Public preferences for prioritizing preventive and curative health care interventions : a discrete choice experiment

Reference:

Luyten Jeroen, Kessels Roselinde, Goos Peter, Beutels Philippe.- Public preferences for prioritizing preventive and curative health care interventions : a discrete choice experiment

Value in health - ISSN 1098-3015 - 18:2(2015), p. 224-233

DOI: http://dx.doi.org/doi:10.1016/j.jval.2014.12.007

Public preferences for prioritizing preventive and curative health care interventions: A discrete choice experiment

Jeroen Luyten[1,2,3,*], Roselinde Kessels[4,5], Peter Goos[4,5,6,7], Philippe Beutels[1,8]

[1]Centre for Health Economics Research & Modelling Infectious Diseases, Vaccine & Infectious Disease Institute, Faculty of Medicine & Health Sciences, University of Antwerp, Antwerpen, Belgium;
[2]Centre for Economics and Ethics, Institute of Philosophy, University of Leuven, Leuven, Belgium;
[3]Department of Social Policy, London School of Economics and Political Science, London, UK;
[4]Faculty of Applied Economics, Department of Economics, University of Antwerp, Antwerpen, Belgium;
[5]StatUa Center for Statistics, University of Antwerp, Antwerpen, Belgium;
[6]Department of Econometrics, Erasmus School of Economics, Erasmus University Rotterdam, Rotterdam, The Netherlands;
[7]Faculty of Bioscience Engineering, Department of Biosystems, University of Leuven, Leuven, Belgium;
[8]School of Public Health and Community Medicine, The University of New South Wales, Sydney, Australia

[*]Corresponding author: jeroen.luyten@uantwerpen.be

Abstract

Background: Setting fair health care priorities counts among the most difficult ethical challenges our societies are facing.

Objective: To elicit through a discrete choice experiment the Belgian adult population's (18-75 years; N = 750) preferences for prioritizing health care and investigate whether these preferences are different for prevention versus cure.

Methods: We used a Bayesian D-efficient design with partial profiles, which enables considering a large number of attributes and interaction effects. We included the following attributes: 1) type of intervention (cure vs. prevention), 2) effectiveness, 3) risk of adverse effects, 4) severity of illness, 5) link between the illness and patient's health-related lifestyle, 6) time span between intervention and effect, and 7) patient's age group.

Results: All attributes were statistically significant contributors to the social value of a health care program, with patient's lifestyle and age being the most influential ones. Interaction effects were found, showing that prevention was preferred to cure for disease in young adults, as well as for severe and lethal disease in people of any age. However, substantial differences were found in the preferences of respondents from different age groups, with different lifestyles and different health states.

Conclusions: Our study suggests that according to the Belgian public, contextual factors of health gains such as patient's age and health-related lifestyle should be considered in priority setting decisions. The studies, however, revealed substantial disagreement in opinion between different population subgroups.

Keywords: distribution, efficiency, equity, prevention, QALY, treatment.

Introduction

One of the greatest challenges for the future consists in finding a fair match between ever-increasing medical needs and possibilities on the one hand and finite health care budgets on the other. Consensus exists that such priority setting should reflect a concern for both efficiency (making maximal use of valuable resources) and equity (avoiding that some people become deprived of their deserved share) [1,2]. Over the past decades, the concern for efficiency has been operationalized in cost-utility analysis, informing decision makers on the ratio between incremental costs and incremental Quality-Adjusted Life Years (QALYs) attributable to interventions [3]. Equity, however, remains a much more elusive concept because a large number of contextual considerations of patients, illnesses, or interventions could justify a more or less favorable weighing in rationing decisions [4]. Therefore, an important research objective remains to clarify which distributive principles carry social support.

The aim of this article is to contribute to the empirical literature describing the general public's distributive preferences regarding health care. We do so by means of a discrete choice experiment (DCE) held in a representative sample of the Belgian adult population. We pay specific attention to the following two issues. First, published studies about preferences for health care resource allocation largely ignored the difference in nature between prevention and cure. Unlike cure, however, prevention 1) avoids the intangible costs of experiencing ill-health; 2) can give rise to substantial externalities, with consequences for both efficiency and equity [5] (e.g., herd immunity through vaccination [6]); 3) is closely related to social justice (e.g., by adjusting social determinants of health) [7]; and 4) is attributed only a small fraction (< 5%) of the health care budget in most countries [8], and may be the first to be cut in times of scarcity. In this study, we pay specific attention to the relative value of either type of health care and investigate whether their nature affects rationing principles. Second, an important criticism against studies eliciting social preferences is that aggregation covers up important heterogeneity in the ethical views of different respondents [7]. We therefore pay much attention to differences in the preferences of relevant subgroups, via the inclusion of several respondent characteristics as covariates in our analysis.

Methods

DCEs are a widely used technique to quantify individuals' preferences by observing their stated choices in a number of hypothetical scenarios [9-11]. Respondents are confronted with a sequence of choice sets consisting of two or more competing options. For each choice set, they have to indicate the option they like best. The options are described in terms of a fixed set of attributes or dimensions that differ in their levels. The data from a DCE allow the assessment of the relative importance of each attribute in the total value attributed to the options under valuation.

DCEs are predominantly used to elicit personal preferences (for a general review of applications, see [12]), but, in a number of studies, they have also been used to explore a population's social and ethical views regarding priority setting in health care (e.g., [13,14]; for specific reviews, see [15-17]). One motivation for using DCEs in the latter context is that respondents are forced to consider the consequences of their choice (choosing one option implies foregoing the other), which avoids that they simply ignore the fact that health care resources are limited.

Conducting a DCE involves the following steps: 1) identification of the attributes and attribute levels, 2) experimental design of the choice sets, 3) survey development, 4) sample selection and survey administration, and 5) data analysis.

Identification of the Attributes and Attribute Levels

For our research objective, it was important to identify a number of decontextualized, generic characteristics that provide a workable description of both preventive and curative interventions. These characteristics should enable respondents to make a meaningful judgment regarding the necessity to reimburse a given intervention. We considered literature review and expert opinion the preferable sources of information. Reviews have classified considerations, potentially relevant for rationing health care programs, into three groups: characteristics belonging to the patient, the intervention and the health condition [15,18]. We updated a review of DCEs about priority setting [17] and identified 12 DCEs exploring the social value of health care [13,14,17,19-27]. We reviewed these studies focusing on the attributes used. We observed that all studies used combinations of attributes to indicate what would happen when a patient would not receive care (severity of illness, expressed in morbidity and/or mortality) and what would happen in case a patient received care

(effectiveness of the intervention/health improvement). In addition, the studies involved a cost or budget impact attribute, the number of patients affected, alternative treatment options, and characteristics of the recipient (mainly age or health-related lifestyle).

The reviewed studies, however, mainly focused on cure, either explicitly or implicitly by shaping a context that is intuitively associated with curing patients, rather than with preventing illness. Therefore, we carried out a separate review of studies aiming to elicit preferences for prevention to find additional attributes. In a review of 114 DCEs [12], we found nine specifically applied to preventive interventions such as screening tests or vaccines [28-36]. These nine studies suggested the inclusion of two additional attributes in our DCE, namely, the intervention's risk on adverse effects and the time span between the intervention and its clinical effect.

In sum, our literature review suggests the following list of nine attributes as most useful to include in our DCE: type of intervention (curative or preventive), effectiveness of the intervention, adverse effects associated with the intervention, severity of illness, cost of the intervention, number of patients, relation to health-related lifestyle, time span between the intervention and the expected effect, and age group of the patient.

Subsequently, we organized group discussions with convenience samples consisting of researchers (N = 10) and lay persons (N = 14) in which we presented interventions in terms of these nine characteristics to investigate whether we overlooked potentially important attributes and whether the descriptions we used for the attributes and their levels allowed a realistic mental image of a health care program. No additional attributes were considered essential. However, when we tested exploratory choice sets, it appeared that inclusion of all nine attributes made the cognitive burden too large for respondents. Respondents not only had to compare the characteristics of the intervention and the disease but also had to consider scale differences between both programs (cost and number of patients). This extra dimension required respondents to make calculations and made them raise questions for clarification. Therefore, we decided to exclude the attributes cost and number of patients by mentioning in every choice set that the interventions had the same cost and were beneficial for the same number of patients.

The next challenge was to refine the wording used to describe the attributes and their levels, and to consider other than verbal presentations of the attribute levels. First, we presented all attributes to our convenience sample in various formulations to determine which one was

easiest to understand. Because the use of attributes representing risks or chances is cognitively demanding, we considered using visualizations for the levels of the attributes effectiveness, risk on adverse effects, and lifestyle instead of verbal descriptions [37]. However, we learnt that a verbal description was most reliable because it minimized the cognitive burden imposed on the respondents while still bringing across the intended meaning. Also, for the other attributes, we experimented by describing levels using numbers and percentages, and found that the choice task was most intuitive when we described levels verbally. Terms such as "rarely" and "often" are more judgmental than numbers and chances (e.g., 1 adverse effect per 100 interventions), and they may translate into different numerical equivalents in different respondents. Using probabilities, however, does not guarantee equal interpretation (e.g., is a chance of 1 per 100 rare or often?). For our purpose, the qualitative rather than quantitative judgment of the respondent was what mattered, and, therefore, we opted for qualitative descriptions for a limited number of attribute levels.

We used three levels for each attribute, except for the attributes type of intervention, which has two levels, and age group of the patient, which has five levels. For the age attribute, we decided against covering all ages because this would make the age groups very wide. Instead, we opted for equally wide age intervals at different stages of life. **Table 1** presents the descriptions of the attributes and their levels used. We presented the attributes one by one to the members of our convenience sample and asked them how they interpreted each attribute and attribute level. We encountered no difficulties in understanding.

Table 1: Attributes and levels.

Attribute	**Level**
What type of intervention is it?	1. Preventive (aiming to prevent healthy persons from becoming ill) 2. Curative (aiming to cure people who are ill)
How big is the probability of success of the intervention?	1. 1 in 3 is successful (33%) 2. 2 in 3 is successful (66%) 3. Always successful (100%)
How often do adverse effects occur?	1. Often 2. Rarely 3. Never
How severe is the illness for which the intervention is developed?	1. Not lethal, but everyone who gets the disease will experience a short period of illness without lasting effects (not severe) 2. Not lethal, but everyone who gets the disease will experience a severe and lasting reduction in quality of life (severe) 3. Lethal, everyone who gets the disease will die from it (lethal)
Does the patient cause the disease through his or her own lifestyle?	1. Fully 2. Partly 3. Not at all

How long does it take before the patient becomes ill/ shows signs/symptoms of illness?	1. After 20 years 2. After 5 years 3. Within a year
At what age does the patient become ill?	1. 80 – 90 years 2. 60 – 70 years 3. 40 – 50 years 4. 20 – 30 years 5. 0 – 10 years

Experimental Design of the Choice Sets

The DCE presented respondents with 14 choice sets of two competing medical interventions, termed "profiles" henceforth. The profiles are combinations of levels of the seven attributes in **Table 1**. To limit the cognitive burden imposed on the respondents, we used "partial profiles" [38-40], that is, we varied the levels of only four of the seven attributes in the choice sets and kept the levels of the other attributes constant. However, different attributes are held constant across choice sets at levels that change between choice sets. We did show the constant attributes to the respondents. This improves the validity of the parameter estimates on the one hand [41], and allows for the estimation of interaction effects on the other hand. To facilitate the choice tasks, we highlighted the varying attributes in each choice set. **Figure 1** shows an example choice set in which respondents had to choose between two interventions A and B.

Figure 1: Example of a choice set.

Medical interventions A and B are exactly equally expensive and they apply to a similar number of patients. If you were forced to make a choice, which of the two interventions should be reimbursed by the government? To make it easier for you, we have highlighted in yellow (gray in this figure) the characteristics that differ between both interventions. There are no right or wrong answers; we are interested in your opinion.

	A	B
What type of intervention is it?	Curative (meant to cure patients who are ill)	Preventive (meant to prevent healthy persons from becoming ill)
How big is the probability of success of the intervention?	2 in 3 is successful	Always successful
How often do adverse effects occur?	Often	Often
How severe is the illness for which the intervention is developed?	Not lethal, but everyone who gets the disease will experience a severe and lasting reduction in quality of life	Lethal, everyone who gets the disease will die from it
Does the patient cause the disease through his or her own lifestyle?	Not at all	Not at all
How long does it take before the patient becomes ill/ shows signs/symptoms of illness?	Within a year	Within a year
At what age does the patient become ill?	0 to 10 years	40 to 50 years
YOUR PREFERENCE:	◘	◘

To maximize the information content of the DCE, we created three different surveys by constructing a partial profile design involving 42 choice sets and dividing it into three groups of 14 choice sets such that each group or survey has a similar partial profile design structure (see below). In constructing the design profiles, we excluded four unrealistic combinations of levels of two attributes (shown in **Appendix A** [which also includes the three surveys]). We ensured that each survey was filled out an equal number of times. As pointed out by Sándor and Wedel [42], using 42 instead of 14 different choice sets results in a larger amount of information on the respondents' preferences and therefore in more precise estimates of the relative importance of the attributes and attribute levels.

Besides the estimation of the main effects of the attributes, we were interested in estimating the interactions between "type of intervention" and any other attribute. However, because of the disallowed level combinations associated with the attribute "time span" (shown in **Appendix A**), the interaction between "type of intervention" and "time span" cannot be estimated. As a discrete choice model, we used a multinomial logit (MNL) model, which is common practice in discrete choice design and analysis [10]. The partial profile design in **Appendix A** is D-efficient or D-optimal for the MNL model, meaning that it guarantees precise estimates of the main effects and the interactions between "type of intervention" and five other attributes [43].

Each choice set of the D-efficient partial profile design varies the levels of four of the seven attributes. These varying attributes differ from choice set to choice set. We determined them using the attribute balance approach that attempts to vary each attribute in an equal number of choice sets and to pair varying attributes an equal number of times [39,40]. That is why each attribute is varied in eight choice sets of each survey of the partial profile design.

The D-efficient partial profile design takes into account prior knowledge concerning the respondents' preferences. For our DCE, for example, it generally holds that the expected priority ranking for reimbursement of interventions is, from low to high, related to a mild disease, followed by a severe, but not lethal disease, and finally, a lethal disease. Similarly, for all other attributes, we took into account expert prior information about the most logical ordering of the levels of the attributes, from low priority to high priority, the result of which is presented in **Table 1**. We also ranked the attributes in order of expected importance and expressed our uncertainty regarding the a priori orderings of the attributes and attribute levels in a multivariate normal prior distribution. In **Appendix B**, we discuss in detail how we

obtained that multivariate normal prior distribution to optimize the design. The design that maximizes the information content of the DCE (as measured by the log-determinant of the information matrix; see [43]), when averaged over that prior distribution, is called a Bayesian D-efficient design. The Bayesian D-efficient design approach is increasingly considered a state-of-the-art approach for DCEs (see, e.g., [43-48]; see also **Appendix C**). Major benefits of Bayesian D-efficient designs are that, using a proper prior distribution, they avoid choice sets in which one profile is completely dominating the other profile(s) on every attribute [49], and (as demonstrated below) such designs can be constructed to efficiently estimate interaction effects.

Survey Development

We provided respondents with a Web link that allowed them to carry out the choice tasks at their earliest convenience. To help respondents, we presented a thorough explanation of the choice tasks at the beginning of the DCE to familiarize them with 1) the context of increasing scarcity in health care, the problem of setting fair priorities, and our objective to investigate how the general population thinks about this difficult ethical policy issue and 2) all seven attributes and their levels and how they are used in the description of a treatment or a preventive intervention. We asked respondents to choose between two interventions of which only one could be reimbursed by the government.

After the DCE, we asked respondents a number of background questions about their age, sex, height, weight, educational attainment, family size, experience as health care worker, smoking status, and experience with severe illness (personal or within the family). To have an estimate of the respondent's current health state, the respondents also had to complete the EuroQol health survey (i.e., the Visual Analogue Scale (VAS) and the generic EQ-5D-5L) [50,51]. These are all variables that we a priori considered to be of potential relevance to someone's health care preferences.

Sample Selection and Survey Administration

In June 2012, a sample was drawn, representative of the Belgian population in terms of age, sex, region, and educational attainment, from an actively managed, continuously updated panel of 10,753 Belgians. Participation was incentivized with credits by means of which, after a number of positively evaluated survey participations, gifts could be bought and a lottery organized on a per-survey basis. Only one respondent per household was allowed. The market

research company guarantees high-quality data through checks and ensures that only "serious" respondents are included. To this end, fraudulent, inattentive, hyperactive, or conditioned respondents were removed from the sample, for example, respondents who complete the survey unreasonably fast ("speedsters"), consistently give the same answer ("straightliners"), and so on.

A total of 30% (N = 3,160) of the 10,753 contacted individuals agreed to participate, of which 937 were selected on the basis of quota requirements. Of the selected individuals, 149 did not finish the survey and 38 did not meet the quality criteria. This left us with a sample of 750 respondents (250 respondents for each of the three versions of the survey), and 10,500 observed choices in total, that is, 14 per respondent. Respondents were distributed proportionally over the three survey versions according to language, sex, and age. **Table 2** compares basic characteristics of the sample to those of the population, showing overall good agreement. Given the societal context of this study, and a lack of clarity about the criteria (additional to the ones described in the previous paragraph) that can identify irrational response data in a DCE without imposing preferences [52,53], all 750 respondents were included in our analysis.

Table 2: Sample characteristics relative to those of the Belgian population.

Characteristic	**Sample (%)**	**Belgian population (%)**
Language		
Dutch	56	56
French	44	44
Sex		
Male (M)	50	50
Female (F)	50	50
Sex per age group *		
18 – 24 M	6	6
18 – 24 F	6	6
25 – 34 M	9	9
25 – 34 F	10	9
35 – 44 M	10	11
35 – 44 F	10	10
45 – 54 M	10	10
45 – 54 F	11	10
55 – 64 M	9	8
55 – 64 F	10	8
65 – 75 M	6	6
65 – 75 F	4	6
Level of education †		
None or lower education	8	19
Lower secondary education	10	20
Higher secondary education	31	33

Higher non-university education	35	18
University education	15	10
Province		
Antwerp	15	16
West Flanders	10	11
East Flanders	13	13
Limburg	8	8
Hainault	13	12
Liege	10	10
Luxemburg	3	2
Namur	5	4
Brussels	10	10
Flemish Brabant	11	10
Walloon Brabant	3	3
Smoking status ‡		
Never smoked	45	54
Ex-smoker	30	22
Smoker	25	25

Source Belgian Data: Federale Overheidsdienst Economie [71].

* Age: The percentages reported are proportions in the selected population (18-75 years), representing 71% of the total Belgian population.

† Education: The percentages reported for the Belgian population are for the age group 15 years or older. The percentages for our sample are only for the age group 18 to 75 years. The overrepresentation of higher educated respondents in our sample as compared to the total population can be explained by our exclusion of the group 15 to 18 years that is too young for higher education, and the age group 75 years or older for which higher education was less democratically accessible.

‡ Smoking percentages from the population are based on the Scientific Institute of Public Health's 2008 Health Survey [72] and are representative of the population aged 15 years or older.

Data Analysis

The data we collected through our DCE allow quantification of and statistical inference about the relative importance of the attributes and attribute levels in assessing the priority ranking of a health intervention. This is done by estimating the respondents' utility function as part of the MNL model, using a maximum likelihood approach. The utility function is represented by the sum of the utilities of the attributes' main and interaction effects under study. The overall significance of the attributes was computed using likelihood ratio (LR) tests, and the relative importance of the attributes was measured by –log (*P* value of the LR test). We started our analysis by estimating the a priori MNL model, that is, the model that seemed most useful when planning the entire study and which was used as a basis for constructing the D-efficient design for the DCE. That model includes the main effects of all attributes and the interactions between type of intervention and five other attributes. Next, we dropped the insignificant model terms until we obtained a final model in which all remaining effects had significant explanatory value. Preference heterogeneity was assessed –in a separate analysis– by adding

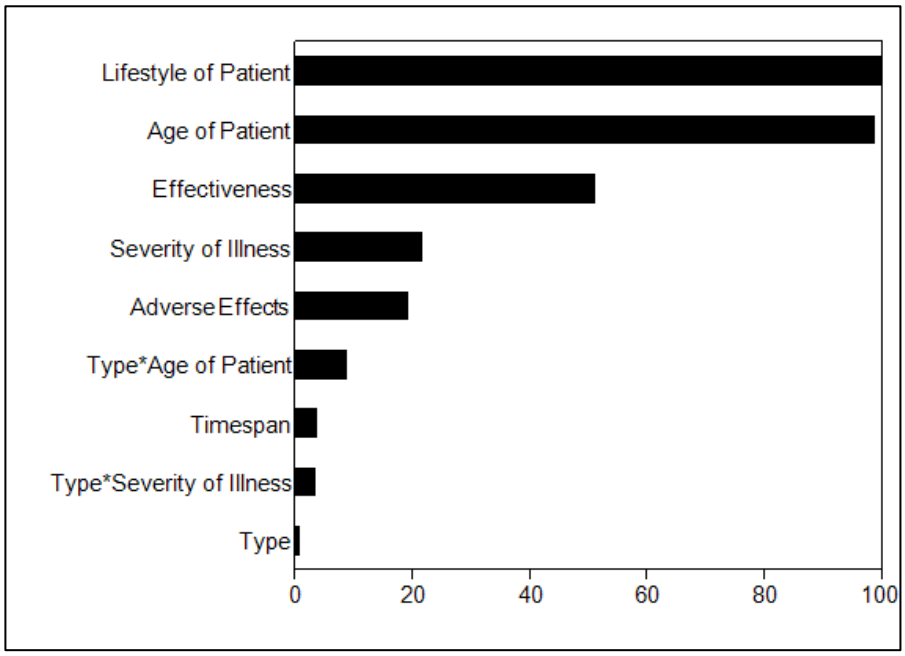
an interaction term in the model with a variable that we a priori determined to be of potential interest (e.g., sex or respondent's age). We carried out the entire data analysis using the Choice Modeling platform in the statistical software package JMP 10.

Results

Main and Interaction Effects

Figure 2 and **Table 3** summarize the results of our analysis. All seven attributes are statistically significant contributors to the social value of a health care intervention, meaning that none of them is considered irrelevant to priority setting. As shown in **Figure 2**, the most influential attributes (based on the LR test) are the patient's health-related lifestyle and age. They are about twice as important as the intervention's effectiveness and about four times as important as severity of illness. Type of intervention is also an important attribute for the model because it appears in two significant interaction effects: between "type of intervention" and "patient's age" and between "type of intervention" and "severity of illness" (see below). Its importance is shown by the LR test for the joint significance of the three effects involving this attribute (LR χ^2 = 26.19; DF = 7; P = 0.0005). Time span is the attribute that is least important.

Figure 2: Importance of the seven attributes (main and interaction effects) to the social value of a health care program relative to the most important attribute "Lifestyle of Patient", the importance of which is set to 100.



The main-effect estimates in **Table 3** represent the utilities attached to the different levels of the attributes. The direction of the coefficients across the levels of each attribute is in line with our a priori expectations. The attractiveness for reimbursement increased as the intervention was more effective or had a lower risk of adverse effects, when the disease was more severe or occurred earlier in time, and when the patient's age and the link between the disease and the patient's lifestyle was lower. **Figure 4** visualizes this for the two most influential attributes: patient's age and link with patient's lifestyle (for the total sample and for specific subgroups, see section Preference Heterogeneity). Looking at the 95% confidence intervals for the utility estimates of each level, we judge that most coefficients are different on a statistically significant level. In a few cases, the level estimates did not significantly differ: the adverse effects levels "never" and "rarely", the patient's age levels "0-10 years" and "20-30 years," and the timing levels "within a year" and "after 5 years."

Table 3: Estimates of coefficients in the MNL model, their 95% CIs, and overall significances of the attributes using *P* values obtained from likelihood ratio tests.

Term	**Estimate**	**95% CI**	***P***
Lifestyle of Patient			
Fully	-0.3742	-0.4324; -0.3163	
Partly	0.0483	0.0046; 0.0921	<0.0001
not at all	*0.3259*[*]	0.2685; 0.3833	
Age of Patient (years)			
80 – 90	-0.6160	-0.7088; -0.5241	
60 – 70	-0.0067	-0.0829; 0.0695	
40 – 50	0.1168	0.0479; 0.1857	<0.0001
20 – 30	0.2206	0.1485; 0.2929	
0 – 10	*0.2853*[*]	0.2012 ; 0.3694	
Effectiveness (%)			
33	-0.2584	-0.3171; -0.2001	
66	0.0282	-0.0177; 0.0741	<0.0001
100	*0.2302*[*]	0.1765; 0.2840	
Severity of Illness			
Not severe	-0.2210	-0.2946; -0.1476	
Severe	0.0653	0.0245; 0.1061	<0.0001
Lethal	*0.1557*[*]	0.0812; 0.2303	
Adverse Effects			
Often	-0.1582	-0.2131; -0.1037	
Rarely	0.0856	0.0426; 0.1287	<0.0001
Never	*0.0726*[*]	0.0235; 0.1221	
Type*Age of Patient (years)			
Preventive*80 – 90	0.0120	-0.0397; 0.0638	
Preventive*60 – 70	-0.0676	-0.1190; -0.0162	
Preventive*40 – 50	-0.0480	-0.1038; 0.0078	0.0005
Preventive*20 – 30	0.1103	0.0537; 0.1669	
Preventive*0 – 10	*-0.0067*[*]	-0.0732; 0.0600	

Curative*80 – 90	*-0.0120*[*]	-0.0638; 0.0397	
Curative*60 – 70	*0.0676*[*]	0.0162; 0.1190	
Curative*40 – 50	*0.0480*[*]	-0.0078; 0.1038	
Curative*20 – 30	*-0.1103*[*]	-0.1669; -0.0537	
Curative*0 – 10	*0.0067*[*]	-0.0600; 0.0732	
Time Span			
After 20 years	-0.0843	-0.1529; -0.0155	0.0404
After 5 years	0.0234	-0.0273; 0.0741	
Within a year	*0.0609*[*]	0.0069; 0.1148	
Type*Severity of Illness			
Preventive*Not severe	-0.0431	-0.0801; -0.0059	0.0487
Preventive*Severe	0.0345	-0.0041; 0.0731	
Preventive*Lethal	*0.0086*[*]	-0.0330; 0.0503	
Curative*Not severe	*0.0431*[*]	0.0059; 0.0801	
Curative*Severe	*-0.0345*[*]	-0.0731; 0.0041	
Curative*Lethal	*-0.0086*[*]	-0.0503; 0.0330	
Type			
Preventive	0.0137	-0.0297; 0.0571	0.5376
Curative	*-0.0137*[*]	-0.0571; 0.0297	

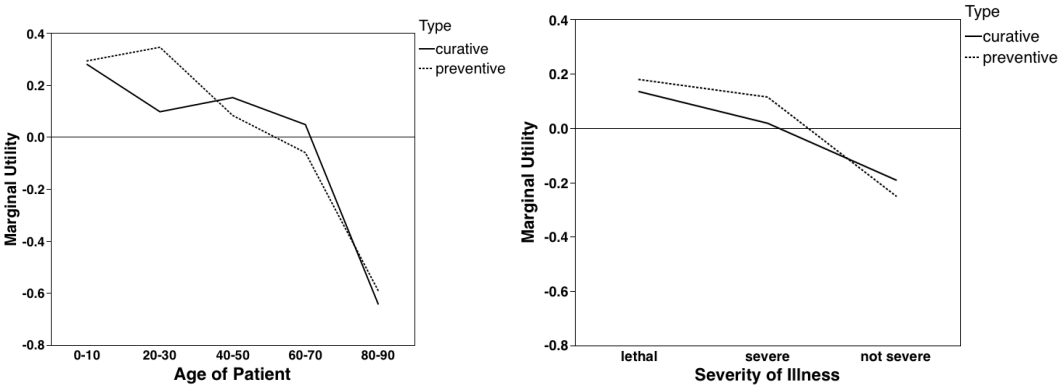
CI, confidence interval; MNL, multinomial logit.

[*]Coefficient estimates corresponding to the last level of an attribute, either as main effect or involved in an interaction, are indicated in italic to stress that they are calculated as minus the sum of the estimates for the other levels of that attribute. To illustrate, the value of 0.0086 for the interaction effect Type[preventive]*Severity of Illness[lethal] is obtained as –(–0.0431 + 0.0345).

In **Figure 3**, the two interactions involving type of intervention are plotted. Our data suggest that the relative value of cure versus prevention differed as a function of the age group of the patient and the severity of the health problem. As shown in **Figure 3A**, however, the interaction with age is mainly due to the upward kink at 20 to 30 years in the value of prevention and the downward kink at the same age category in cure. For patients older than 30 years, the utility of prevention declined, whereas the value of cure remained relatively stable over the age interval from 20 to 70 years. However, we judge that this difference in utility between prevention and cure in older age groups is statistically not significant because from the 95% confidence intervals we observe that the interaction effects "type[preventive]*age of patient[40-50 years]" and "type[preventive]*age of patient[60-70 years]" belong to each other's confidence interval. Regarding the second interaction, **Figure 3B** shows that severity of illness had a larger impact on the utility of a preventive intervention than on that of a curative one. Prevention was valued less than cure in case of a non-severe, transient illness. However, it was valued more for severe, long-lasting, and life-threatening diseases. Our respondents did not consider effectiveness, risk of adverse effects, or lifestyle to be of differential importance in choosing between prevention and cure (i.e., these attributes did not interact with type of intervention).

Figure 3: Marginal utility values for different combinations of "Age of Patient" and "Type of Intervention" (A) and for different combinations of "Severity of Illness" and "Type of Intervention" (B).

A. B.



Preference Heterogeneity

We found many differences in the preferences of various subgroups (indicated by a statistically significant interaction effect). The largest interactions (based on the LR test) were found between respondent's age and patient's age and between respondent's smoking status and patient's lifestyle.

Respondent's age

Younger respondents attributed significantly more importance to patient's age ($P < 0.001$), severity of illness ($P < 0.001$) and patient's lifestyle ($P = 0.0001$). To visualize the impact of the respondent's age, we partitioned the respondents of our data set into three age groups: 18 to 35 years, 36 to 60 years, and 61 to 75 years (see **Figure 4)**. For the youngest respondents, the social value of an intervention depended more strongly on the patient's age (**Figure 4A**, dotted line) and the link between the disease and the patient's lifestyle (**Figure 4B**, dotted line) than for the older respondents.

Sex

The covariate sex appeared in two significant interactions, one involving the attribute patient's age ($P = 0.003$) and one involving the attribute effectiveness ($P = 0.01$). Female respondents attached less value to interventions for older patients than did male respondents. Male respondents attributed a higher value to effectiveness.

Level of education and experience in health care

Respondents with a degree of higher education (university or non-university) were more willing to ration on the basis of age ($P < 0.001$) and attribute more importance to severity of illness ($P < 0.0001$). In addition, they attributed more value to prevention ($P = 0.02$). No differences in opinion were found between higher and lower educated respondents for the attributes lifestyle, effectiveness, and adverse effects. Present or past experience as health care worker (14% of the sample) did not result in significant interactions.

Household

The larger the household of the respondent, the more prevention was preferred to cure ($P = 0.003$) and the more the value of health care decreased as a function of patient's age ($P = 0.002$). Respondents living with children gave more importance to patient's age. No significant interactions were found with other attributes.

Health state

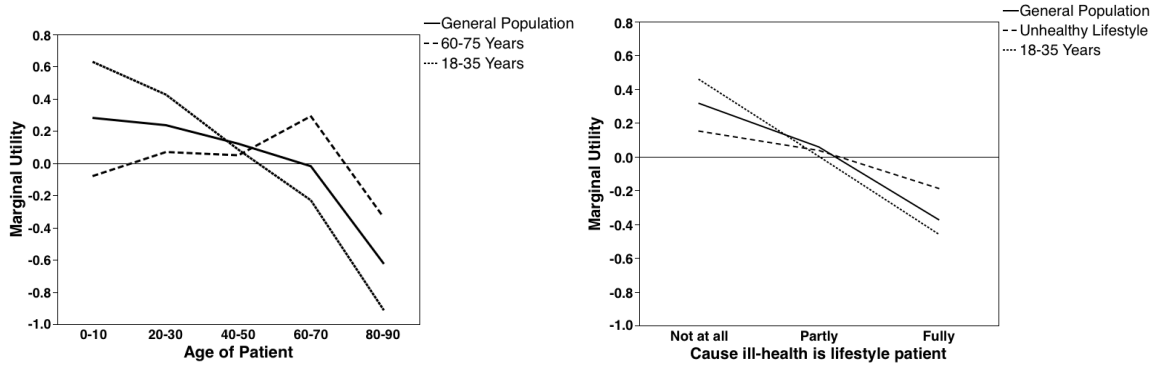
We found significant interactions with the respondent's VAS and EQ-5D-5L scores. The lower the VAS score, the lower the importance of the patient's age group ($P < 0.0001$) and the stronger the preference for cure over prevention ($P < 0.008$). The lower their EQ-5D-5L score, the more likely respondents were to prefer cure over prevention ($P < 0.006$) and to prefer current health gains over future ones ($P < 0.01$). Respondents who reported having had personal experience with severe illness attached greater value to cure than to prevention ($P < 0.002$), and to current rather than to future health gains ($P < 0.01$). They attributed less importance to adverse effects ($P < 0.04$) and were less inclined to take the age of patients into account ($P < 0.006$). We also partitioned our sample into a "good health" group (74% of the sample) and a group with "present or past health problems" (26%), depending on whether they had personal experience with severe illness or had an EQ-5D-5L score below 0.6 or a VAS score below 60. Both groups differ in that the "good health" group preferred prevention ($P < 0.002$), discounted future health gains to a lesser extent ($P = 0.04$), and was more willing to ration on the basis of age ($P = 0.01$). Respondents' choices did not differ significantly on the basis of experience with severe illness in their family.

Lifestyle

Whether or not the respondent is a smoker is a highly influential covariate and, hence, an important factor explaining preference heterogeneity among the respondents. Smokers clearly attributed a lower importance to the attributes lifestyle ($P < 0.0001$) and patient's age ($P = 0.0008$), preferred cure to prevention ($P < 0.0001$), and discounted future health gains to a greater extent ($P = 0.04$). Also, "Body Mass Index" (BMI) turns out to be an important covariate. The higher the BMI score, the less a respondent takes into account the patient's age ($P < 0.0032$), severity of illness ($P = 0.0006$), and the disease's link with lifestyle ($P = 0.04$) and the more he or she prefers cure to prevention ($P = 0.0034$) and discounts future health gains ($P = 0.008$). To visualize the differential valuation by respondents with a "healthy" lifestyle and patients with an "unhealthy" lifestyle, we partitioned the respondents into two groups, one group (in total 38% of the sample) containing respondents with a BMI exceeding 30 (i.e., the obesity threshold [54]) as well as smokers and one group containing non-smokers with a BMI lower than 30. The "unhealthy" lifestyle group preferred cure to prevention ($P < 0.0001$), attributed a lower weight to lifestyle ($P < 0.0001$) and patient's age ($P < 0.0001$) (as illustrated in **Figure 4B**), and, remarkably, attached more importance to the risk of adverse effects ($P = 0.02$).

Figure 4: Marginal utility values for the attributes "Age of Patient" (A) and "Lifestyle of Patient" (B) for the entire sample (general population) and for different respondent subgroups (respondents aged 18-35 years, aged 60-75 years, and a group defined as having an "unhealthy" lifestyle [i.e., being smoker or obese]).

A. B.



Discussion

The objective of our study was to investigate on which basis the Belgian population wants to set health care priorities. Although characteristics of the intervention (effectiveness and risk of adverse effects) and of the illness (severity of illness and time span) were found to matter, it were mainly the characteristics of the recipient that drove respondents' preferences. Priority was given to younger patients and to those who have not somehow caused their own illness. We also detected substantial heterogeneity in the preferences: young, healthy, highly educated or more health-conscious adults responded in a markedly different way than did older, unhealthy, less well educated, and health-unconscious ones.

Our results confirm studies in other countries indicating that the context shapes the social value of QALYs, and that the general public's distributive preferences diverge from a simple health maximization approach, as would be prescribed by cost-utility analysis (i.e., minimizing cost/QALY) (e.g., [13,14,16,17,19,20,24,25,55]; for reviews, see [15,16,56]). Many of these studies also observe a public preference for prioritizing younger patients over older ones, and several ones describe how a substantial number of participants want to account for self-inflicted illness. However, our results seem to diverge from these other studies in the strong impact of the lifestyle attribute, and the relatively limited impact of severity of illness to priority setting.

We paid specific attention to the difference between prevention and cure. A few studies in the literature also compared stated preferences for both types of health care [24,57-62]. These studies found no preference [57,58], a preference for prevention [24,60,61], or a preference for cure [59,62]. Our sample valued prevention higher than cure only when it is targeted at relatively young age groups and when it protects against more severe illness. However, as the self-inflicted nature of a health condition was a factor of major relevance in our study, indirectly, our results can be interpreted as providing further support for prevention in general. An allocation scheme that accounts for individual responsibility would mainly ration on curative treatments because accountability for lifestyle is less relevant for (not) providing prevention, especially when it comes to primary prevention. Preventive programs can incentivize, or even *enable*, citizens to adopt healthy and responsible lifestyles before their lifestyle-associated risk exposure requires cure. Currently, preventive "lifestyle" policies such as alcohol, fat, sugar, or smoking taxes are gaining interest [63,64]. Such measures, if effective, would increase short-term government income and reduce lifestyle-related

morbidity.

Some limitations of our study must be mentioned. First and foremost, priority setting in health care requires societal support [65]. But the majority is not necessarily right [66,67]. We observed that the support for age-based rationing and accounting for lifestyle depends on the age and lifestyle of the respondents themselves, indicating some degree of self-serving answers. Whereas we conveyed to respondents to answer as citizens (a societal view) and not as potential health care consumers (an individual view), and although self-serving answers are not necessarily unjust, this does raise suspicion of partiality. How this can be avoided, and how we can construct a more effective "veil of ignorance" in social preference studies, remains a challenge for further research. But even with impartial answers, age-based rationing and accounting for lifestyle remain controversial grounds for setting priorities (for an elaborate discussion, see [68]). The results of public opinion research in this area should always be complemented by ethical considerations. Therefore, instead of being directly useful to priority-setting decisions, our survey in the first place supports the need for a more extensive public debate about the appropriate role of age and lifestyle in health care rationing.

Second, although our sample was broadly representative of the Belgian population, respondents were recruited from an online panel. This excluded respondents older than 75 years and membership of the panel may be associated with unobservable characteristics. Third, we surveyed our sample on a complex topic, in a single recording. We encouraged respondents to think thoroughly about their answer, and evidence suggests that respondents' answers to DCEs such as ours are reliable [21]. Nonetheless, it would be interesting to repeat this study in a non-panel sample (e.g., generated through random digit dialing of telephone numbers), and to organize a follow-up study in the same sample to compare the results. Fourth, an inherent limitation of DCEs is that only a limited number of attributes can be used. Although we included a relatively large number of attributes, the choice alternatives we presented remain simplified versions of real health care programs. Fifth, because this simplified the choice tasks, we excluded the cost attribute. A disadvantage of this was that it became impossible to quantify willingness to pay for changes in the attributes levels. Such inferences, however, were not our primary objective, and, moreover, they have also been shown to be less informative than expected on some occasions [69,70].

Acknowledgements

Source of financial support: We acknowledge funding from the Research Foundation – Flanders (FWO, project no. G098911N and Roselinde Kessels' postdoctoral fellowship) and Pfizer's European HTAcademy prize competition (2011).

References

[1] Wagstaff A. QALYs and the equity-efficiency trade-off. J Health Econ 1991;10:21-41.
[2] Sen A. Why health equity? Health Econ 2002;11:659-66.
[3] Gold E, Siegel J, Russell L, Weinstein M. Cost-Effectiveness in Health and Medicine. Oxford, UK: Oxford University Press, 1996.
[4] McIntyre D, Mooney G. The Economics of Health Equity. New York: Cambridge University Press, 2007.
[5] Carande-Kulis V, Getzen TE, Thacker SB. Public goods and externalities: a research agenda for public health economics. J Public Health Manag Pract 2007;13:227-32.
[6] Beutels P, Van Doorslaer E, Van Damme P, Hall J. Methodological issues and new developments in the economic evaluation of vaccines. Expert Rev Vaccines 2003;2:649-60.
[7] Powers M, Faden RR. Social Justice: The Moral Foundations of Public Health and Health Policy. Oxford: Oxford University Press, 2008.
[8] Organisation for Economic Co-operation and Development. Health expenditure and financing. 2010. Available from: http://www.oecd.org/health/health-systems/oecdhealthdata.htm. [Accessed May 1, 2013].
[9] Louviere J, Hensher D, Swait J. Stated Choice Methods: Analysis and Applications. Cambridge: Cambridge University Press, 2000.
[10] Ryan M, Gerard K, Amaya-Amaya M. Using Discrete Choice Experiments to Value Health and Health Care. Dordrecht: Springer, 2008.
[11] Bridges JFP, Hauber AB, Marshall D, et al. Conjoint analysis applications in health—a checklist: a report of the ISPOR Good Research Practices for Conjoint Analysis Task Force. Value Health 2011;14:403-13.
[12] de Bekker-Grob EW, Ryan M, Gerard K. Discrete choice experiments in health economics: a review of the literature. Health Econ 2012;21:145-72.
[13] Lancsar E, Wildman J, Donaldson C, et al. Deriving distributional weights for QALYs through discrete choice experiments. J Health Econ 2011;30:466-78.
[14] Norman R, Hall J, Street D, Viney R. Efficiency and equity: A stated preference approach. Health Econ 2012;22:568-81.
[15] Schwappach DL. Resource allocation, social values and the QALY: a review of the debate and empirical evidence. Health Expect 2002;5:210-22.
[16] Dolan P, Shaw R, Tsuchiya A, Williams A. QALY maximisation and people's preferences: a methodological review of the literature. Health Econ 2005;14:197-208.
[17] Green C, Gerard K. Exploring the social value of health-care interventions: a stated preference discrete choice experiment. Health Econ 2009;18:951-76.
[18] Olsen JA, Richardson J, Dolan P, Menzel P. The moral relevance of personal characteristics in setting health care priorities. Soc Sci Med 2003;57:1163-72.
[19] Bryan S, Roberts T, Heginbotham C, McCallum A. QALY-maximisation and public preferences: results from a general population survey. Health Econ 2002;11:679-93.

[20] Schwappach DL. Does it matter who you are or what you gain? An experimental study of preferences for resource allocation. Health Econ 2003;12:255-67.
[21] 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;25:432-48.
[22] Baltussen R, Stolk E, Chisholm D, Aikins M. Towards a multi-criteria approach for priority setting: an application to Ghana. Health Econ 2006;15:689-96.
[23] Tappenden P, Brazier J, Ratcliffe J, Chilcott J. A stated preference binary choice experiment to explore NICE decision making. Pharmacoeconomics 2007;25:685-93.
[24] 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.
[25] Ratcliffe J, Bekker HL, Dolan P, Edlin R. Examining the attitudes and preferences of health care decision-makers in relation to access, equity and cost-effectiveness: a discrete choice experiment. Health Policy 2009;90:45-57.
[26] Mentzakis E, Stefanowska P, Hurley J. A discrete choice experiment investigating preferences for funding drugs used to treat orphan diseases: an exploratory study. Health Econ Policy Law 2011;6:405-33.
[27] Defechereux T, Paolucci F, Mirelman A, et al. Health care priority setting in Norway a multicriteria decision analysis. BMC Health Serv Res 2012;12:39.
[28] Ryan M, Hughes J. Using conjoint analysis to assess women's preferences for miscarriage management. Health Econ 1997;6:261-73.
[29] Hall J, Kenny P, King M, et al. Using stated preference discrete choice modelling to evaluate the introduction of varicella vaccination. Health Econ 2002;11:457-65.
[30] Salkeld G, Solomon M, Short L, et al. Evidence-based consumer choice: a case study in colorectal cancer screening. Aust N Z J Public Health 2003;27:449-55.
[31] Bishop AJ, Marteau TM, Armstrong D, et al. Women and health care professionals' preferences for Down's Syndrome screening tests: a conjoint analysis study. BJOG 2004;111:775-9.
[32] Ryan M, Diack J, Watson V, Smith N. Rapid prenatal diagnostic testing for Down syndrome only or longer wait for full karyotype: the views of pregnant women. Prenat Diagn 2005;25:1206-11.
[33] Wordsworth S, Ryan M, Skatun D, Waugh N. Women's preferences for cervical cancer screening: a study using a discrete choice experiment. Int J Technol Assess Health Care 2006;22:344-50.
[34] Hall J, Fiebig DG, King MT, et al. What influences participation in genetic carrier testing? Results from a discrete choice experiment. J Health Econ 2006;25:520-37.
[35] Bishai D, Brice R, Girod I, et al. Conjoint analysis of French and German parents' willingness to pay for meningococcal vaccine. Pharmacoeconomics 2007;25:143-54.
[36] Lancsar EJ, Hall JP, King M, et al. Using discrete choice experiments to investigate subject preferences for preventive asthma medication. Respirology 2007;12:127-36.
[37] Ancker JS, Senathirajah Y, Kukafka R, Starren JB. Design features of graphs in health risk communication: a systematic review. J Am Med Inform Assoc 2006;13:608-18.
[38] Green P. On the design of choice experiments involving multifactor alternatives. J Consum Res 1974;1:61-8.
[39] Kessels R, Jones B, Goos P. Bayesian optimal designs for discrete choice experiments with partial profiles. J Choice Modelling 2011;4:52-74.
[40] Kessels R, Jones B, Goos P. An improved two-stage variance balance approach for constructing partial profile designs for discrete choice experiments, Appl Stochastic Models Bus Ind 2015. http://dx.doi.org/10.1002/asmb.2065.

[41] Dellaert BGC, Donkers B, van Soest A. Complexity effects in choice experiment-based models. J Marketing Res 2012;49:424-34.
[42] Sándor Z, Wedel M. Heterogeneous conjoint choice designs. J Marketing Res 2005;42:210-8.
[43] Kessels R, Jones B, Goos P, Vandebroek M. The usefulness of Bayesian optimal designs for discrete choice experiments. Appl Stochastic Models Bus Ind 2011;27:173-88.
[44] Kessels R, Goos P, Vandebroek M. A comparison of criteria to design efficient choice experiments. J Marketing Res 2006;43:409-19.
[45] Kessels R, Jones B, Goos P, Vandebroek M. Recommendations on the use of Bayesian optimal designs for choice experiments. Qual Reliability Eng Int 2008;24:737-44.
[46] Rose JM, Bliemer MCJ. Constructing efficient stated choice experimental designs. Transport Rev 2009;29:587-617.
[47] Reed Johnson F, 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;16:3-13.
[48] Kessels R, Jones B, Goos P, Vandebroek M. Rejoinder: the usefulness of Bayesian optimal designs for discrete choice experiments. Appl Stochastic Models Bus 2011;27:197-203.
[49] Crabbe M, Vandebroek M. Using appropriate prior information to eliminate choice sets with a dominant alternative from D-efficient designs. J Choice Modelling 2012;5:22-45.
[50] EuroQol Group. EuroQol–a new facility for the measurement of health-related quality of life. Health Policy 1990;16:199-208.
[51] Herdman M, Gudex C, Lloyd A, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). Qual Life Res 2011;20:1727-36.
[52] Lancsar E, Louviere J. Deleting 'irrational' responses from discrete choice experiments: a case of investigating or imposing preferences? Health Econ 2006;15:797-811.
[53] Ryan M, Watson V, Entwistle V. Rationalising the 'irrational': a think aloud study of discrete choice experiment responses. Health Econ 2009;18:321-36.
[54] Aasheim ET, Sovik TT. Global trends in body-mass index. Lancet 2011;377:1916-7; author reply 1917-18.
[55] Dolan P, Cookson R. A qualitative study of the extent to which health gain matters when choosing between groups of patients. Health Policy 2000;51:19-30.
[56] Shah K. Severity of illness and priority setting in healthcare: a review of the literature. Health Policy 2009;93:77-84.
[57] Ubel PA, Spranca MD, Dekay ML, et al. Public preferences for prevention versus cure: what if an ounce of prevention is worth only an ounce of cure? Med Decis Making 1998;18:141-8.
[58] Corso P. Prevention just in case or treatment just because: measuring societal preferences. Harvard Health Policy Rev 2006;7:32-41.
[59] Schwappach DL. The equivalence of numbers: the social value of avoiding health decline: an experimental Web-based study. BMC Med Inform Decis Mak 2002;2:3.
[60] Bosworth R, Cameron TA, DeShazo JR. Is an ounce of prevention worth a pound of cure? Comparing demand for public prevention and treatment policies. Med Decis Mak 2010;30:E40-56.
[61] Johannesson M, Johansson PO. A note on prevention versus cure. Health Policy 1997;41:181-7.
[62] Corso PS, Hammitt JK, Graham JD, et al. Assessing preferences for prevention versus treatment using willingness to pay. Med Decis Making 2002;22:S92-101.

[63] Dharmasena S, Capps O Jr. Intended and unintended consequences of a proposed national tax on sugar-sweetened beverages to combat the U.S. obesity problem. Health Econ 2012;21:669-94.
[64] Cook PJ, Durrance CP. The virtuous tax: lifesaving and crime-prevention effects of the 1991 federal alcohol-tax increase. J Health Econ 2013;32:261-7.
[65] Claxton K, Culyer AJ. Not a NICE fallacy: a reply to Dr Quigley. J Med Ethics 2008;34:598-601.
[66] Richardson J. The poverty of ethical analyses in economics and the unwarranted disregard of evidence. In: Murray C, ed., Summary Measures of Population Health. Geneva, Switzerland: World Health Organization, 1997.
[67] Hausman DM. Polling and public policy. Kennedy Inst Ethics J 2004;14:241-7.
[68] Denier Y, Gastmans C, Vandevelde A. Justice, luck and responsibility in health care: philosophical background and ethical implications for end-of-life care. Dordrecht: Springer, 2013.
[69] Slothuus Skjoldborg U, Gyrd-Hansen D. Conjoint analysis. The cost variable: an Achilles'heel? Health Econ 2003;12:479-91.
[70] Johnson FR, Mohamed AF, Ozdemir S, et al. How does cost matter in health-care discrete-choice experiments? Health Econ 2011;20:323-30.
[71] Federale Overheidsdienst Economie. Belgische bevolking: feiten en cijfers. 2012. Available from: http://statbel.fgov.be/nl/modules/publications/statistiques/bevolking/cijfers_bevolking_1_1_2009.jsp. [Accessed May 1, 2013].
[72] Scientific Institute of Public Health. Health Survey Belgium 2008. 2008. Available from: http://www.vvsg.be/sociaal_beleid/gezondheidsbeleid/Documents/SamenvattingGezondheidstoestand.pdf. [Accessed May, 1, 2013].

Appendix A: Bayesian D-efficient partial profile design consisting of three surveys for the discrete choice experiment.

The design of the discrete choice experiment involved three surveys of 14 choice sets with two alternative medical interventions. The surveys appear in **Tables A.1**, **A.2** and **A.3**. The choice sets in each survey were presented in a randomized order to the respondents. Each survey was filled out by 250 respondents. The choice sets are described by four attributes of which the levels are varied and three attributes of which the levels are kept constant. The levels of the varying attributes are indicated in yellow (as in **Figure 1**). The constant attributes are shown to the respondents to present actual alternative interventions as well as to be able to estimate interactions. In each survey, each attribute is held constant in six choice sets and is varied in eight choice sets.

When constructing the Bayesian D-efficient partial profile design, we excluded the following four combinations of attribute levels, because we believe they are unrealistic:

i. type of intervention [curative] & time span [after 20 years],
ii. type of intervention [curative] & time span [after 5 years],
iii. lifestyle of patient [fully] & age of patient [0 – 10 years],
iv. lifestyle of patient [partly] & age of patient [0 – 10 years].

Table A.1: Survey 1 of the Bayesian D-efficient partial profile design.

Choice set	Type of intervention	Effectiveness	Adverse effects	Severity of illness	Lifestyle of patient	Time span	Age of patient
1	preventive	always successful	never	lethal	fully	after 5 years	40 – 50 years
1	preventive	always successful	never	severe	not at all	within a year	60 – 70 years
2	preventive	2 in 3 is successful	often	not severe	fully	after 5 years	20 – 30 years
2	preventive	2 in 3 is successful	rarely	severe	fully	after 20 years	60 – 70 years
3	preventive	2 in 3 is successful	rarely	severe	partly	within a year	40 – 50 years
3	preventive	1 in 3 is successful	rarely	severe	not at all	after 20 years	0 – 10 years
4	curative	1 in 3 is successful	rarely	not severe	not at all	within a year	0 – 10 years
4	curative	2 in 3 is successful	often	not severe	partly	within a year	40 – 50 years
5	preventive	always successful	never	not severe	not at all	after 20 years	40 – 50 years
5	preventive	1 in 3 is successful	often	severe	not at all	after 5 years	40 – 50 years
6	curative	always successful	rarely	not severe	partly	within a year	80 – 90 years
6	curative	1 in 3 is successful	often	lethal	not at all	within a year	80 – 90 years
7	curative	1 in 3 is successful	never	not severe	partly	within a year	60 – 70 years
7	preventive	1 in 3 is successful	never	severe	not at all	within a year	40 – 50 years
8	curative	1 in 3 is successful	rarely	lethal	partly	within a year	20 – 30 years
8	preventive	1 in 3 is successful	never	lethal	not at all	after 5 years	20 – 30 years
9	preventive	2 in 3 is successful	never	lethal	partly	after 5 years	60 – 70 years
9	curative	2 in 3 is successful	often	lethal	not at all	within a year	60 – 70 years
10	preventive	always successful	rarely	severe	fully	within a year	20 – 30 years
10	curative	always successful	often	lethal	fully	within a year	60 – 70 years
11	curative	1 in 3 is successful	rarely	severe	not at all	within a year	80 – 90 years
11	preventive	2 in 3 is successful	rarely	severe	not at all	after 5 years	20 – 30 years
12	preventive	2 in 3 is successful	often	lethal	not at all	after 5 years	0 – 10 years
12	curative	always successful	often	not severe	not at all	within a year	0 – 10 years
13	curative	1 in 3 is successful	often	severe	partly	within a year	60 – 70 years
13	preventive	always successful	often	lethal	fully	within a year	60 – 70 years
14	curative	2 in 3 is successful	rarely	not severe	not at all	within a year	80 – 90 years
14	preventive	always successful	often	not severe	not at all	within a year	0 – 10 years

Table A.2: Survey 2 of the Bayesian D-efficient partial profile design.

Choice set	Type of intervention	Effectiveness	Adverse effects	Severity of illness	Lifestyle of patient	Time span	Age of patient
1	preventive	2 in 3 is successful	often	lethal	not at all	within a year	80 – 90 years
1	preventive	2 in 3 is successful	rarely	lethal	fully	after 5 years	60 – 70 years
2	curative	2 in 3 is successful	rarely	lethal	partly	within a year	40 – 50 years
2	curative	2 in 3 is successful	never	severe	fully	within a year	20 – 30 years
3	preventive	2 in 3 is successful	often	not severe	not at all	after 5 years	0 – 10 years
3	preventive	1 in 3 is successful	often	lethal	not at all	after 20 years	20 – 30 years
4	preventive	2 in 3 is successful	rarely	lethal	partly	after 20 years	80 – 90 years
4	preventive	always successful	rarely	severe	fully	after 5 years	80 – 90 years
5	preventive	2 in 3 is successful	often	not severe	partly	after 20 years	20 – 30 years
5	preventive	always successful	rarely	not severe	partly	after 5 years	80 – 90 years
6	curative	always successful	rarely	severe	not at all	within a year	80 – 90 years
6	curative	2 in 3 is successful	never	lethal	fully	within a year	80 – 90 years
7	preventive	2 in 3 is successful	rarely	not severe	not at all	after 5 years	80 – 90 years
7	curative	2 in 3 is successful	rarely	not severe	fully	within a year	60 – 70 years
8	preventive	always successful	never	not severe	not at all	after 5 years	40 – 50 years
8	curative	always successful	never	severe	fully	within a year	40 – 50 years
9	preventive	1 in 3 is successful	rarely	lethal	not at all	within a year	0 – 10 years
9	curative	1 in 3 is successful	never	severe	not at all	within a year	40 – 50 years
10	preventive	always successful	rarely	lethal	partly	after 20 years	20 – 30 years
10	curative	always successful	never	not severe	partly	within a year	20 – 30 years
11	curative	always successful	rarely	lethal	fully	within a year	80 – 90 years
11	preventive	2 in 3 is successful	rarely	lethal	not at all	within a year	40 – 50 years
12	curative	2 in 3 is successful	often	severe	not at all	within a year	0 – 10 years
12	preventive	always successful	often	lethal	not at all	within a year	40 – 50 years
13	preventive	always successful	often	lethal	partly	after 5 years	80 – 90 years
13	curative	1 in 3 is successful	never	lethal	partly	within a year	80 – 90 years
14	preventive	1 in 3 is successful	never	severe	partly	within a year	80 – 90 years
14	curative	always successful	often	severe	fully	within a year	80 – 90 years

Table A.3: Survey 3 of the Bayesian D-efficient partial profile design.

Choice set	Type of intervention	Effectiveness	Adverse effects	Severity of illness	Lifestyle of patient	Time span	Age of patient
1	preventive	1 in 3 is successful	often	severe	partly	after 20 years	80 – 90 years
1	preventive	1 in 3 is successful	often	not severe	fully	within a year	60 – 70 years
2	curative	2 in 3 is successful	often	not severe	not at all	within a year	20 – 30 years
2	curative	2 in 3 is successful	never	severe	partly	within a year	60 – 70 years
3	preventive	2 in 3 is successful	never	not severe	fully	within a year	80 – 90 years
3	preventive	1 in 3 is successful	never	not severe	partly	after 5 years	40 – 50 years
4	preventive	always successful	often	lethal	not at all	after 5 years	60 – 70 years
4	preventive	1 in 3 is successful	never	lethal	not at all	within a year	0 – 10 years
5	preventive	always successful	never	severe	fully	after 20 years	20 – 30 years
5	preventive	1 in 3 is successful	rarely	not severe	fully	within a year	20 – 30 years
6	curative	always successful	never	not severe	not at all	within a year	60 – 70 years
6	curative	2 in 3 is successful	rarely	severe	partly	within a year	60 – 70 years
7	curative	2 in 3 is successful	never	not severe	not at all	within a year	0 – 10 years
7	preventive	2 in 3 is successful	never	severe	not at all	after 5 years	20 – 30 years
8	preventive	1 in 3 is successful	never	not severe	partly	after 5 years	60 – 70 years
8	curative	1 in 3 is successful	rarely	not severe	fully	within a year	60 – 70 years
9	curative	1 in 3 is successful	often	not severe	partly	within a year	80 – 90 years
9	preventive	1 in 3 is successful	rarely	not severe	not at all	after 20 years	80 – 90 years
10	preventive	2 in 3 is successful	rarely	severe	partly	within a year	40 – 50 years
10	curative	2 in 3 is successful	never	not severe	partly	within a year	20 – 30 years
11	curative	1 in 3 is successful	rarely	not severe	fully	within a year	40 – 50 years
11	preventive	always successful	rarely	not severe	partly	within a year	20 – 30 years
12	curative	2 in 3 is successful	never	not severe	not at all	within a year	80 – 90 years
12	preventive	1 in 3 is successful	never	lethal	fully	within a year	80 – 90 years

13	preventive	2 in 3 is successful	often	severe	fully	after 20 years	40 – 50 years
13	curative	1 in 3 is successful	often	not severe	fully	within a year	40 – 50 years
14	preventive	always successful	often	severe	not at all	within a year	0 – 10 years
14	curative	2 in 3 is successful	rarely	severe	not at all	within a year	20 – 30 years

Appendix B: Multivariate normal prior parameter distribution used to construct the Bayesian D-efficient partial profile design for the discrete choice experiment.

To construct the Bayesian D-efficient partial profile design for the discrete choice experiment shown in **Appendix A**, we used a multivariate normal prior distribution that reflects the prior beliefs about the unknown parameter values associated with the levels of the seven attributes. Based on expert interviews and literature review, we ranked the seven attributes in order of importance and specified mean parameter values and variances for the multivariate normal prior distribution accordingly.

Table B.1 shows the seven attributes in order of importance, as suggested by a literature review and a panel of experts prior to our study. Based on these orders of importance, we specified prior mean utility values for the main effects of the seven attributes. The more important an attribute, the larger in magnitude the a priori mean utility values specified for the main effects of that attribute. We adopted the same ordering of the levels of the attributes as shown in Table 1, where they are ranked from least preferred to most preferred. Note that the utility values associated with the levels of each attribute sum to zero. This is because we used effects-type coding for the attribute levels, which means that the levels of the 2-level attribute, "type of intervention", are coded as 1 and -1, the levels of every 3-level attribute as [1 0], [0 1] and [-1 -1], and the levels of the 5-level attribute, "age of patient" as [1 0 0 0], [0 1 0 0], [0 0 1 0], [0 0 0 1] and [-1 -1 -1 -1]. The a priori mean utility values for each attribute are symmetric around zero, except for the attributes "time span" and "adverse effects". For these two attributes, the literature and the experts suggested that the prior utility values for the last two levels lie closer to each other than to the utility value for the first level.

Table B.1: A priori order of importance of the main effects of the seven attributes and conversion into mean utility values used in the multivariate normal prior distribution.

Rank	**Attribute**	**Prior mean**				
		Level 1	Level 2	Level 3	Level 4	Level 5
1	Severity of illness	-0.8	0	0.8		
2	Age of patient	-0.5	-0.25	0	0.25	0.5
	Lifestyle of patient	-0.5	0	0.5		
	Effectiveness	-0.5	0	0.5		
	Time span	-0.5	0.2	0.3		
3	Adverse effects	-0.4	0.1	0.3		
	Type of intervention	-0.4	0.4			

For the prior variances around the mean utility values for the main effects of the attributes, we used a value of 0.09 for all levels of all attributes. This way, we ensured that the multivariate normal prior distribution preserved the preference ordering for the levels of any given

attribute as much as possible. Following a suggestion of Kessels et al. [1], we specified negative covariances of -0.045 for the 3-level attributes and of -0.0225 for the 5-level attribute.

Regarding the interaction effects between "type of intervention" and any other attribute except for "time span", we had no prior information about people's preferences. Therefore, we specified zero mean utility values for these effects. For ease of computation, we also assumed zero prior variances around the utility values for the interaction effects, allowing for no uncertainty around these values. This implies that the prior parameter specification of the interaction effects corresponds to a local instead of a Bayesian approach.

Appendix C: A note on robust and inaccurate priors

As indicated by Huber and Zwerina [2] and Carlsson and Martinsson [3], a design based on an inaccurate prior parameter can perform worse than a design based on a flat prior parameter. The authors, however, make this point for locally D-optimal designs, which are more vulnerable to misspecification of the prior distribution than Bayesian D-optimal designs, which explicitly account for uncertainty about the parameter values. However, the design we generated is not a locally D-optimal design, but a Bayesian D-optimal design that assumes a prior distribution of likely parameter values and optimizes the design over that distribution. Hence, our optimal design is robust to parameter uncertainty, which means that it will also be efficient in case the prior means are misspecified to a certain extent. Sándor and Wedel [4], who originated this type of choice design, demonstrated that Bayesian optimal designs are generally more efficient than locally optimal designs. By the specification of the prior variance-covariance matrix, the prior relative attribute ranking may deviate to some extent from the true one without having much effect on the design efficiency. What should be avoided, however, is that the prior relative importance of the attribute levels is consistently wrongly specified. In our discrete choice study, the natural ordering of the attribute levels specified proved to be the estimated ordering, so that the efficiency of the design is not degraded.

References

[1] Kessels R, Jones B, Goos P, Vandebroek M. Recommendations on the use of Bayesian optimal designs for choice experiments. Qual Reliability Eng Int 2008;24:737-44.
[2] Huber J, Zwerina K. The importance of utility balance inefficient choice designs. J Marketing Res 1996;33:307-17.
[3] Carlsson F, Martinsson P. Design techniques for stated preference methods in health economics. Health Econ 2003;12:281-94.
[4] Sándor Z, Wedel M. Heterogeneous conjoint choice designs. J Marketing Res 2005;42:210-8.