 0.33)**

CI, confidence interval.
[a]Results of a multinomial logistic regression model.
**$P < 0.01$. ***$P < 0.001$.

Table 5 Societal Need: Model Summary for the General Population and Decision Maker Sample

Attribute	Level	General Population (Estimated Coefficient[a] With Confidence Interval and Significance Level)	Decision Makers (Estimated Coefficient[a] With Confidence Interval and Significance Level)
Prevalence	Rare	−0.68 (CI: −0.77, −0.59)	−0.92 (CI: −1.35, −0.48)
	Not so frequent	−0.22 (CI: −0.29, −0.14)***	0.13 (CI: −0.23, 0.49)
	Rather frequent	0.33 (CI: 0.26, 0.40)***	0.22 (CI: −0.14, 0.59)
	Very frequent	0.57 (CI: 0.49, 0.65)***	0.57 (CI: 0.18, 0.95)**
Public expenditure	Little public expenditures per patient	−0.52 (CI: −0.57, −0.47)	−0.38 (CI: −0.60, −0.16)
	Much public expenditures per patient	0.52 (CI: 0.48, 0.56)***	0.38 (CI: 0.20, 0.56)***

CI, confidence interval.
[a]Results of a multinomial logistic regression model.
**$P < 0.01$. ***$P < 0.001$.

frequent and very frequent disease were positive, meaning that a higher prevalence contributed to a higher perceived societal need (Table 5).

Added Value of New Treatments. The added value of new treatments was considered to be influenced most by changes in quality of life and prevalence (weights of 0.37 and 0.36, respectively, in the general population model; Table 3). For decision makers, the relative importance of a reduction in prevalence as compared with an improvement in quality of life was lower (weight of 0.29 for impact on prevalence compared with 0.39 for impact on quality of life). The results also showed that the weight for changes in the quality of life was almost 2.5 times higher than the weight for changes in life expectancy and more than 5 times higher than the weight for impact on public expenditures and inconvenience of treatment.

The coefficients of the model for the general population are presented in Table 6.

Table 6 Added Value: Model Summary for the General Population and Decision Maker Sample

Attribute	Level	General Population (Estimated Coefficient[a] With Confidence Interval and Significance Level)	Decision Makers (Estimated Coefficient[a] With Confidence Interval and Significance Level)
Impact on public expenditure	Increases public expenditure	−0.37 (CI: −0.40, −0.33)	−0.50 (CI: −0.73, −0.26)
	Does not change public expenditure	0.07 (CI: 0.03, 0.10)***	0.12 (CI: −0.08, 0.31)
	Reduces public expenditure	0.3 (CI: 0.26, 0.34)***	0.38 (CI: 0.14, 0.62)**
Change in quality of life	Reduction	−0.83 (CI: −0.87, −0.78)	−1,02 (CI: −1.31, −0.73)
	No change	−0.006 (CI: −0.04, 0.03)	−0,11 (CI: −0.3, 0.08)
	Improvement	0.83 (CI: 0.79, 0.87)***	1.13 (CI: 0.88, 1.38)***
Change in life expectancy	Does not change	−0.41 (CI: −0.43, −0.38)	−0.64 (CI: −0.83, −0.46)
	Increase	0.41 (CI: 0.38, 0.43)***	0.64 (CI: 0.48, 0.80)***
Treatment inconvenience	More	−0.35 (CI: −0.39, −0.32)	−0.29 (CI: −0.46, −0.11)
	As much	0.03 (CI: −0.007, 0.067)	0.08 (CI: −0.13, 0.29)
	Less	0.32 (CI: 0.29, 0.36)***	0.21 (CI: 0.01, 0.40)*
Change in prevalence	Cures fewer	−0.89 (CI: −0.94, −0.83)	−0.92 (CI: −1.22, −0.61)
	Cures an equal number	0.082 (CI: 0.05, 0.12)***	−0.07 (CI: −0.27, 0.12)
	Cures more	0.80 (CI: 0.76, 0.84)***	0.99 (CI: 0.75, 1.23)***

CI, confidence interval.

[a]Results of a multinomial logistic regression model.

*$P < 0.05$. $P < 0.01$. $P < 0.001$.

Subgroup Analysis

Respondents in the citizens' group aged 80 to 89 years had different preferences than the other age groups (Figure 3). For therapeutic need, respondents between 80 and 89 years of age gave much more importance to the criterion of inconvenience of current treatment and less to the criterion of quality of life under current treatment than the other age groups. For societal need, the 80 to 89 year olds gave a higher weight to prevalence than to public expenditures, unlike all other age groups. As for the judgment of the added value of new treatments, the 80 to 89 year olds gave relatively more weight to improvements in quality of life than the other age groups. At the same time, improvements in treatment inconvenience were more important than changes in life expectancy for this group as well as for the 70 to 79 years old. This means that these age groups valued living better more than living longer, whether "better life" was defined by better quality of life or less treatment inconvenience. In contrast, the other age groups typically gave more weight to improvements in life expectancy than to reductions in inconvenience, but they also gave more weight to improvements in quality of life than to increases in life expectancy. The respondents in the youngest age group (20–29 years) gave relatively more weight to reductions in public expenditures compared with the other age groups, although this criterion also for this age group remained the least important for the assessment of the added value of a new intervention.

People who reported being currently in good health gave slightly more weight to quality of life when judging therapeutic need than to inconvenience of current treatment. Respondents who reported not being in good health found it more important to reduce treatment inconvenience than to increase overall quality of life. Both subgroups gave the lowest weight to reductions in life expectancy due to the disease. A full report of all subgroup analyses performed on the data is publicly available in Cleemput et al.[20]

Discussion

The aim of the current study was to determine the importance of a set of criteria for the appraisal of therapeutic need, societal need, and added value of new treatments, from the perspective of the general public in Belgium.

The results of the study indicate that the general public gives the highest weight to the impact of a disease on quality of life when assessing therapeutic need. As only few studies are available that study public preferences like in this study, a comparison with existing literature is difficult. In general, medical need and health benefits of treatment are described as the two most important criteria among all priority setting criteria.[29–37] Medical need refers to the severity of a disease if untreated. Therapeutic need, as defined in this study, has received relatively little attention in the empirical literature.

However, we argue that therapeutic need is a more relevant criterion than the absolute concept of medical need, as in many cases a somewhat effective treatment is already available.

In our study, disease-related public expenditures were found more important than prevalence of disease when assessing societal need. Like in previous studies, it was also found that reimbursement of treatment for more prevalent diseases has higher societal need than that of less prevalent diseases. Literature and the results of this study do not support the claim that "rarity of the disease" is an important separate criterion.[34,37–40] Finally, the results of this study indicate that the impact of treatment on quality of life is most important when considering added value. In contrast, previous studies found that impact of treatment on life expectancy is a more important criterion than impact on the quality of life.[39,41,42]

However, older respondents attached higher importance to inconvenience of current treatment than to its impact on quality of life when judging the therapeutic need. An intriguing observation is that people seem to dichotomize between "lethal" and "nonlethal" diseases when confronted with a DCE for judging therapeutic need. Therapeutic need was considered higher in diseases leading to premature death as compared with diseases not resulting in premature death, but no distinction was made between diseases leading to immediate death and diseases leading to a reduction of life expectancy of 5 years.

The impact of nonresponse to the survey is hard to predict. We found a statistical difference for age but not for gender between nonresponders and responders. The oldest age groups (>60 years) was slightly overrepresented in the nonresponders group. In our subgroup analysis, we found differential effects for the group of respondents aged 70 or older; hence, a slight bias in the weights might be expected. However, since no other information is available on these nonresponders, it is impossible to quantify if and how our results would be different if these non-responders would have participated.

The current study is unique in several aspects. First and foremost, to our knowledge, no such large-scale public preference study was conducted before.[23] Moreover, the public preferences were obtained using discrete choice experiments rather than with more traditional surveying techniques like, for instance, Likert-type scales. DCE requires people to think about and compare hypothetical scenarios. Literature suggests that a stated preference technique like a DCE is an appropriate way to obtain relative preference values for different criteria, if the assumption is made that people actually have a utility function and hence can make a choice.[37] Compared with, for example, a rating and ranking exercise, DCE has the advantage to let all attributes be weighted in the decision at once, and actual tradeoffs between criteria need to be made by the respondents. By including a pretest, a pilot test, and a test-retest phase in the survey development process, comprehension, presentation, and feasibility of the questions was substantially improved.

Methodologically, our approach with level-independent attribute weights has not been used before, although it has been suggested as a theoretical possibility by Lancsar and colleagues.[23] How to derive attribute weights from part-worth utility data is actually a big gap in the scientific literature. We applied an approach based on the log-likelihood because we considered this approach methodologically sound and intuitively appealing. However, more research on the robustness of the approach is needed.

While a DCE does not presume preference independence between attributes, the proposed application of the weights in the MCDA do assume such preference independence. This is a weakness of the proposed MCDA. However, an MCDA should always be complemented with deliberation about the MCDA results, not only to deal with aspects such as possible preference dependencies but also to deal with criteria that are not included in the MCDA but nevertheless considered relevant. To fuel this deliberation, qualitative patient input might be needed. Patient input might provide evidence on the decision criteria included in the MCDA, allows decision makers to make better judgements on the "performance levels" of diseases or interventions, and might highlight possible dependencies between criteria.

Contrary to many examples in literature, the MCDA application developed in this study is different, in that we propose to apply an MCDA to each cluster of criteria, being therapeutic need (disease-related criteria from patients' point of view), societal need (disease-related criteria from the societal point of view), and added value (intervention-related criteria). Most MCDA models described in literature aim at one weighted score covering all relevant clusters at once.[6,43,44] We have several reasons to suggest the use of a multilayered MCDA. First, we presumed that the willingness to reimburse a new treatment out of public resources is a function of the level of therapeutic and societal need. A new treatment with a presumably high added value could still not be worthwhile reimbursing because there simply is no need for a new treatment. In case of a high need and a high added value, decision makers will be more inclined to consider reimbursement than in case of a low need and a

low added value. However, there are several situations in which a conditional decision might be taken. For example, in case of a low therapeutic and societal need and a high added value, the authorities might still want to reimburse a new intervention, under the condition that the overall cost of the treatment is the same as that of the comparator. An economic evaluation can provide this information. Or, when no active alternative treatment is available but the only available alternative is best supportive care, decision makers might still want to reimburse a new promising intervention, with currently a limited added value, to keep the door open for further improvements in the development of the intervention. In such cases, specific conditions for reimbursement will often have to be defined (i.e., who gets reimbursement, under which conditions) and a re-assessment after some time will have to be scheduled.

Second, in a hierarchical decision-making process, the number of criteria to be considered per step in the process diminishes as compared with an all-encompassing one-step decision-making process. This makes the appraisal process more manageable from the cognitive point of view.

The weights derived in this study are independent of the disease or treatment under consideration. If the same criteria weights are used for all diseases and treatments, this will result in more consistent decision making than when varying weights are used over decisions. It was a normative choice to use unlabeled states in the DCE. On the one hand, it might reduce confounding due to consideration of additional implicit criteria, but on the other hand, it reduces the specificity of the preferences. For applicability in real life it is, however, important not to have to repeat such a large exercise for every decision that needs to be made.

While weights indicate to what extent a criterion should be taken into account in the decision-making process, an important next step is the development of scoring rules for the criteria in an MCDA tool.

The criteria scores will vary across diseases and treatments, as they depend on the impact of a disease or the effect of treatment on for instance quality of life and life expectancy. In other words, the clinical significance is reflected in the scoring, and the weights indicate to what extent a clinical significant or insignificant effect should matter for the decision. Appraisal committees should develop a habit in scoring scientific evidence on a particular scale. Guidance on how to deal with missing or low-quality evidence should also be developed.

In future studies, we will envisage the weights derived from this study to be used in the multi-criteria decision framework described in the introduction to include the public perspective in the decision-making process. The application of an MCDA using the weights presented in this study has been pilot tested in Belgium for the appraisal of unmet therapeutic and societal needs.[45] The members of the commission who participated in the pilot study were overall very positive about the experience and decided to apply the methodology in real life. It is important to consider that while the weights derived in this study hold promise to be used in the current appraisal process, other considerations that have not been considered can play a role in the decision process (e.g., impact of a disease on the well-being of the patients' family, considerations of distributive justice). If other criteria are considered important for a decision, they should be made explicit by the appraisal committee. Committees should explain how these additional criteria modify the ranking of a disease or a treatment based on the MCDA. If not, the decision process will remain opaque and it will be unclear whether the preferences of the population eventually really mattered. Helping decision makers to make better informed decision making by providing data on the relative importance of decision criteria according to the general public is the key objective of the current study. The application in MCDA will increase transparency of the decision making process.

Conclusion

In a democratic system, decision makers might wish to take societal preferences into account when making health care reimbursement decisions. Our study showed that the general public gives the highest weight to the impact of a disease on quality of life when assessing therapeutic need, to disease-related public expenditures when assessing societal need, and to impact of a new treatment on quality of life when assessing the added value. The weights presented in this study could be used in a multi-criteria decision approach that could increase the legitimacy of decision-making processes and the acceptance of the decisions. Whether or not this promise will hold true if the results of these study were to be implemented in the Belgian context remains to be studied.

References

1. Clark S, Weale A. Social values in health priority setting: a conceptual framework. *J Health Organ Manag.* 2012;26(3):293–316.
2. Devlin N, Sussex J. *Incorporating Multiple Criteria in HTA. Methods and Processes.* London: Office of Health Economics; 2011.

3. Thokala P, Devlin N, Marsh K, 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.

4. Jaramillo HEC, Goetghebeur M, Moreno-Mattar O. Testing multi-criteria decision analysis for more transparent resource-allocation decision making in Colombia. *Int J Technol Assess Health Care*. 2016;32(4):307–14.

5. Kolasa K, Zwolinski KM, Kalo Z, Hermanowski T. Potential impact of the implementation of multiple-criteria decision analysis (MCDA) on the Polish pricing and reimbursement process of orphan drugs. *Orphanet J Rare Dis*. 2016;11:23.

6. Schey C, Krabbe PF, Postma MJ, Connolly MP. Multi-criteria decision analysis (MCDA): testing a proposed MCDA framework for orphan drugs. *Orphanet J Rare Dis*. 2017;12(1):10.

7. Wagner M, Khoury H, Bennetts L, et al. Appraising the holistic value of lenvatinib for radio-iodine refractory differentiated thyroid cancer: a multi-country study applying pragmatic MCDA. *BMC Cancer*. 2017;17(1):272.

8. le Polain M, Franken M, Koopmanschap M, Cleemput I. *Drug Reimbursement Systems: International Comparison and Policy Recommendations*. Brussels: Belgian Health Care Knowledge Centre; 2010.

9. Cleemput I, Franken M, Koopmanschap M, le Polain M. European drug reimbursement systems' legitimacy: five-country comparison and policy tool. *Int J Technol Assess Health Care*. 2012;28(4):358–66.

10. Health Technology Assessment International. Interest groups. Patient and citizen involvement [cited June 25, 2018]. Available from: https://htai.org/interest-groups/pcig/

11. Silva AS. Involving public and patients in the process of evaluation of health technologies in Brazil [cited June 25, 2018]. *ISPOR News Across: Latin America*. Available from: http://press.ispor.org/LatinAmerica/2017/09/involving-public-and-patients-in-the-process-of-evaluation-of-health-technologies-in-brazil/

12. Bastian H. Speaking up for ourselves. The evolution of consumer advocacy in health care. *Int J Technol Assess Health Care*. 1998;14(1):3–23.

13. Thomas V, Livingstone H, Norburn L, Thomas L, Leng G. England. In: Facey KM, Hansen HP, Single ANV, eds. *Patient Involvement in Health Technology Assessment*. Singapore: Adis, Springer Nature; 2017. p 275–88.

14. Single ANV, Macpherson K, Fearns N, Dickson J, Facey KM. Scotland. In: Facey KM, Hansen HP, Single ANV, eds. *Patient Involvement in Health Technology Assessment*. Singapore: Adis, Springer Nature; 2017. p 321–31.

15. Farmakas A, Theodorou M, Galanis P, et al. Public engagement in setting healthcare priorities: a ranking exercise in Cyprus. *Cost Eff Resour Alloc*. 2017;15:16.

16. Baltussen R, Mitton C, Danis M, Williams I, Gold M. Global developments in priority setting in health. *Int J Health Policy Manag*. 2017;6(3):127–8.

17. Weale A, Kieslich K, Littlejohns P, et al. Introduction: priority setting, equitable access and public involvement in health care. *J Health Organ Manag*. 2016;30(5):736–50.

18. Hunter DJ, Kieslich K, Littlejohns P, et al. Public involvement in health priority setting: future challenges for policy, research and society. *J Health Organ Manag*. 2016;30(5):796–808.

19. McKie J, Richardson J. Social preferences for prioritizing the treatment of severely ill patients: the relevance of severity, expected benefit, past health and lifetime health. *Health Policy*. 2017;121(8):913–22.

20. Cleemput I, Devriese S, Kohn L, et al. *Incorporating Societal Preferences in Reimbursement Decisions—Relative Importance of Decision Criteria According to Belgian Citizens*. KCE Report 234. Health Services Research (HSR). Brussels: Belgian Health Care Knowledge Centre; 2014.

21. de Bekker-Grob EW, Hol L, Donkers B, et al. Labeled versus unlabeled discrete choice experiments in health economics: an application to colorectal cancer screening. *Value Health*. 2010;13(2):315–23.

22. Johnson FR, Lancsar E, Marshall D, et al. Constructing experimental designs for discrete-choice experiments: report of the ISPOR Conjoint Analysis Experimental Design Good Research Practices Task Force. *Value Health*. 2013;16(1):3–13.

23. Lancsar E, Louviere J, Flynn T. Several methods to investigate relative attribute impact in stated preference experiments. *Soc Sci Med*. 2007;64(8):1738–53.

24. Croissant Y. mlogit: Multinomial Logit Model (R package version 0.2-4). Available from: http://CRAN.R-project.org/package=mlogit

25. Hensher DA, Rose JM, Greene WH. *Applied Choice Analysis: A Primer*. Cambridge: Cambridge University Press; 2005.

26. Bech M, Gyrd-Hansen D. Effects coding in discrete choice experiments. *Health Econ*. 2005;14(10):1079–83.

27. Statistics Belgium. ICT usage in households. StatBel. 2018; https://statbel.fgov.be/en/themes/households/ict-usage-households (accessed September 3, 2018).

28. Tafforeau J. Subjectieve gezondheid. In: Van der Heyden J, Charafeddine R, eds. *Gezondheidsenquête 2013 Rapport 1: Gezondheid en Welzijn*. Brussel: WIV-ISP; 2014.

29. Diederich A, Swait J, Wirsik N. Citizen participation in patient prioritization policy decisions: an empirical and experimental study on patients' characteristics. *PLoS One*. 2012;7(5):e36824.

30. Gallego G, Taylor SJ, McNeill P, Brien JA. Public views on priority setting for high cost medications in public hospitals in Australia. *Health Expect*. 2007;10(3):224–35.

31. Green C. Investigating public preferences on "severity of health" as a relevant condition for setting healthcare priorities. *Soc Sci Med*. 2009;68(12):2247–55.

32. Kasemsup V, Schommer JC, Cline RR, Hadsall RS. Citizen's preferences regarding principles to guide health-care allocation decisions in Thailand. *Value Health*. 2008;11(7):1194–202.

33. Lim MK, Bae EY, Choi SE, Lee EK, Lee TJ. Eliciting public preference for health-care resource allocation in South Korea. *Value Health*. 2012;15(1 Suppl.):S91–S94.

34. Linley WG, Hughes DA. Societal views on NICE, cancer drugs fund and value-based pricing criteria for prioritising medicines: a cross-sectional survey of 4118 adults in Great Britain. *Health Econ*. 2013;22(8):948–64.

35. Mortimer D, Segal L. Is the value of a life or life-year saved context specific? Further evidence from a discrete choice experiment. *Cost Eff Resour Alloc*. 2008;6:8.

36. Ubel PA. How stable are people's preferences for giving priority to severely ill patients? *Soc Sci Med*. 1999;49(7): 895–903.

37. Watson V, Carnon A, Ryan M, Cox D. Involving the public in priority setting: a case study using discrete choice experiments. *J Public Health (Oxf)*. 2012;34(2):253–60.

38. Bryan S, Roberts T, Heginbotham C, McCallum A. QALY-maximisation and public preferences: results from a general population survey. *Health Econ*. 2002;11(8): 679–93.

39. 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.

40. Schwappach DL, Strasmann TJ. "Quick and dirty numbers?" The reliability of a stated-preference technique for the measurement of preferences for resource allocation. *J Health Econ*. 2006;25(3):432–48.

41. Lees A, Scott N, Scott SN, MacDonald S, Campbell C. Deciding how NHS money is spent: a survey of general public and medical views. *Health Expect*. 2002;5(1):47–54.

42. Shmueli A. Survival vs. quality of life: a study of the Israeli public priorities in medical care. *Soc Sci Med*. 1999;49(3): 297–302.

43. Marsh K, Dolan P, Kempster J, Lugon M. Prioritizing investments in public health: a multi-criteria decision analysis. *J Public Health (Oxf)*. 2013;35(3):460–6.

44. Marsh K, IJzerman M, Thokala P, et al. Multiple criteria decision analysis for health care decision making—emerging good practices: report 2 of the ISPOR MCDA Emerging Good Practices Task Force. *Value Health*. 2016;19(2):125–37.

45. Cleemput I, Devriese S, Christiaens W, Kohn L. *Multi-Criteria Decision Analysis for the Appraisal of Medical Needs: A Pilot Study*. KCE Report 272. Health Services Research (HSR). Brussels: Belgian Health Care Knowledge Centre; 2016.