PharmacoEconomics

Measuring public preferences for health outcomes and expenditures in a context of healthcare resources re-allocation

--Manuscript Draft--

Manuscript Number:	PECA-D-18-00185R3
Full Title:	Measuring public preferences for health outcomes and expenditures in a context of healthcare resources re-allocation
Article Type:	Original Research Article
Funding Information:	
Abstract:	Backgound: The final outcome of any resource allocation decision in healthcare cannot be determined in advance. Thus decision makers, in deciding which new programme to implement (or not), need to accommodate the uncertainty of different potential outcomes (i.e., change in both health and costs) that can occur, the size and nature (i.e., “bad” or “good”) of these outcomes and how they are being valued. Using the decision making plane (DMP), which explicitly incorporates opportunity costs and relaxes the assumptions of perfect divisibility and constant returns to scale of the cost-effectiveness plane, all the potential outcomes of each resource allocation decision can be described. Objective: In this study we describe the development and testing of an instrument, using a discrete choice experiment methodology, allowing the measurement of public preferences for potential outcomes falling in different quadrants of the DMP. Method: In a sample of 200 participants providing 4,200 observations we compared four versions of the preferences-elicitation instrument using a range of indicators. Results: We identified one version that was well accepted by the participants and with good measurement properties. Conclusion:This validated instrument can now be used in a larger representative sample to study the preferences of the public for potential outcomes stemming from re-allocation of healthcare resources.
Corresponding Author:	Nicolas Krucien, PhD University of Aberdeen Aberdeen, UNITED KINGDOM
Corresponding Author Secondary Information:	
Corresponding Author's Institution:	University of Aberdeen
Corresponding Author's Secondary Institution:	
First Author:	Nicolas Krucien, PhD
First Author Secondary Information:	
Order of Authors:	Nicolas Krucien, PhD
	Amiram Gafni, PhD
	Nathalie Pelletier-Fleury, PhD
Order of Authors Secondary Information:	
Author Comments:	Dear editorial office, we have made the last editorial change requested by the journal co-editor.
Response to Reviewers:	Dear editorial office, we have made the last editorial change requested by the journal co-editor.
Suggested Reviewers:	

Aberdeen, UK
13/11/2018

Dear Tim Wrightson,
Co-Editor of PharmacoEconomics journal,

As requested we have added to the final version of the questionnaire (as online supplementary material) in the main text (i.e., Discussion section; 1st paragraph).

Sincerely,

Dr Nicolas KRUCIEN (PhD)
Health Economics Research Unit
Institute of Applied Health Sciences
University of Aberdeen
Aberdeen, AB25 2QN
Tel: +44(0)1-224-437-892

FULL TITLE

Measuring public preferences for health outcomes and expenditures in a context of healthcare resources re-allocation

SHORT TITLE

Public preferences for health outcomes and expenditures

AUTHORS

Nicolas KRUCIEN, PhD

Health Economics Research Unit, Institute of Applied Health Sciences, University of Aberdeen, Aberdeen (AB25 2QN), United Kingdom; nicolas.krucien@abdn.ac.uk

Nathalie PELLETIER-FLEURY, PhD, MD

Centre de Recherche en Epidémiologie et Santé des Populations, Université Paris-Sud, UVSQ, INSERM, Université Paris-Saclay, Villejuif, France ; nathalie.pelletier-fleury@inserm.fr

Amiram GAFNI, PhD

Centre for Health Economics and Policy Analysis, Department of Health Research Methods, Evaluation and Impact, McMaster University, Hamilton, Canada; gafni@mcmaster.ca

CORRESPONDING AUTHOR

Nicolas KRUCIEN, PhD
Health Economics Research Unit
Institute of Applied Health Sciences
University of Aberdeen
Aberdeen, AB25 2QN
Tel: +44(0)1-224-437-892
Fax: +44(0)1-224-437-195
Email: nicolas.krucien@abdn.ac.uk

ACKNOWLEDGMENTS

We thank all participants who took part in the study. We would like to thank the two anonymous reviewers for their comments which helped us to improve the quality of this article.

Compliance with ethical standards

Financial support for this study was provided by the French National Institute of Health and Medical Research (INSERM). The funding agreement ensured the authors' independence in designing the study, interpreting the data, writing, and publishing the report. The authors (Nicolas KRUCIEN; Nathalie PELLETIER-FLEURY; Amiram GAFNI) have no conflict of interest to declare.

ABSTRACT

Backgound: The final outcome of any resource allocation decision in healthcare cannot be determined in advance. Thus decision makers, in deciding which new programme to implement (or not), need to accommodate the uncertainty of different potential outcomes (i.e., change in both health and costs) that can occur, the size and nature (i.e., "bad" or "good") of these outcomes and how they are being valued. Using the decision making plane (DMP), which explicitly incorporates opportunity costs and relaxes the assumptions of perfect divisibility and constant returns to scale of the cost-effectiveness plane, all the potential outcomes of each resource allocation decision can be described.

Objective: In this study we describe the development and testing of an instrument, using a discrete choice experiment methodology, allowing the measurement of public preferences for potential outcomes falling in different quadrants of the DMP.

Method: In a sample of 200 participants providing 4,200 observations we compared four versions of the preferences-elicitation instrument using a range of indicators.

Results: We identified one version that was well accepted by the participants and with good measurement properties.

Conclusion:This validated instrument can now be used in a larger representative sample to study the preferences of the public for potential outcomes stemming from re-allocation of healthcare resources.

Key points for decision makers

1. A validated preferences-elicitation instrument was developed that can be used in a representative sample of the general population.
2. This study found preliminary evidence of non-linearities in public valuation of outcomes stemming from reallocation of healthcare resources.
3. This study describes a comprehensive methodology for testing and comparing the properties of preferences-elicitation instruments, which can be applied to other discrete choice experiments.

1: INTRODUCTION

If healthcare budgets were unlimited, all most effective treatments could be adopted. However, resources allocated to healthcare are scarce, hence health policy decision-makers (HPDMs) need to decide how to best allocate them. Resources scarcity can occur in different contexts (i.e., fixed budgets, shrinking budget with less resources allocated to healthcare, growing budget with more resources allocated to healthcare) as long as the total amount of resources available is not sufficient to support the implementation of all the most effective treatments. As a result of scarcity, HPDMs need to determine where the resources should come from to fund the implementation of new treatments to replace or complement existing treatments. For example, in the case of a fixed budget, HPDMs may decide to cancel existing treatment(s) in order to free up resources to implement the new treatment(s). In the case of a growing budget, because not all new treatments can be implemented, HPDMs would still need to decide which new treatment(s) to implement and which ones to abandon.

This resources allocation decision (RAD) is challenging because it typically requires to trade off potential health gains for patients who will benefit from the new treatment against potential health losses for those who will see their current treatment being cancelled or replaced (or potential new treatment not adopted). In this context, HPDMs need to consider the opportunity costs of their decisions to "*ensure that the value of what is gained from an activity* [e.g., implementing the new treatment] *outweighs the value of what has to be sacrificed* [e.g., cancelling an existing treatment]"[1].

Cost–Effectiveness Analysis (CEA) is widely advocated as a tool to help HPDMs to allocate the resources available in a way that maximizes the health benefits produced to the population. The analytical tool of CEA is the incremental cost-effectiveness ratio (ICER), which is then compared with a threshold ICER to determine whether the new treatment should be implemented. Assuming that healthcare resources are efficiently used, this threshold ICER should in principle correspond to the ICER of the last treatment adopted [2] and would then correspond to the shadow price of the budget constraint. However, it has been shown that this approach would lead to an optimal use of healthcare resources only under the strong assumptions of perfect divisibility and constant returns to scale of all treatments [3]–[5]. Birch & Gafni (B&G) have suggested an alternative approach relaxing these two questionable assumptions [2],[3]. Other studies also questioned the validity of the "*ICER of the last treatment adopted*" as threshold to guide the RAD [6]. Eckerman & Pekarsky showed that the shadow price is impacted by different factors such as type of financing (i.e., expansion of healthcare budget vs. displacement of existing resources) and whether existing resources were optimally allocated or not. For instance, when funding is done by expansion of the healthcare budget in an economically efficient system, the shadow price should correspond to the ICER of the "best" (i.e., most cost-effective) existing programme. When the funding is done by displacement of existing resources in an economically inefficient system, the shadow price should depend on the ICERs of the "best" programme, of the "worst" (i.e., least cost-effective) programme and of the displaced programme [6]. The B&G approach is based on the key concept of opportunity cost. The differences are: (i) It does not require the use of underlying unrealistic assumptions and their consequences [3], and (ii) it does not require the use of an ICER and a ICER threshold. The B&G approach identifies the source of the additional resource requirements of the new program and makes recommendation regarding the adoption of the new program based on a direct comparison of the total additional benefits produced from the new program with the total benefits forgone. In doing so it ensures, that if followed and under conditions of certainty, "*the value of what is gained from an activity outweighs the value of what has to be sacrificed"* [1].

In the past two decades it has been recognized that both costs and effects of all programs are stochastic, and then the B&G approach has been extended to account for the uncertainty in costs and effects of re-allocating resources [7], [8]. Visually, it takes the form of a Decision Making Plane (DMP) allowing to describe all the possible outcomes stemming from resource reallocation due to the uncertainty [7]. The cost-effectiveness plane (CEP) describes only the difference in health outcomes (E) and costs (C) of a candidate treatment for implementation (A1) with a reference one (A0) using measures of incremental effectiveness ($\Delta E_A = E_{A1} - E_{A0}$) and increment

costs ($\Delta C_A = C_{A1} - C_{A0}$). Those measures are used to compute the ICER and compare it to the ICER threshold. The DMP "extends" the CEP by also comparing a candidate treatment(s) for cancellation (B1) with another reference treatment(s) (B0)[1] (i.e., the explicit consideration of the source of additional resources), leading thus to another set of incremental effects ($\Delta E_B = E_{B1} - E_{B0}$) and costs ($\Delta C_B = C_{B1} - C_{B0}$). All these incremental measures are used to compute net changes in health outcomes ($\Delta E = \Delta E_A - \Delta E_B$) and costs ($\Delta C = \Delta C_A - \Delta C_B$) which are then mapped into the DMP (**Figure 1A**). The DMP is divided into four quadrants which will affect the RAD. Quadrant I (QI) describes situations where the joint decision to replace A0 by A1 and B1 by B0 allows improving the population health (i.e., $\Delta E > 0$) for an overall lower level of medical expenditures (i.e., $\Delta C < 0$). At the opposite, quadrant III (Q_{III}) describes situations where population health is decreased (i.e., $\Delta E < 0$) and medical expenditures increased (i.e., $\Delta C > 0$). Quadrant II (Q_{II}) describes situations where both the population health and level of medical expenditures are decreased (i.e., $\Delta E < 0$; $\Delta C < 0$). Quadrant IV (Q_{IV}) describes situations where both the population health and level of medical expenditures are increased (i.e., $\Delta E > 0$; $\Delta C > 0$).

In terms of health policy decision-making, the decision to replace existing treatments in order to free up resources for the implementation of a new treatment should be made, ideally, only if the final outcome will be located in QI. However, this cannot be guaranteed because, as explained, net changes in population health and medical expenditures are uncertain, such that RAD becomes a risky decision. Each proposed way of allocating healthcare resources might have a non-null probability to end up in each of the four quadrants of the DMP. This uncertainty can be represented by a joint distribution of net changes over the DMP (**Figure 1B**). Given this element of risk, the RAD will depend not only on the probabilities of falling in the four DMP quadrants, but also on how HPDMs value each possible situations. It would be too restrictive to assume that HPDMs view all potential situations as being equally desirable [8]. The valuation of each (ΔE; ΔC) situation is likely to depend on the specific quadrant that it falls in and the exact location within the quadrant. Assuming, for example, that HPDMs positively value an improvement in population health and a decrease in the level of medical expenditures, a situation falling in Q_{III} should be perceived as "bad" (i.e., to have a negative value), and likewise a Q_I situation should be seen as "good" (i.e., to have a positive value). But this description of HPDMs' preferences for changes in population health and medical expenditures remains largely incomplete. It is unknown whether: (i) HPDMs would be more concerned by a "bad" situation rather than a comparable (same-size) "good" situation; (ii) all situations falling in Q_I (Q_{III}) should be seen as equally "good" ("bad"). Also (ΔE; ΔC) situations can fall in quadrants II and IV where one outcome is "good" and the other is "bad". The answers to these questions depend on both the sign and size of HPDMs' preferences for net changes in population health and medical expenditures. To the best of our knowledge such valuation function, that describes the preferences (or value attributed) for every potential outcome in each quadrant of the DMP, does not exist. In order to measure such function one first needs to develop and validate a tool which will provide reliable measures of preferences for net changes in population health and medical expenditures. This is the objective of this study. In this study we report the development and testing of a preference-elicitation instrument (PEI) which can be used to measure preferences for changes in population health and medical expenditures in a context of resources scarcity.

2: METHODS

2.1: Developing a preference elicitation instrument (PEI)

2.1.1: Choice experiment

The discrete choice experiment (DCE) methodology was used to measure preferences for net changes in health outcomes and costs. DCEs are commonly used in health for eliciting preferences for a wide range of policy

[1] The DMP can also be extended to the case where more than one existing treatments have to be replaced in order to free up resources for the implement of the new treatment.

questions [9], [10]. We use the concept of healthy year equivalent (HYE)[2] [11] to describe net changes in health outcomes (ΔE) and amount of Euros for net changes in the medical expenditures (ΔC). In our study, we used a modified DCE format known as best-worst scaling (BWS) case III. Unlike the standard DCE approach which only asks participants to identified their most preferred choice option (i.e., BEST choice), the BWS approach also asks them to identify their least preferred option (i.e., WORST choice). This approach allows for a full rank ordering of the situations and then provide more information about individuals' preferences for the same number of choice tasks [12].

Whilst there are studies in the DCE literature explaining how to identify relevant attributes, to the best of our knowledge there is no comparable evidence regarding the selection of attributes' levels. Furthermore, methodological research on the designing of DCEs showed that individuals' preferences were not invariant to changes in the range of attributes' levels [13]. Unfortunately this issue has been overlooked in the DCE literature and there is no validated approach to identify the "best set" of attributes' levels. In our study we addressed this issue by testing and comparing four different versions of the BW-DCE questionnaire that only differ in terms of attributes' levels: The ranges of possible values for the ΔE attribute were {-4; -2; 0; +2; +4} and {-8; -4; 0; +4; +8} for versions 1 & 2 ($V_{1\text{-}2}$) and versions 3 & 4 ($V_{3\text{-}4}$) respectively; The ranges of possible values for the ΔC attribute were {-60,000; -30,000; 0; +30,000; +60,000} and {-120,000; -60,000; 0; +60,000; +120,000} for V_1 and $V_{2\text{-}4}$ respectively. All four versions were based on a D-Efficient design [14], [15] allowing for the estimation of all main effects and one continuous interaction effect between ΔE and ΔC. However, we used non-informative (i.e., null) priors about participants' preferences to generate the list of choice tasks for $V_{1\text{-}3}$ and used results from V_3 as informative (i.e., non-null) priors to design the tasks for V_4. This last version also included 12 experimental tasks (instead of 10) to allow for the estimation of two alternative-specific constants in addition to the other effects[3].

In every choice task we included three generic situations (i.e., scenario 1; scenario 2; scenario 3) (**Figure 2**) to reflect the uncertainty in the consequences of reallocating healthcare resources. In the information sheet of the questionnaire, the participants were told that deciding to reallocate resources would have uncertain consequences which are represented by the different scenarios. One of these generic situations was designed to correspond to the origin point of the DMP, hereafter *neutral changes situation* (NCS). The specification of the BW-DCE was completed by manually designing two quality checks. For the estimation of individuals' preferences, only answers to the experimentally designed choice tasks were considered. The order of the choice tasks within the questionnaire and the order of the alternatives within the choice tasks were randomised across participants to control for potential order effects (e.g., left-to-right, learning/fatigue).

2.1.2: Sampling

The development of our PEI took place in France. In 2016, we contacted a market research company to recruit a total of 200 participants from the general population. V_4 of the instrument was administered two months after the first three versions[4]. Participants were randomly allocated to V_1-V_3. We used the same recruitment procedure for

[2] Those who prefer to use quality-adjusted life years (QALYs) as a measure of health outcome can use the methodology described in this paper but will need to change the description of the health outcome in the instrument.

[3] We used the same experimental design for V1-3 because we specified null preferences for the ΔE and ΔC attributes, making thus the D-efficiency measure insensitive to changes in the magnitude only of the attributes' levels. The purpose of V4 was to investigate whether a "better" (i.e., statistically more efficient) design would allow building a better PEI. The gain in statistical efficiency was obtained by relaxing the assumption of null preferences for ΔE and ΔC, using V3 as non-null priors for the designing of V4.

[4] The V4 was administered two months after the three other versions because we first needed to analyse data obtained from V3 before being able to improve the statistical efficiency of the V4 design (by using V3 results as non-null priors).

all four versions. Following Louviere et al (2010) [16] formulae[5], we needed to recruit a minimum of 44 participants per version (rounded up at 50).

2.2: Empirical testing of the preferences-elicitation instrument

As there is no validated approach regarding the selection of the "best set" of attributes' levels, we used different criteria to compare the four DCE versions in terms of statistical performance (predictive validity), behavioural realism (properties of participants' choices) and acceptability by the participants. This multi-criteria analysis should increase our chance to identify the best PEI.

2.2.1: Debriefing questions

The participants were asked to rate the interest in and difficulty of the questionnaire on a 5-points scale. Then we asked them to answer questions related to how they made their choices: (i) Decision objective (i.e., random choice; decision to minimise ΔC; decision to maximise ΔE; to find a compromise between ΔE and ΔC); (ii) Minimum acceptable level of ΔE, and maximum acceptable level of ΔC; (iii) Importance of ΔE and ΔC. Differences between the four choice experiments were investigated using Chi-2 tests.

2.2.2: Properties of participants' choices

In addition to the experimental choice tasks, we also included two additional tasks to control for the quality of participants' choices. One task was used to check the monotonicity of participants' choices. In the monotonicity task, one option was the best and one was the worst in terms of both ΔE and ΔC. Participants' were expected to choose the best option as "most preferred" and the worst one as "least preferred". The second quality check consisted was a stability task. We tested the stability of choices by repeating task #2 as 2nd last task. Participants were expected to pass the stability test when at least one of their choices was repeated.

We also defined two other quality measures based on serial non-participation (i.e., participants who systematically select either the left, middle or right-located option) and response time (RT). A participant was classified as serial non-participant when s/he selected the NCS situation as BEST in more than 80% of the choice tasks. We recorded the RTs at the task level to identify "speedsters" (i.e., participants who tended to answer the choice tasks "too quickly"). A "quick decision" was defined as a choice with a RT falling in the 1st quintile of the corresponding RT distribution. A participant was considered as a "speedster" when s/he made *quick* decisions in at least 80% of the choice tasks. We compared the proportions of participants who pass/fail the quality checks using Chi-2 tests.

2.2.3: Behavioural realism

Multi-attribute choices are typically analysed using the random utility maximisation (RUM) framework [17], [18].

The "base case" model can be written:

$$U_{ntj} = \beta_1 \Delta E_{ntj} + \beta_2 \Delta C_{ntj} + \varepsilon_{ntj} \quad \textbf{(Eq. 1)}$$

Where U_{ntj} corresponds to the utility (U) derived by respondent (n) at task (t) from the option (j), ε_{ntj} are modelling errors, and the (β_1, β_2) parameters capture the main effect of a 1-unit change in ΔE attribute (i.e., +1 HYE) and in ΔC (i.e., +10,000 euros) on participants' choices respectively. We expect to find positive preferences for ΔE, meaning that on average participants would positively respond to better health outcomes, and negative preferences for ΔC, meaning that participants would negatively respond to increase in level of medical expenditures.

However, as suggested by Gafni et al., there is *a priori* no guarantee that this reference specification provides the best account of participants' decisions [8]. For each version of the choice experiment, we estimated 32 different

[5] The formulae is for choice proportions and it allows testing whether observed proportions significantly differ from proportions that would be obtained by chance (In our case, 33% as there is three choice options per task): H_0: proportion = 33%; H_1: proportion ≠ 33%.

specifications allowing for more flexible choice behaviours: (i) Non-linear preferences (in ΔE and/or ΔC); (ii) Interaction effect between ΔE and ΔC; (iii) NCS bias; (iv) Choices inconsistency. As the number of parameters differs across the models, we used the Bayesian Information Criterion (BIC) to identify the best performing specification for each version (i.e., the specification associated with the smallest BIC value).

The more sophisticated specification of the choice model would take the following form:

$$U_{ntj} = \exp(\beta_1 TYPE_{ntj})[\beta_2 NCS_{ntj} + [\beta_3 \Delta E\{\text{Max loss}\} + \beta_4 \Delta E\{\text{Min loss}\} + \beta_5 \Delta E\{\text{Min gain}\} + \beta_6 \Delta E\{\text{Max gain}\}] + [\beta_7 \Delta C\{\text{Max loss}\} + \beta_8 \Delta C\{\text{Min loss}\} + \beta_9 \Delta C\{\text{Min gain}\} + \beta_{10} \Delta C\{\text{Max gain}\}] + \beta_{11}(\Delta E_{njt} \times \Delta C_{ntj})] + \varepsilon_{ntj} \quad \textbf{(Eq. 2)}$$

Where (β_1) captures an effect of the type of choices (i.e., BEST vs. WORST) on the errors variance, (β_2) a preference for the NCS above and beyond the preferences for ΔE and ΔC, (β_3-β_6) are parameters capturing the preferences for the different ΔE values relative to a null change, (β_7-β_{10}) are similar parameters for the ΔC attribute, and finally (β_{11}) is a parameter capturing an interaction effect between ΔE and ΔC.

In terms of *behavioural realism* of the choice model, the best version is expected to be the one associated with: (i) Non-linear preferences for at least one attribute; (ii) Non-significant bias towards NCS; (iii) Significant interaction effect between preferences for ΔE and ΔC ([19], [20], [21]).

2.2.4: Predictive validity

The relative performance of choice models across the four versions of the choice experiment can be compared in terms of ability to predict individuals' choices. We use a cross-validation (CV) procedure to determine the level of predictive validity of each model on its corresponding version. The CV procedure consists in randomly splitting the sample into two groups, namely an estimation sample and a validation sample. The observations from the estimation sample are used to estimate the choice model, and the estimates are then used to predict choices observed in validation sample. The predictive validity corresponds to the % of correct matches between predicted and observed choices. Because sample sizes are limited, we proceed to a 75%-25% repartition of the respondents between the estimation and validation samples respectively. The CV procedure was repeated 10,000 times to compute mean score of predictive validity and associated 95% confidence interval.

3: RESULTS

3.1: Samples of respondents

The proportion of men was approximately 50% in all four samples (44%-50%). About a quarter of the respondents reported a less than good health status (20%-30%) and a third declared at least one chronic condition (28%-44%). The samples mainly included respondents with a higher level of education (either University or college degree) (62%-84%). In overall the differences in sample characteristics across the four versions don't reach significance, thus suggesting there is no sample selection bias (i.e., some profiles of participants are not significantly more represented in one group compared to the others).

3.2: Debriefing questions

Results are reported in **Table 1**. The descriptive analysis of debriefing questions indicates that overall the participants considered the questionnaire as being interesting (≈ 70%) and easy to answer (≈ 50%). A majority of participants were willing to trade net changes in health outcomes (ΔE) against net changes in medical expenditures (ΔC). V_4 appeared to be more interesting (+ 20 points) (P = 0.004) than the other ones. This version also has an impact on the individuals' perception of the ΔE and ΔC attributes. In this version, 48% of the participants declare not being willing to accept a net change in health outcomes below +8 HYE (which also corresponds to the maximum value presented to the participants). However we don't find a similar effect for net changes in medical

expenditures with only 16% of the participants not being willing to accept an option offering a level of change below the maximum value.

3.3: Properties of participants' choices

Results are presented in **Table 2**. Regarding the monotonicity of preferences, all four versions of the choice experiment were associated with high levels of performance. However, V_4 appeared to perform better than the other ones (+10-16 points) but this difference did not reach significance (P = 0.147). V4 outperformed the other versions in terms of stability (P < 0.001); it achieved better performance in terms of serial non-participation and response time (+6-8 points) but the difference did not reach significance (P = 0.283 and P = 0.580, respectively).

Regarding the results of the RTs analysis, we found a similar pattern of RTs across the four versions of the choice experiment. The 1st choice task (task #1) is associated with significantly longer RTs, and then RTs tend to slightly decrease over the sequence of tasks. Although most of RT differences between the four versions did not reach significance, V_4 appeared to be systematically associated with longer RT at every task. This last result might indicate that participants might have been more engaged in the completion of the choice tasks [22].

3.4: Behavioural realism

Results are presented in **Table 3**[6]. Regarding V_1 and V_3, the best fitting choice model appeared to be a model allowing for linear preferences for both ΔE and ΔC. At the opposite final model for V_2 and V_4 allowed for non-linearities in preferences for ΔE and/or ΔC. In V_1 and V_3, there was evidence of a NCS bias. In all versions but V_4 we found a significant interaction effect between preferences for ΔE and ΔC. Overall the results verify our *a priori* assumptions regarding the nature of respondents' preferences for ΔE and ΔC attributes (i.e., positive effect of gains; negative effect of losses; monotonic preferences for changes in ΔE/ΔC).

3.5: Predictive validity

Results are presented in **Figure 3**. With a level of predictive validity close to 78%, the version V_4 appeared to perform significantly better than the other ones. This high level of predictive validity indicates that most of participants made choices that can be well explained by the RUM hypothesis, providing thus evidence that participants were actually making trade-offs between the ΔE and ΔC attributes.

4: DISCUSSION

The objective of this study was to develop and test a preferences-elicitation instrument (PEI) that can then be used in large representative sample of the general population to identify the outcome valuation function needed to help the decision makers to decide whether a given distribution of potential situations is acceptable or not, which will affect the decision to reallocate (or not) resources. After having compared four different versions of the PEI in terms of data quality, behavioural realism and predictive validity, version 4 (V_4) was identified as best. To the best of our knowledge, it is the first time that such an instrument is being developed. A copy of the instrument is provided as online supplementary material.

We find that public preferences were sensitive both to the range of values for medical expenditures and health outcomes, and to the design of the PEI. The two versions based on "extended" sets of attributes' values (i.e., V_3 and V_4) outperformed V_1 which was based on a "narrow" set of values. Increasing the level of statistical efficiency, and thus presumably making the choice tasks more difficult, also had a positive effect on the preferences elicitation. In the DCE literature, it is usually argued that participants would respond to an increase in task difficulty by adopting simplifying decision rules or making more random decisions [23]. In our case making the choice tasks more statistically efficient (and presumably more difficult) was not correlated with an increase in perceived difficulty (as reported by the participants). This result suggests that an increase in the statistical efficiency of the tasks could also have made the choice situations more realistic and more engaging for the

[6] Summary information about all model specifications can be found in online supplementary material.

participants, leading this to better quality data. Overall our study results indicate that *small* changes in the design of the choice tasks can have significant effects on stated preferences. This result is consistent with the hypothesis of ill-defined/malleable preferences [24], [25] following which individuals would not know *a priori* how much they value the different product attributes and therefore would "build their preferences on the fly". Previous studies have investigated the effect of manipulating some experimental features, such as level of statistical efficiency, on respondents' choices [26], [27] and provide mixed evidence.

In the past, studies have already used the DCE methodology to investigate public preferences for the allocation of healthcare resources [28]–[36]. We cannot compare our results to this literature because our study differs in one central feature: the way the resource constraint was described and incorporated. Whilst the previous studies were motivated by the context of resources scarcity (i.e., because resources are limited it becomes important to understand public preferences regarding how these scarce resource should be used), they did not explicitly incorporate the resource constraint in the decision problem (i.e., the question asked). Our study is different because it forces the respondent to deal with the consequences of taking into account the opportunity costs of implementing a new program (e.g., having to make a decision where a programme(s) have to be cancelled to free up resources to implement a new programme and what would be the outcome of such decision). We explained, carefully, to all participants the meaning of the concepts of resources scarcity and opportunity costs and their implications when making a decision about reallocation of resources.

Our study is not exempt from limitations. First, due to sample size limitation we were unable to explore the impact of respondents' characteristics on their preferences for health outcomes and medical expenditures. Erdem & Thompson (2014) used a latent class approach to investigate preferences heterogeneity and found the existence of three different classes of preferences [29]. As this flexible modelling of respondents' choices requires a large amount of data, we plan to repeat this analysis in a bigger sample. Second, the recruitment of the participants was done at two different points in time (i.e., participants to the V_4 were recruited two months after those for V_1-V_3) which might have introduced a potential bias in our comparison. However, this seems unlikely as (i) the recruitment procedure was the same for all four versions, (ii) the main samples socio-demographic characteristics did not significantly differ between the four samples, and (iii) the time lag was relatively short (i.e., only two months). Our instrument measures preferences for two core elements of the decision making process, namely changes in health outcomes and medical expenditures. In real life situations, decision makers are likely to take into account more factors (e.g., profile of the patients who will benefit from the new intervention and those who will lose, etc.). However, a priori we had no guarantee that members of the general population would be willing to make such difficult decisions and then we decided to focus on the core dimensions of resources reallocation. There is no point making the the instrument more complex by including other factors in the decision making process if public already struggles to trade medical expenditures against health outcomes. Building on this work, future studies could further improve the quality of our instrument by including additional factors in the decision making problem. While we had no formal way of controlling whether participants did consider other factors when making their decisions, we tried to prevent/decrease a potential omission bias by explicitly prompting them to only consider the information about health outcomes and medical expenditures.

This validated PEI can now be used in a larger sample of participants to measure their preferences for net changes in health outcomes and medical expenditures. This information about public preferences can then be used in combination with the information described in the DMP to inform the resources allocation decision (RAD). Knowing the joint distribution of net changes in health outcomes and medical expenditures, the preferences information can be used, for example, to compute the expected value of implementing the new treatment. The HPMs may decide to adopt the new intervention only if, for example, the expected value is positive. By following this approach the decision-making process would become more transparent and would allow public to have a say in the management of healthcare system by letting their preferences for population health and medical expenditures influence the RAD.

Data Availability Statement

The datasets generated during and/or analysed during the current study are available from the corresponding author on request.

Author Contributions

All three authors (Nicolas KRUCIEN, Nathalie PELLETIER-FLEURY, Amiram GAFNI) were involved in the designing of the study and the writing up the article. NK was in charge of the data analysis.

5: References

[1] A. Williams, "The economic role of health indicators," in *Measuring the Social Benefits of Medicine*, London: Office of Health Economic: G. Teeling Smith, 1983, pp. 63–67.

[2] M. Weinstein and R. Zeckhauser, "Critical ratios and efficient allocation," *J. Public Econ.*, vol. 2, no. 2, pp. 147–157, Apr. 1973.

[3] A. Gafni and S. Birch, "Incremental cost-effectiveness ratios (ICERs): The silence of the lambda," *Soc. Sci. Med.*, vol. 62, no. 9, pp. 2091–2100, May 2006.

[4] S. Birch and A. Gafni, "Cost effectiveness/utility analyses. Do current decision rules lead us to where we want to be?," *J. Health Econ.*, vol. 11, no. 3, pp. 279–296, Oct. 1992.

[5] A. Gafni and S. Birch, "Guidelines for the adoption of new technologies: a prescription for uncontrolled growth in expenditures and how to avoid the problem," *Can. Med. Assoc. J.*, vol. 148, no. 6, pp. 913–917, 1993.

[6] S. Eckermann and B. Pekarsky, "Can the Real Opportunity Cost Stand Up: Displaced Services, the Straw Man Outside the Room," *PharmacoEconomics*, vol. 32, no. 4, pp. 319–325, Apr. 2014.

[7] P. Sendi, A. Gafni, and S. Birch, "Opportunity costs and uncertainty in the economic evaluation of health care interventions," *Health Econ.*, vol. 11, no. 1, pp. 23–31, Jan. 2002.

[8] A. Gafni, S. Walter, and S. Birch, "UNCERTAINTY AND THE DECISION MAKER: ASSESSING AND MANAGING THE RISK OF UNDESIRABLE OUTCOMES: UNCERTAINTY AND THE DECISION MAKER," *Health Econ.*, vol. 22, no. 11, pp. 1287–1294, Nov. 2013.

[9] E. W. de Bekker-Grob, M. Ryan, and K. Gerard, "Discrete choice experiments in health economics: a review of the literature," *Health Econ.*, vol. 21, no. 2, pp. 145–172, 2012.

[10] M. D. Clark, D. Determann, S. Petrou, D. Moro, and E. W. de Bekker-Grob, "Discrete Choice Experiments in Health Economics: A Review of the Literature," *PharmacoEconomics*, vol. 32, no. 9, pp. 883–902, Sep. 2014.

[11] A. Gafni and S. Birch, "QALYs and HYEs Spotting the differences," *J. Health Econ.*, vol. 16, no. 5, pp. 601–608, Oct. 1997.

[12] J. J. Louviere, D. Street, L. Burgess, N. Wasi, T. Islam, and A. A. J. Marley, "Modeling the choices of individual decision-makers by combining efficient choice experiment designs with extra preference information," *J. Choice Model.*, vol. 1, no. 1, pp. 128–164, Jan. 2008.

[13] M. R. Mørkbak, T. Christensen, and D. Gyrd-Hansen, "Choke Price Bias in Choice Experiments," *Environ. Resour. Econ.*, vol. 45, no. 4, pp. 537–551, Apr. 2010.

[14] J. M. Rose and M. C. J. Bliemer, "Constructing Efficient Stated Choice Experimental Designs," *Transp. Rev.*, vol. 29, no. 5, pp. 587–617, Sep. 2009.

[15] F. Reed Johnson *et al.*, "Constructing Experimental Designs for Discrete-Choice Experiments: Report of the ISPOR Conjoint Analysis Experimental Design Good Research Practices Task Force," *Value Health*, vol. 16, no. 1, pp. 3–13, Jan. 2013.

[16] J. J. Louviere, D. A. Hensher, J. D. Swait, and W. Adamowicz, *Stated choice methods: analysis and applications*, 7. printing. Cambridge: Cambridge Univ. Press, 2010.

[17] D. McFadden, "Conditional logit analysis of qualitative choice behavior," in *Frontier in econometrics*, New York: Academic Press, 1974, pp. 105–142.

[18] K. Train, *Discrete choice methods with simulation*, 2nd ed. Cambridge ; New York: Cambridge University Press, 2009.

[19] D. Kahneman and A. Tversky, "Prospect Theory: An Analysis of Decision under Risk," *Econometrica*, vol. 47, no. 2, p. 263, Mar. 1979.

[20] W. Samuelson and R. Zeckhauser, "Status Quo Bias in Decision Making," *J. Risk Uncertain.*, vol. 1, no. 1, pp. 7–59, 1988.

[21] D. Kahneman, J. L. Knetsch, and R. H. Thaler, "Anomalies: The Endowment Effect, Loss Aversion, and Status Quo Bias," *J. Econ. Perspect.*, vol. 5, no. 1, pp. 193–206, 1991.

[22] P. Bonsall and B. Lythgoe, "Factors affecting the amount of effort expended in responding to questions in behavioural choice experiments," *J. Choice Model.*, vol. 2, no. 2, pp. 216–236, 2009.

[23] J. J. Louviere, T. Islam, N. Wasi, D. Street, and L. Burgess, "Designing Discrete Choice Experiments: Do Optimal Designs Come at a Price?," *J. Consum. Res.*, vol. 35, no. 2, pp. 360–375, Aug. 2008.

[24] I. J. Bateman, D. Burgess, W. G. Hutchinson, and D. I. Matthews, "Learning design contingent valuation (LDCV): NOAA guidelines, preference learning and coherent arbitrariness," *J. Environ. Econ. Manag.*, vol. 55, no. 2, pp. 127–141, 2008.

[25] B. Day *et al.*, "Ordering effects and choice set awareness in repeat-response stated preference studies," *J. Environ. Econ. Manag.*, vol. 63, no. 1, pp. 73–91, Jan. 2012.

[26] R. T. Yao, R. Scarpa, J. M. Rose, and J. A. Turner, "Experimental Design Criteria and Their Behavioural Efficiency: An Evaluation in the Field," *Environ. Resour. Econ.*, vol. 62, no. 3, pp. 433–455, Nov. 2015.

[27] R. Viney, E. Savage, and J. Louviere, “Empirical investigation of experimental design properties of discrete choice experiments in health care,” *Health Econ.*, vol. 14, no. 4, pp. 349–362, Apr. 2005.
[28] A. Diederich, J. Swait, and N. Wirsik, “Citizen Participation in Patient Prioritization Policy Decisions: An Empirical and Experimental Study on Patients’ Characteristics,” *PLoS ONE*, vol. 7, no. 5, p. e36824, May 2012.
[29] S. Erdem and C. Thompson, “Prioritising health service innovation investments using public preferences: a discrete choice experiment,” *BMC Health Serv. Res.*, vol. 14, no. 1, Dec. 2014.
[30] M. K. Lim, E. Y. Bae, S.-E. Choi, E. K. Lee, and T.-J. Lee, “Eliciting Public Preference for Health-Care Resource Allocation in South Korea,” *Value Health*, vol. 15, no. 1, pp. S91–S94, Jan. 2012.
[31] P. A. Scuffham *et al.*, “Engaging the public in healthcare decision-making: quantifying preferences for healthcare through citizens’ juries,” *BMJ Open*, vol. 4, no. 5, p. e005437, Apr. 2014.
[32] D. L. B. Schwappach and T. J. Strasmann, “‘Quick and dirty numbers’?,” *J. Health Econ.*, vol. 25, no. 3, pp. 432–448, May 2006.
[33] D. L. B. Schwappach, “Does it matter who you are or what you gain? an experimental study of preferences for resource allocation,” *Health Econ.*, vol. 12, no. 4, pp. 255–267, Apr. 2003.
[34] C. Green and K. Gerard, “Exploring the social value of health-care interventions: a stated preference discrete choice experiment,” *Health Econ.*, vol. 18, no. 8, pp. 951–976, Aug. 2009.
[35] C. D. Skedgel, A. J. Wailoo, and R. L. Akehurst, “Choosing vs. allocating: discrete choice experiments and constant-sum paired comparisons for the elicitation of societal preferences,” *Health Expect.*, vol. 18, no. 5, pp. 1227–1240, Oct. 2015.
[36] C. Skedgel, A. Wailoo, and R. Akehurst, “Societal Preferences for Distributive Justice in the Allocation of Health Care Resources: A Latent Class Discrete Choice Experiment,” *Med. Decis. Making*, vol. 35, no. 1, pp. 94–105, Jan. 2015.

FIGURES LEGEND

For Figure 1:

Source: Sendi, P., Gafni, A. & Birch, S. Opportunity costs and uncertainty in the economic evaluation of health care interventions. *Health Econ.* **11,** 23–31 (2002).

Reading: ΔC(A) indicates the incremental costs for programme A; ΔC(B) indicates the incremental costs for programme B; ΔE(A) indicates the incremental health outcomes for programme A; ΔE(B) indicates the incremental health outcomes for programme B; The Latin numbers (I, II, III, IV) are used to describe the four quadrants of the DMP: {Increase in health outcomes; Decrease in medical expenditures}; {Decrease in health outcomes; Decrease in medical expenditures}; {Decrease in health outcomes; Increase in medical expenditures}; {Increase in health outcomes; Increase in medical expenditures}. Panel B illustrates the posterior joint distribution taken from an example used in Sendi et al (2002) of introducing programme A and cancelling programme B in the decision making plane.

Figure 1. The Decision Making Plane (DMP)

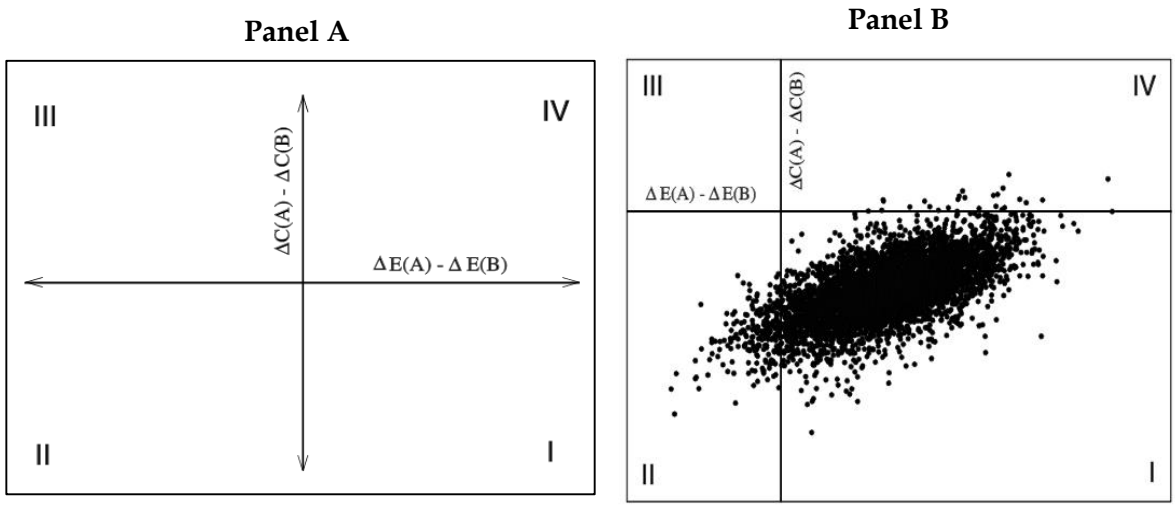


Figure 2. Illustration of choice task format

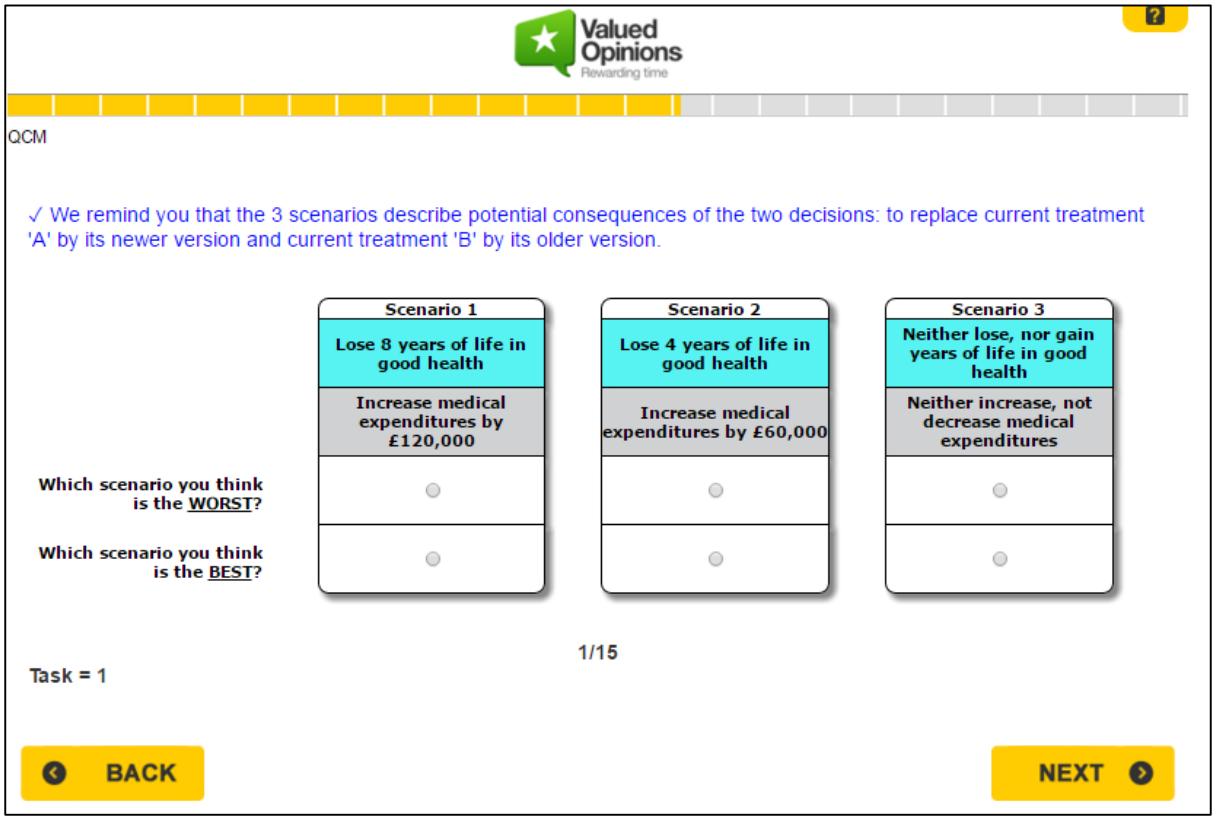


Figure 3. Comparison of predictive performance values across the four versions

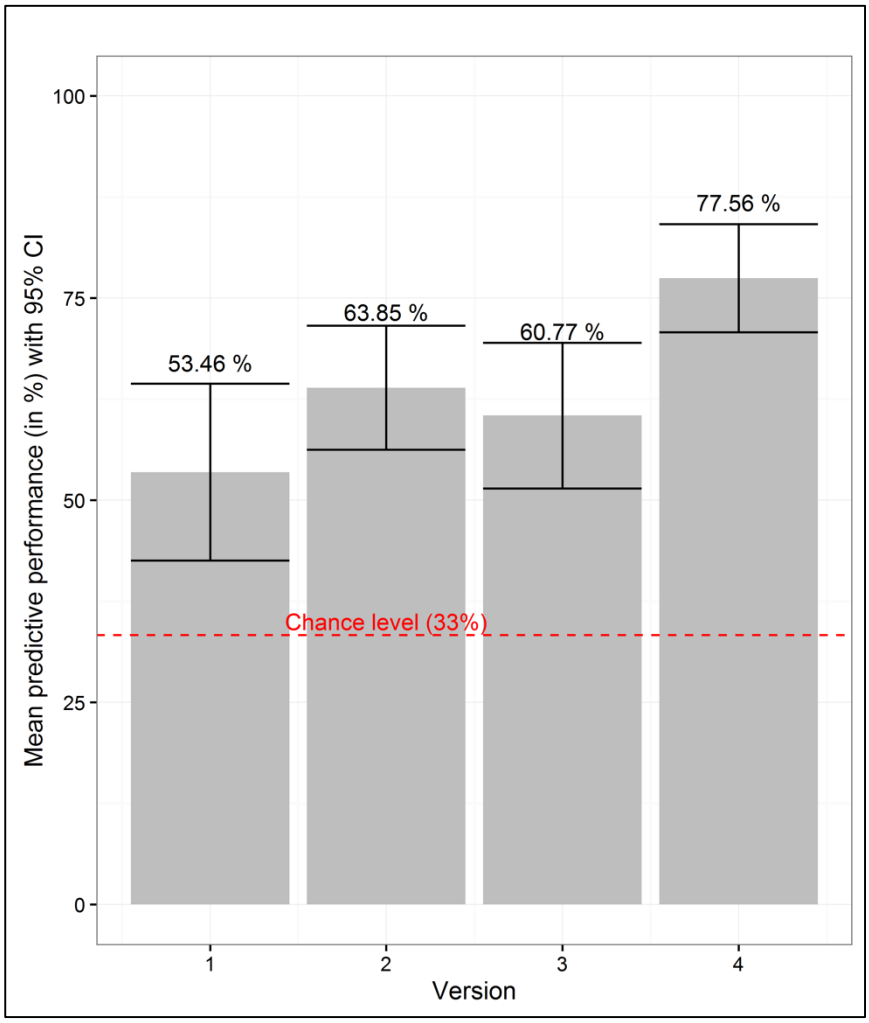


Table 1. Comparison of debriefing questions across the four versions

Characteristic*	Level	V1 (N=50)	V2 (N=50)	V3 (N=50)	V4 (N=50)
1. Interest of the questionnaire (P-value = 0.004)					
	Very interesting	36%	26%	32%	58%
	Interesting	38%	54%	32%	20%
	Moderately/Slightly/Not at all interesting	26%	20%	36%	22%
2. Difficulty of the questionnaire (P-value = 0.109)					
	Very easy	12%	8%	14%	26%
	Easy	44%	48%	40%	32%
	Neither easy, nor difficult	22%	26%	30%	36%
	Difficult/Very difficult	22%	18%	16%	6%
3. Decision making objective (P-value = 0.205)					
	To minimise health expenditures	16%	10%	6%	4%
	To maximise health outcomes	42%	32%	42%	32%
	To find a compromise between health expenditures and outcomes	42%	58%	52%	64%
4. Minimum acceptable level of health outcomes (P-value = 0.039)					
	1st worst/2nd worst level (Loss)	18%	8%	6%	6%
	Neutral level (No gain, no loss)	40%	28%	44%	24%
	2nd best level (Gain)	18%	34%	20%	22%
	1st best level (Gain)	24%	30%	30%	48%
5. Minimum acceptable level of health expenditures (P-value = 0.584)					
	1st worst level (Loss/Increase)	18%	8%	16%	22%
	2nd worst level (Loss/Increase)	12%	10%	10%	8%
	Neutral level (No gain, no loss)	40%	36%	36%	42%
	2nd best level (Gain/Decrease)	6%	22%	12%	12%
	1st best level (Gain/Decrease)	24%	24%	26%	16%
6. Importance of health outcomes (P-value = 0.667)					
	Very important	48%	50%	56%	64%
	Important	32%	34%	32%	26%
	Moderately/Slightly/Not at all important	20%	16%	12%	10%
7. Importance of health expenditures (P-value = 0.485)					
	Very important	20%	14%	26%	14%
	Important	38%	52%	34%	44%
	Moderately important	30%	28%	30%	24%
	Slightly/Not at all important	12%	6%	10%	18%

** P-value of Chi-2 test reported in brackets*

Table 2. Comparison of choices properties across the four versions

Characteristic	Level	V1	V2	V3	V4
1. Monotonicity (P-value = 0.147)					
	No	20%	16%	22%	6%
	Yes	80%	84%	78%	94%
2. Stability (P-value < 0.001)					
	No	32%	20%	8%	8%
	Partial	16%	36%	38%	14%
	Full	52%	44%	54%	78%
3. Serial non-participation (P-value = 0.283)					
	No	94%	94%	86%	96%
	Yes	6%	6%	14%	4%
4. Response time (P-value = 0.580)					
	"Normal"	86%	88%	86%	94%
	"Speedster"	14%	12%	14%	6%

Table 3. Best fitting choice model for the different versions of the choice experiment

Parameters		Version 1			Version 2			Version 3			Version 4		
		MLE	**CRSE**	**P-val.**	**MLE**	**CRSE**	**P-val.**	**MLE**	**CRSE**	**P-val.**	**MLE**	**CRSE**	**P-val.**
1. Estimated preferences													
HYE	*(Continuous)*	0.287	0.053	< 0.001	-	-	-	0.254	0.035	< 0.001	-	-	-
	Max. loss	-	-	-	-0.617	0.268	0.022	-	-	-	-3.399	0.335	< 0.001
	Min. loss	-	-	-	-1.284	0.195	< 0.001	-	-	-	-1.459	0.161	< 0.001
	No loss, lo gain (Ref.)	-	-	-	-0.060	-	-	-	-	-	0.627	-	-
	Min. gain	-	-	-	1.296	0.170	< 0.001	-	-	-	1.764	0.189	< 0.001
	Max. gain	-	-	-	0.664	0.228	0.004	-	-	-	2.468	0.284	< 0.001
KCOST	*(Continuous)*	0.118	0.028	< 0.001	-	-	-	0.102	0.017	< 0.001	0.091	0.011	< 0.001
	Max. loss	-	-	-	-1.236	0.267	< 0.001	-	-	-	-	-	-
	Min. loss	-	-	-	-0.431	0.153	0.005	-	-	-	-	-	-
	No loss, lo gain (Ref.)	-	-	-	0.650	-	-	-	-	-	-	-	-
	Min. gain	-	-	-	1.138	0.229	< 0.001	-	-	-	-	-	-
	Max. gain	-	-	-	-0.120	0.343	0.726	-	-	-	-	-	-
2. Other effects													
Decision bias: Neutral situation		0.234	0.100	0.019	-	-	-	0.683	0.145	< 0.001	-	-	-
Interaction: HYE ´ KCOST		0.014	0.005	0.003	-0.014	0.005	0.005	0.003	0.002	0.089	-	-	-
Choice consistency: BEST vs WORST		-	-	-	-	-	-	-	-	-	-	-	-
3. Model statistics													
# Observations			1,000			1,000			1,000			1,200	
# Parameters			4			9			4			5	
Log-likelihood			-995.9			-862.9			-869.9			-806.0	
BIC			2,019.4			1,787.9			1,767.3			1,647.5	

MLE: Maximum Likelihood Estimate; CRSE: Cluster Robust Standard Errors; BIC: Bayesian Information Criterion

Online supplementary material

- Document: Estimation of different choice model specifications

- Journal: PharmacoEconomics
- Article: “Measuring public preferences for health outcomes and expenditures in a context of healthcare resources re-allocation”
- Authors: Nicolas KRUCIEN *; Nathalie PELLETIER-FLEURY; Amiram GAFNI

* Nicolas KRUCIEN, PhD
Health Economics Research Unit
Institute of Applied Health Sciences
University of Aberdeen
Aberdeen, AB25 2QN
Tel: +44(0)1-224-437-892
Fax: +44(0)1-224-437-195
Email: nicolas.krucien@abdn.ac.uk

Supplementary material: Estimation of different choice model specifications

For each of the four DCE versions it was possible to estimate 32 choice models based on different specifications of the indirect utility function (for a grand total of 32 x 4 = 128 models):

1. The preferences for the HYE (ΔE) attribute are linear vs. non-linear;
2. The preferences for the KCOST (ΔC) attribute are linear vs. non-linear;
3. Participants have a systematic preference for the neutral change situation "NCS" (ASC_SQ) vs don't have;
4. Preferences for the HYE and KCOST attributes are dependent vs. independent;
5. Consistency of participants' choices differ between BEST and WORST choices vs. don't differ

The best specification is the one minimising the Bayesian Information Criterion (BIC). In the below table, the best fitting model is highlighted in grey colour for each DCE version, and detailed results for these best models are presented in the article.

Table. List of all possible choice models (Part 1)

Model	Version	ASC_SQ	HYE	KCOST	Interaction	Scale	#par	#Obs	Loglik.	BIC
17	1	Yes	Linear	Linear	Yes	No	4	1,000	-995.9	2,019.4
25	1	No	Linear	Linear	Yes	No	3	1,000	-1,000.9	2,022.6
31	1	No	Linear	Non-linear	No	No	5	1,000	-995.0	2,024.6
1	1	Yes	Linear	Linear	Yes	Yes	5	1,000	-995.8	2,026.2
21	1	Yes	Linear	Linear	No	No	3	1,000	-1,002.9	2,026.5
27	1	No	Linear	Non-linear	Yes	No	6	1,000	-993.4	2,028.3
30	1	No	Non-linear	Linear	No	No	5	1,000	-996.9	2,028.3
9	1	No	Linear	Linear	Yes	Yes	4	1,000	-1,000.9	2,029.5
15	1	No	Linear	Non-linear	No	Yes	6	1,000	-994.9	2,031.3
23	1	Yes	Linear	Non-linear	No	No	6	1,000	-995.0	2,031.5
5	1	Yes	Linear	Linear	No	Yes	4	1,000	-1,002.6	2,032.8
22	1	Yes	Non-linear	Linear	No	No	6	1,000	-995.7	2,032.9
26	1	No	Non-linear	Linear	Yes	No	6	1,000	-996.4	2,034.3
11	1	No	Linear	Non-linear	Yes	Yes	7	1,000	-993.4	2,035.1
14	1	No	Non-linear	Linear	No	Yes	6	1,000	-996.8	2,035.1
19	1	Yes	Linear	Non-linear	Yes	No	7	1,000	-993.4	2,035.1
7	1	Yes	Linear	Non-linear	No	Yes	7	1,000	-994.9	2,038.2
18	1	Yes	Non-linear	Linear	Yes	No	7	1,000	-995.0	2,038.3
32	1	No	Non-linear	Non-linear	No	No	8	1,000	-991.7	2,038.6
6	1	Yes	Non-linear	Linear	No	Yes	7	1,000	-995.7	2,039.7
29	1	No	Linear	Linear	No	No	2	1,000	-1,013.4	2,040.6
10	1	No	Non-linear	Linear	Yes	Yes	7	1,000	-996.4	2,041.2
3	1	Yes	Linear	Non-linear	Yes	Yes	8	1,000	-993.3	2,041.9
2	1	Yes	Non-linear	Linear	Yes	Yes	8	1,000	-994.9	2,045.2
16	1	No	Non-linear	Non-linear	No	Yes	9	1,000	-991.6	2,045.4
24	1	Yes	Non-linear	Non-linear	No	No	9	1,000	-991.7	2,045.5
28	1	No	Non-linear	Non-linear	Yes	No	9	1,000	-991.7	2,045.5
13	1	No	Linear	Linear	No	Yes	3	1,000	-1,013.1	2,047.0
8	1	Yes	Non-linear	Non-linear	No	Yes	10	1,000	-991.6	2,052.3
12	1	No	Non-linear	Non-linear	Yes	Yes	10	1,000	-991.6	2,052.3
20	1	Yes	Non-linear	Non-linear	Yes	No	10	1,000	-991.7	2,052.4
4	1	Yes	Non-linear	Non-linear	Yes	Yes	11	1,000	-991.6	2,059.2

Table. List of all possible choice models (Part 2)

Model	Version	ASC_SQ	HYE	KCOST	Interaction	Scale	#par	#Obs	Loglik.	BIC
28	2	No	Non-linear	Non-linear	Yes	No	9	1,000	-862.9	1,787.9
32	2	No	Non-linear	Non-linear	No	No	8	1,000	-867.3	1,789.8
30	2	No	Non-linear	Linear	No	No	5	1,000	-878.0	1,790.5
20	2	Yes	Non-linear	Non-linear	Yes	No	10	1,000	-862.0	1,793.1
12	2	No	Non-linear	Non-linear	Yes	Yes	10	1,000	-862.7	1,794.5
22	2	Yes	Non-linear	Linear	No	No	6	1,000	-876.9	1,795.3
24	2	Yes	Non-linear	Non-linear	No	No	9	1,000	-866.9	1,796.0
26	2	No	Non-linear	Linear	Yes	No	6	1,000	-877.3	1,796.0
14	2	No	Non-linear	Linear	No	Yes	6	1,000	-877.3	1,796.1
16	2	No	Non-linear	Non-linear	No	Yes	9	1,000	-867.3	1,796.7
4	2	Yes	Non-linear	Non-linear	Yes	Yes	11	1,000	-862.0	1,799.9
18	2	Yes	Non-linear	Linear	Yes	No	7	1,000	-876.0	1,800.4
6	2	Yes	Non-linear	Linear	No	Yes	7	1,000	-876.3	1,800.9
10	2	No	Non-linear	Linear	Yes	Yes	7	1,000	-877.0	1,802.3
8	2	Yes	Non-linear	Non-linear	No	Yes	10	1,000	-866.9	1,802.9
2	2	Yes	Non-linear	Linear	Yes	Yes	8	1,000	-875.8	1,806.9
27	2	No	Linear	Non-linear	Yes	No	6	1,000	-903.0	1,847.5
31	2	No	Linear	Non-linear	No	No	5	1,000	-907.8	1,850.1
11	2	No	Linear	Non-linear	Yes	Yes	7	1,000	-901.7	1,851.7
19	2	Yes	Linear	Non-linear	Yes	No	7	1,000	-901.9	1,852.2
23	2	Yes	Linear	Non-linear	No	No	6	1,000	-906.9	1,855.2
15	2	No	Linear	Non-linear	No	Yes	6	1,000	-907.2	1,855.9
3	2	Yes	Linear	Non-linear	Yes	Yes	8	1,000	-900.7	1,856.6
7	2	Yes	Linear	Non-linear	No	Yes	7	1,000	-906.4	1,861.2
17	2	Yes	Linear	Linear	Yes	No	4	1,000	-925.0	1,877.7
21	2	Yes	Linear	Linear	No	No	3	1,000	-930.0	1,880.7
1	2	Yes	Linear	Linear	Yes	Yes	5	1,000	-924.9	1,884.3
5	2	Yes	Linear	Linear	No	Yes	4	1,000	-929.2	1,886.1
25	2	No	Linear	Linear	Yes	No	3	1,000	-934.0	1,888.8
9	2	No	Linear	Linear	Yes	Yes	4	1,000	-933.8	1,895.3
29	2	No	Linear	Linear	No	No	2	1,000	-946.6	1,907.1
13	2	No	Linear	Linear	No	Yes	3	1,000	-946.6	1,914.0

Table. List of all possible choice models (Part 3)

Model	Version	ASC_SQ	HYE	KCOST	Interaction	Scale	#par	#Obs	Loglik.	BIC
17	3	Yes	Linear	Linear	Yes	No	4	1,000	-869.9	1,767.3
21	3	Yes	Linear	Linear	No	No	3	1,000	-873.3	1,767.4
5	3	Yes	Linear	Linear	No	Yes	4	1,000	-873.0	1,773.7
1	3	Yes	Linear	Linear	Yes	Yes	5	1,000	-869.7	1,774.0
23	3	Yes	Linear	Non-linear	No	No	6	1,000	-868.1	1,777.6
22	3	Yes	Non-linear	Linear	No	No	6	1,000	-868.2	1,777.8
18	3	Yes	Non-linear	Linear	Yes	No	7	1,000	-868.0	1,784.4
6	3	Yes	Non-linear	Linear	No	Yes	7	1,000	-868.1	1,784.5
7	3	Yes	Linear	Non-linear	No	Yes	7	1,000	-868.1	1,784.5
19	3	Yes	Linear	Non-linear	Yes	No	7	1,000	-868.1	1,784.5
2	3	Yes	Non-linear	Linear	Yes	Yes	8	1,000	-867.9	1,791.1
3	3	Yes	Linear	Non-linear	Yes	Yes	8	1,000	-868.0	1,791.3
24	3	Yes	Non-linear	Non-linear	No	No	9	1,000	-867.0	1,796.2
20	3	Yes	Non-linear	Non-linear	Yes	No	10	1,000	-864.0	1,797.0
30	3	No	Non-linear	Linear	No	No	5	1,000	-881.5	1,797.6
32	3	No	Non-linear	Non-linear	No	No	8	1,000	-873.2	1,801.7
8	3	Yes	Non-linear	Non-linear	No	Yes	10	1,000	-867.0	1,803.0
4	3	Yes	Non-linear	Non-linear	Yes	Yes	11	1,000	-863.9	1,803.8
26	3	No	Non-linear	Linear	Yes	No	6	1,000	-881.3	1,804.1
14	3	No	Non-linear	Linear	No	Yes	6	1,000	-881.5	1,804.4
28	3	No	Non-linear	Non-linear	Yes	No	9	1,000	-872.4	1,806.9
16	3	No	Non-linear	Non-linear	No	Yes	9	1,000	-873.1	1,808.4
10	3	No	Non-linear	Linear	Yes	Yes	7	1,000	-881.3	1,811.0
31	3	No	Linear	Non-linear	No	No	5	1,000	-888.5	1,811.4
12	3	No	Non-linear	Non-linear	Yes	Yes	10	1,000	-872.3	1,813.7
27	3	No	Linear	Non-linear	Yes	No	6	1,000	-887.5	1,816.4
15	3	No	Linear	Non-linear	No	Yes	6	1,000	-888.1	1,817.6
11	3	No	Linear	Non-linear	Yes	Yes	7	1,000	-887.2	1,822.7
25	3	No	Linear	Linear	Yes	No	3	1,000	-903.7	1,828.0
9	3	No	Linear	Linear	Yes	Yes	4	1,000	-903.6	1,834.8
29	3	No	Linear	Linear	No	No	2	1,000	-925.4	1,864.6
13	3	No	Linear	Linear	No	Yes	3	1,000	-925.3	1,871.4

Table. List of all possible choice models (Part 4)

Model	Version	ASC_SQ	HYE	KCOST	Interaction	Scale	#par	#Obs	Loglik.	BIC
30	4	No	Non-linear	Linear	No	No	5	1,200	-806.0	1,647.5
22	4	Yes	Non-linear	Linear	No	No	6	1,200	-805.4	1,653.4
14	4	No	Non-linear	Linear	No	Yes	6	1,200	-806.0	1,654.6
26	4	No	Non-linear	Linear	Yes	No	6	1,200	-806.0	1,654.6
32	4	No	Non-linear	Non-linear	No	No	8	1,200	-799.8	1,656.2
6	4	Yes	Non-linear	Linear	No	Yes	7	1,200	-805.4	1,660.4
18	4	Yes	Non-linear	Linear	Yes	No	7	1,200	-805.4	1,660.5
24	4	Yes	Non-linear	Non-linear	No	No	9	1,200	-798.8	1,661.4
10	4	No	Non-linear	Linear	Yes	Yes	7	1,200	-806.0	1,661.7
16	4	No	Non-linear	Non-linear	No	Yes	9	1,200	-799.6	1,663.0
28	4	No	Non-linear	Non-linear	Yes	No	9	1,200	-799.7	1,663.2
2	4	Yes	Non-linear	Linear	Yes	Yes	8	1,200	-805.4	1,667.5
8	4	Yes	Non-linear	Non-linear	No	Yes	10	1,200	-798.6	1,668.1
20	4	Yes	Non-linear	Non-linear	Yes	No	10	1,200	-798.8	1,668.4
12	4	No	Non-linear	Non-linear	Yes	Yes	10	1,200	-799.6	1,670.0
4	4	Yes	Non-linear	Non-linear	Yes	Yes	11	1,200	-798.6	1,675.2
27	4	No	Linear	Non-linear	Yes	No	6	1,200	-816.4	1,675.3
17	4	Yes	Linear	Linear	Yes	No	4	1,200	-824.2	1,676.8
31	4	No	Linear	Non-linear	No	No	5	1,200	-822.3	1,680.0
11	4	No	Linear	Non-linear	Yes	Yes	7	1,200	-815.9	1,681.3
19	4	Yes	Linear	Non-linear	Yes	No	7	1,200	-816.4	1,682.3
1	4	Yes	Linear	Linear	Yes	Yes	5	1,200	-823.8	1,683.1
21	4	Yes	Linear	Linear	No	No	3	1,200	-831.6	1,684.4
23	4	Yes	Linear	Non-linear	No	No	6	1,200	-822.2	1,686.9
15	4	No	Linear	Non-linear	No	Yes	6	1,200	-822.2	1,687.0
3	4	Yes	Linear	Non-linear	Yes	Yes	8	1,200	-815.8	1,688.4
5	4	Yes	Linear	Linear	No	Yes	4	1,200	-831.6	1,691.5
7	4	Yes	Linear	Non-linear	No	Yes	7	1,200	-822.2	1,694.0
25	4	No	Linear	Linear	Yes	No	3	1,200	-840.6	1,702.6
9	4	No	Linear	Linear	Yes	Yes	4	1,200	-840.4	1,709.2
29	4	No	Linear	Linear	No	No	2	1,200	-853.9	1,722.0
13	4	No	Linear	Linear	No	Yes	3	1,200	-853.8	1,728.8

Online supplementary material

- Document: Copy of the final version of the questionnaire in English language

- Journal: PharmacoEconomics
- Article: “Measuring public preferences for health outcomes and expenditures in a context of healthcare resources re-allocation”
- Authors: Nicolas KRUCIEN *; Nathalie PELLETIER-FLEURY; Amiram GAFNI

* Nicolas KRUCIEN, PhD
Health Economics Research Unit
Institute of Applied Health Sciences
University of Aberdeen
Aberdeen, AB25 2QN
Tel: +44(0)1-224-437-892
Fax: +44(0)1-224-437-195
Email: nicolas.krucien@abdn.ac.uk

PAGE 1

Study about funding of new medical interventions

The objective of the study is to understand what your preferences for funding of new medical interventions are.

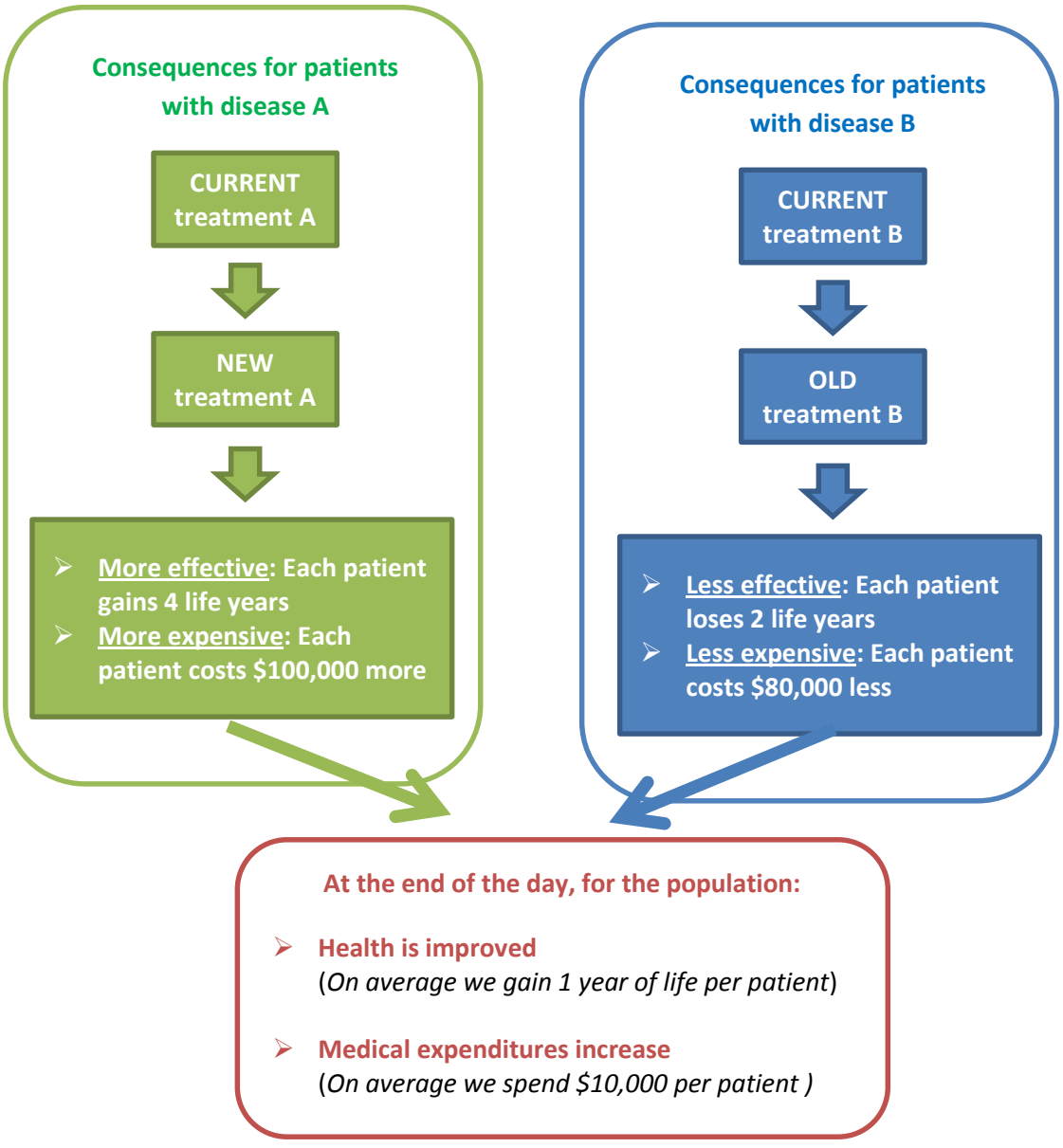
How to fund new interventions when the healthcare budget is limited?

Like any other public expenditures (e.g., education, transportation, preserving the environment), the health budget is limited. Every year new medical treatments are developed. Typically new treatments are more effective but also more expensive than those currently available. Replacing current treatments by the new ones would increase the level of medical expenditures. Because of budget limitations it is impossible to increase medical expenditures without, for example, raising the income tax, decreasing other public expenditures, asking users/patients to pay more when they receive medical services. If health policy makers do not want to use this measures they need to find other options to fund the new medical interventions. One option would be to replace other treatments by their older versions, which are typically less expensive but also less effective, and then to use the freed up money to fund the new treatments. This option of replacing other treatments is the topic of this study.

What are the consequences of funding new medical interventions?

The decision to replace a treatment by its older version to free up enough money for funding the new version of another treatment is difficult to make. The patients who will have access to the newer - *more effective* - treatment will be better off but those who will receive the older - *less effective* - treatment will be worse off. However health policy decision-makers are willing to make this decision because at the end of the day they believe this will improve the overall health of the population (all patients confounded) without having to increase healthcare expenditures (or eventually by lowering them).

In this example we illustrate the consequences of a decision to replace treatments. Suppose there are two different diseases, let's call them "**disease A**" and "**disease B**", which affect the same number of patients. Treatments are currently in place for these two diseases, but a new one has been recently developed for the "**disease A**" and the health policy decision-makers consider the possibility to implement this new treatment that will replace the existing treatment. However there is not enough money left out in the healthcare budget. The health policy decision-maker decides to get the additional money needed by replacing the current version of the treatment for "**disease B**" by an older version, which is less expensive but less effective. This decision to replace the treatments for both "disease A" and "disease B" will have important consequences in terms of health outcomes and medical expenditures. You can see an illustration in the following figure:



PAGE 3

Funding new treatments is a risky decision

As it is the case with other public policy decisions, health decisions can lead to unexpected results (for example, the new treatment may be less effective or more expensive than expected).

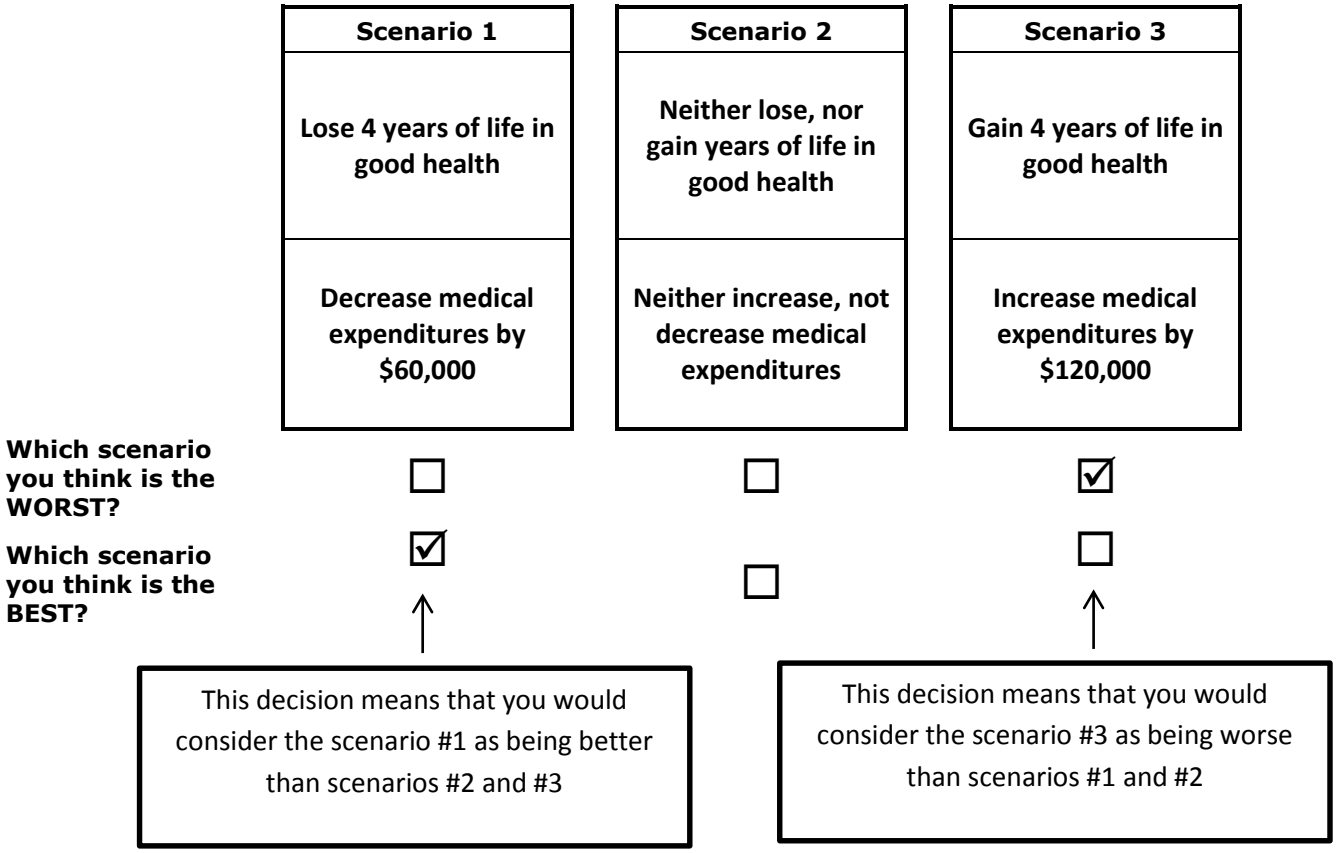
Because of this uncertainty in the consequences of health decisions, the decision to replace current treatments may improve the overall health of the population as predicted, more than predicted, less than predicted or even have the opposite effect and reduce it, affecting health care expenditures can also be either positive (i.e., no additional funds required or even savings) or negative (i.e., additional funds still required).

Health policy decision-makers take into account this uncertainty in the consequences of their decisions by anticipating different "scenarios" that they might face. Given these different scenarios and their likelihood to occur, the health policy makers will finally decide whether or not it is worth replacing the current treatment "A" by its newer version and the current treatment "B" by its older version.

PAGE 4

Now, it's your turn to decide

In the remaining of this questionnaire, we are going to ask you to take the seat of the policy decision-makers and to make your own decisions of funding new medical interventions.
We will give you a list of **choice tasks**. Each choice task is composed of **3 different scenarios** describing the consequences of replacing current treatments A and B by their new and old versions respectively. In every choice task we ask you to select the scenario you think is the **BEST** and the one you think is the **WORST.**



Few last details before making your own decisions

All the information you need to make your decisions is provided within the choice tasks. For your decisions, it is not important to have more details about the exact type of treatments, the profiles and numbers of patients who benefit from treatments A and B.

There is no right or wrong answers. What matters is what you think.

We understand that decisions to fund new medical interventions can be difficult. However we would like to ask you to take your time and to pay attention to the different pieces of information (number of years of life in good health, medical expenditures) before making your own decisions.

Thanks for your attention, now it's your turn!

PAGE 5

- We remind you that the 3 scenarios describe potential consequences of the two decisions: to replace current treatment "A" by its newer version and current treatment "B" by its older version.

CHOICE SET 1

Scenario 1	Scenario 2	Scenario 3
Lose 8 years of life in good health	**Lose 4 years of life in good health**	**Neither lose, nor gain years of life in good health**
Decrease medical expenditures by $120,000	**Decrease medical expenditures by $60,000**	**Neither increase, not decrease medical expenditures**

Which scenario you think is the WORST?

Which scenario you think is the BEST?

PAGE 6

- We remind you that the 3 scenarios describe potential consequences of the two decisions: to replace current treatment "A" by its newer version and current treatment "B" by its older version.

CHOICE SET 2

Scenario 1	Scenario 2	Scenario 3
Lose 4 years of life in good health	**Neither lose, nor gain years of life in good health**	**Neither lose, nor gain years of life in good health**
Decrease medical expenditures by $60,000	**Neither increase, not decrease medical expenditures**	**Neither increase, not decrease medical expenditures**

Which scenario you think is the WORST?

Which scenario you think is the BEST?

PAGE 7

- We remind you that the 3 scenarios describe potential consequences of the two decisions: to replace current treatment "A" by its newer version and current treatment "B" by its older version.

CHOICE SET 3

Scenario 1	Scenario 2	Scenario 3
Lose 8 years of life in good health	**Lose 4 years of life in good health**	**Neither lose, nor gain years of life in good health**
Decrease medical expenditures by $120,000	**Decrease medical expenditures by $60,000**	**Neither increase, not decrease medical expenditures**

Which scenario you think is the WORST?

Which scenario you think is the BEST?

PAGE 8

- We remind you that the 3 scenarios describe potential consequences of the two decisions: to replace current treatment "A" by its newer version and current treatment "B" by its older version.

CHOICE SET 4

Scenario 1	Scenario 2	Scenario 3
Gain 4 years of life in good health	**Gain 8 years of life in good health**	**Neither lose, nor gain years of life in good health**
Neither increase, not decrease medical expenditures	**Decrease medical expenditures by $120,000**	**Neither increase, not decrease medical expenditures**

Which scenario you think is the WORST?

Which scenario you think is the BEST?

PAGE 9

- We remind you that the 3 scenarios describe potential consequences of the two decisions: to replace current treatment "A" by its newer version and current treatment "B" by its older version.

CHOICE SET 5

Scenario 1	Scenario 2	Scenario 3
Lose 8 years of life in good health	**Lose 4 years of life in good health**	**Neither lose, nor gain years of life in good health**
Decrease medical expenditures by $60,000	**Increase medical expenditures by $120,000**	**Neither increase, not decrease medical expenditures**

Which scenario you think is the WORST?

Which scenario you think is the BEST?

PAGE 10

- We remind you that the 3 scenarios describe potential consequences of the two decisions: to replace current treatment "A" by its newer version and current treatment "B" by its older version.

CHOICE SET 6

Scenario 1	Scenario 2	Scenario 3
Gain 8 years of life in good health	**Gain 4 years of life in good health**	**Neither lose, nor gain years of life in good health**
Increase medical expenditures by $120,000	**Decrease medical expenditures by $60,000**	**Neither increase, not decrease medical expenditures**

Which scenario you think is the WORST?

Which scenario you think is the BEST?

PAGE 11

- We remind you that the 3 scenarios describe potential consequences of the two decisions: to replace current treatment "A" by its newer version and current treatment "B" by its older version.

CHOICE SET 7

Scenario 1	Scenario 2	Scenario 3
Neither lose, nor gain years of life in good health	**Gain 8 years of life in good health**	**Neither lose, nor gain years of life in good health**
Decrease medical expenditures by $120,000	**Increase medical expenditures by $60,000**	**Neither increase, not decrease medical expenditures**

Which scenario you think is the WORST?

Which scenario you think is the BEST?

PAGE 12

- We remind you that the 3 scenarios describe potential consequences of the two decisions: to replace current treatment "A" by its newer version and current treatment "B" by its older version.

CHOICE SET 8

Scenario 1	Scenario 2	Scenario 3
Lose 4 years of life in good health	**Lose 8 years of life in good health**	**Neither lose, nor gain years of life in good health**
Increase medical expenditures by $60,000	**Decrease medical expenditures by $120,000**	**Neither increase, not decrease medical expenditures**

Which scenario you think is the WORST?

Which scenario you think is the BEST?

PAGE 13

- We remind you that the 3 scenarios describe potential consequences of the two decisions: to replace current treatment "A" by its newer version and current treatment "B" by its older version.

CHOICE SET 9

Scenario 1	Scenario 2	Scenario 3
Lose 4 years of life in good health	**Gain 4 years of life in good health**	**Neither lose, nor gain years of life in good health**
Neither increase, not decrease medical expenditures	**Increase medical expenditures by $120,000**	**Neither increase, not decrease medical expenditures**

Which scenario you think is the WORST?

Which scenario you think is the BEST?

PAGE 14

- We remind you that the 3 scenarios describe potential consequences of the two decisions: to replace current treatment "A" by its newer version and current treatment "B" by its older version.

CHOICE SET 10

Scenario 1	Scenario 2	Scenario 3
Gain 8 years of life in good health	**Lose 4 years of life in good health**	**Neither lose, nor gain years of life in good health**
Decrease medical expenditures by $60,000	**Decrease medical expenditures by $120,000**	**Neither increase, not decrease medical expenditures**

Which scenario you think is the WORST?

Which scenario you think is the BEST?

PAGE 15

- We remind you that the 3 scenarios describe potential consequences of the two decisions: to replace current treatment "A" by its newer version and current treatment "B" by its older version.

CHOICE SET 11

Scenario 1	Scenario 2	Scenario 3
Lose 8 years of life in good health	**Lose 8 years of life in good health**	**Neither lose, nor gain years of life in good health**
Increase medical expenditures by $120,000	**Increase medical expenditures by $60,000**	**Neither increase, not decrease medical expenditures**

Which scenario you think is the WORST?

Which scenario you think is the BEST?

PAGE 16

- We remind you that the 3 scenarios describe potential consequences of the two decisions: to replace current treatment "A" by its newer version and current treatment "B" by its older version.

CHOICE SET 12

Scenario 1	Scenario 2	Scenario 3
Gain 4 years of life in good health	**Neither lose, nor gain years of life in good health**	**Neither lose, nor gain years of life in good health**
Decrease medical expenditures by $120,000	**Decrease medical expenditures by $60,000**	**Neither increase, not decrease medical expenditures**

Which scenario you think is the WORST?

Which scenario you think is the BEST?

PAGE 17

- We remind you that the 3 scenarios describe potential consequences of the two decisions: to replace current treatment "A" by its newer version and current treatment "B" by its older version.

CHOICE SET 13

Scenario 1	Scenario 2	Scenario 3
Neither lose, nor gain years of life in good health	**Lose 8 years of life in good health**	**Neither lose, nor gain years of life in good health**
Increase medical expenditures by $60,000	**Neither increase, not decrease medical expenditures**	**Neither increase, not decrease medical expenditures**

Which scenario you think is the WORST?

Which scenario you think is the BEST?

PAGE 18

- We remind you that the 3 scenarios describe potential consequences of the two decisions: to replace current treatment "A" by its newer version and current treatment "B" by its older version.

CHOICE SET 14

Scenario 1	Scenario 2	Scenario 3
Lose 8 years of life in good health	**Lose 4 years of life in good health**	**Neither lose, nor gain years of life in good health**
Decrease medical expenditures by $120,000	**Decrease medical expenditures by $60,000**	**Neither increase, not decrease medical expenditures**

Which scenario you think is the WORST?

Which scenario you think is the BEST?

PAGE 19

- We remind you that the 3 scenarios describe potential consequences of the two decisions: to replace current treatment "A" by its newer version and current treatment "B" by its older version.

CHOICE SET 15

Scenario 1	Scenario 2	Scenario 3
Lose 8 years of life in good health	**Gain 8 years of life in good health**	**Neither lose, nor gain years of life in good health**
Increase medical expenditures by $120,000	**Decrease medical expenditures by $120,000**	**Neither increase, not decrease medical expenditures**

Which scenario you think is the WORST?

Which scenario you think is the BEST?

PAGE 20

The following questions will help us to understand what you have learned from this study about the funding of new medical interventions. Please answer each question by RIGHT or WRONG.

		WRONG	RIGHT
1.	The healthcare budget is unlimited	☐	☐
2.	One can increase medical expenditures without raising income tax, cutting down other public expenditures or asking patients/users to pay more	☐	☐
3.	It is impossible to replace current treatments by their older versions in order to free some resources for other purposes	☐	☐
4.	The two decisions to replace treatments consist of replacing a current treatment “B” by its newer version and another current treatment “B” by its older version	☐	☐
5.	The two decisions to replace treatments has consequences for different groups of patients, say patients with disease« A » and patients with disease « B »	☐	☐
6.	Health policy decision-makers are not interested in improving he overall health of the population (all patients confounded)	☐	☐
7.	The consequences of the decision to replace treatments cannot be known for sure in advance and thus can lead to unexpected results such as decrease in overall health of the population	☐	☐
8.	Health policy decision-makers take into account the uncertainty in the consequences of their decisions by anticipating different scenarios they might face at the end of the day	☐	☐
9.	The consequences of the decision to replace treatments can be mainly described in terms of both gains/losses in the numbers of years of life in good health and increases/decreases in the level of medical expenditures	☐	☐

In taking part to this study, would you say that your knowledge about funding of new medical interventions has changed?

Not at all										**Extremely**
0	1	2	3	4	5	6	7	8	9	10
☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

PAGE 21

What are your views about the questionnaire?

How would you rate this questionnaire?

Not at all interesting										**Extremely interesting**
0	1	2	3	4	5	6	7	8	9	10
☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

How would you rate the difficulty of this questionnaire?

Extremely difficult										**Extremely easy**
0	1	2	3	4	5	6	7	8	9	10
☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

How would you rate the quality of this questionnaire?

Extremely good										**Extremely bad**
0	1	2	3	4	5	6	7	8	9	10
☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

Do you think the public should be involved in the decisions to fund new medical interventions?

Not at all										**Yes quite**
0	1	2	3	4	5	6	7	8	9	10
☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

How satisfied are you with the healthcare system?

Fully unsatisfied										**Fully satisfied**
0	1	2	3	4	5	6	7	8	9	10
☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

Which statement best describes how you answered the choice questions? (Please tick one)

- ☐ I was only concerned with selecting the smallest level of medical expenditures
- ☐ I was only concerned with selecting the largest number of years of life in good health
- ☐ I tried to find a balance (compromise) between medical expenditures and number of years of life in good health
- ☐ I randomly selected one of the scenarios

What is the minimum number of years of life in good health that you were willing to accept in your decisions? (Please tick one)

- ☐ -8 years (Lose 8 years of life in good health)
- ☐ -4 years (Lose 4 years of life in good health)
- ☐ +0 year (Neither lose, nor gain years of life in good health)
- ☐ +4 years (Gain 4 years of life in good health)
- ☐ +8 years (Gain 8 years of life in good health)

What is the maximum level of medical expenditures that you were willing to accept in your decisions? (Please tick one)

- ☐ -$120,000 (Decrease medical expenditures by $120,000)
- ☐ -$60,000 (Decrease medical expenditures by $60,000)
- ☐ +$0 (Neither decrease nor increase medical expenditures)
- ☐ +$60,000 (Increase medical expenditures by $60,000)
- ☐ +$120,000 (Increase medical expenditures by $120,000)

What importance did you attach to the number of years of life in good health in your decisions? (Please tick one)

- ☐ Extremely important
- ☐ Very important
- ☐ Moderately important
- ☐ Slightly important
- ☐ Not at all important

What importance did you attach to the medical expenditures in your decisions? (Please tick one)

- ☐ Extremely important
- ☐ Very important
- ☐ Moderately important
- ☐ Slightly important
- ☐ Not at all important

PAGE 22

Questions about your attitudes towards funding of medical interventions

These questions will help us to understand your views regarding the funding of medical interventions.
(For each statement please tick the appropriate box)

Statement	I agree	I partially agree	I am uncertain	I partially disagree	I disagree
I am willing to pay more tax to increase the health care budget!					
People should pay more health care services out of their own pocket!					
We need a larger health care budget in the future!					
It is unfair that healthy people pay for the treatment of unhealthy people!					
New medical interventions should be adopted at all costs!					
Life years of some people should not be decrease to increase the life years of others!					
Patients should be more involved in the decision whether a new medical intervention will be adopted!					
The society should be more involved in the decision whether a new medical intervention will be adopted!					
Health care budget could be saved if doctors would stop prescribing too much medicine!					

PAGE 23

Questions about yourself

Age

Gender

How is your health in general? (Please tick one)

- ☐ Very good
- ☐ Good
- ☐ Fair
- ☐ Bad
- ☐ Very bad

W**hat is your highest level of educational qualifications? (Please tick one)**

- ☐ No formal qualifications
- ☐ Secondary/high school qualifications
- ☐ University/College degree
- ☐ Other, please specify

Which group represents your total household income including any benefits received and before any deductions (e.g. tax)? (Please tick one)

- ☐ Up to £5,199 per year
- ☐ £5,200 and up to £10,399 per year
- ☐ £10,400 and up to £15,599 per year
- ☐ £15,600 and up to £20,799 per year
- ☐ £20,800 and up to £25,999 per year
- ☐ £26,000 and up to £31,199 per year
- ☐ £31,200 and up to £36,399 per year
- ☐ £36,400 and up to £51,999 per year
- ☐ £52,000 and above per year
- ☐ Prefer not to say

Do you have any of the following chronic conditions (select all those applicable)?

- ☐ Arthrosis
- ☐ Asthma
- ☐ Cancer
- ☐ Chronic Pulmonary Disease
- ☐ Cognitive disorders (Alzheimer disease, Parkinson disease)
- ☐ Diabetis
- ☐ Heart failure
- ☐ Hypertension
- ☐ Kydney disease
- ☐ Mental illness (for example schizophrenia, depression, etc.)
- ☐ Osteoporosis
- ☐ Viral diseases (for example, hepatitis C, HIV/AIDS)
- ☐ Other(s)
- ☐ Prefer not to say

AUTHOR DECLARATION FORM

At submission, **EVERY AUTHOR** listed in the manuscript must **READ** and **COMPLETE** the following statements on: (A) Authorship Responsibility, (B) Authorship Criteria, (C) Authorship Contribution, (D) Funding Disclosures, (E) Contributor Disclosures/Acknowledgments, and (F) Conflicts of Interest Disclosures.

It is important that you return this form as early as possible in the publication process. **EVERY AUTHOR MUST COMPLETE AN INDIVIDUAL COPY OF THE FORM, AND EVERY SECTION OF THE FORM MUST BE COMPLETED.** We will **NOT** consider your manuscript for publication until every author has completed the form and returned it to us.

Your name (please print): Amiram Gafni E-mail: gafni@mcmaster.ca

Journal name: PharmacoEconomics Corresponding author: Nicolas KRUCIEN

Manuscript title: Measuring public preferences for health outcomes and expenditures in a context of healthcare resources re-allocation

A. AUTHORSHIP RESPONSIBILITY

☒ I certify that **ALL** of the following statements are correct **(PLEASE CHECK THE BOX)**.

- The manuscript represents valid work; neither this manuscript nor one with substantially similar content under my authorship has been published or is being considered for publication elsewhere (except as described in the manuscript submission); and copies of any closely related manuscripts are enclosed in the manuscript submission; ***AND***
- For manuscripts with more than one author, I agree to allow the corresponding author to serve as the primary correspondent with the editorial office and to review and sign off on the final proofs prior to publication; or, if I am the only author, I will be the corresponding author and agree to serve in the roles described above.
- For manuscripts that are a report of a study, I confirm that this work is an accurate representation of the trial results.

B. AUTHORSHIP CRITERIA

To fulfil all of the criteria for authorship, every author of the manuscript must have made substantial contributions to **ALL** of the following aspects of the work:

- Conception and planning of the work that led to the manuscript or acquisition, analysis and interpretation of the data, or both; ***AND***
- Drafting and/or critical revision of the manuscript for important intellectual content; ***AND***
- Approval of the final submitted version of the manuscript.

☒ I certify that I fulfill **ALL** of the above criteria for authorship **(PLEASE CHECK THE BOX)**.

C. AUTHORSHIP CONTRIBUTION

I certify that I have participated sufficiently in the work to take public responsibility for **(PLEASE CHECK 1 OF THE 2 BOXES BELOW)**:

☐ Part of the content of the manuscript; ***OR***
☒ The entire content of the manuscript.

D. FUNDING DISCLOSURES

PLEASE CHECK 1 OF THE 2 BOXES BELOW:

☐ I certify that no funding has been received for the conduct of this study and/or preparation of this manuscript; ***OR***
☒ I certify that all financial and material support for the conduct of this study and/or preparation of this manuscript is clearly described in the Compliance with Ethical Standards section of the manuscript.

Some funding organizations require that authors of manuscripts reporting research deposit those manuscripts with an approved public repository.

☐ Please check here if you have received such funding.

E. CONTRIBUTOR DISCLOSURES

All persons who have made substantial contributions to the work reported in the manuscript (e.g. data collection, data analysis, or writing or editing assistance) but who do not fulfill the authorship criteria **MUST** be named with their specific contributions in the Acknowledgments section of the manuscript. Groups of persons who have contributed may be listed under a heading such as 'Clinical investigators' and their function described. Because readers may infer their endorsement of the manuscript, all persons named in the Acknowledgments section **MUST** give the authors their written permission to be named in the manuscript.

☒ I certify that all persons who have made substantial contributions to this manuscript but who do not fulfill the authorship criteria are listed with their specific contributions in the Acknowledgments section in the manuscript, and that all persons named in the Acknowledgments section have given me written permission to be named in the manuscript.

F. CONFLICT OF INTEREST DISCLOSURES

A conflict of interest exists when professional judgment concerning a primary interest (such as patients' welfare or the validity of research) may be influenced by a secondary interest (such as financial gain or personal rivalry). A conflict of interest may arise for authors when they have a financial interest that may influence – probably without their knowing – their interpretation of their results or those of others. We believe that to make the best decision on how to deal with a manuscript we should know about any such conflict of interest that the authors may have. We are not aiming to eradicate conflicts of interests – they are almost inevitable. We will not reject manuscripts simply because the authors have a conflict of interest, but we will publish a declaration in the manuscript as to whether or not the authors have conflicts of interests.

All authors **MUST** complete the following checklist:

Category of potential conflict of interest	If you have had any of the listed relationships with an entity that has a financial interest in the subject matter discussed in this manuscript, please check the appropriate "Yes" box below and provide details. If you do not have a listed relationship, please check the appropriate "No" box. When completing this section, please take into account the last 36 months through to the foreseeable future.		
	No (√)	**Yes (√)**	**Details**
Employment	√		
Grant received/grants pending	√		
Consulting fees or honorarium	√		
Support for travel to meetings for the study, manuscript preparation or other purposes	√		
Fees for participation in review activities such as data monitoring boards, etc	√		
Payment for writing or reviewing the manuscript	√		
Provision of writing assistance, medicines, equipment or administrative support	√		
Payment for lectures including service on speakers bureaus	√		
Stock/stock options	√		
Expert testimony	√		
Patents (planned, pending or issued)	√		
Royalties	√		
Other (err on the side of full disclosure)	√		

Every author **MUST** complete option 1 or option 2 as appropriate below. If you answered "Yes" to any of the questions relating to financial conflicts of interests in the table above (or if you wish to disclose a non-financial conflict of interest), you **MUST** write a suitable statement in the box below and include this statement in the Compliance with Ethical Standards section of the manuscript.

☒ I have no conflicts of interest to declare; ***OR***

☐ The following statement regarding conflicts of interest and financial support for conduct of this study and/or preparation of this manuscript is to be published in the Compliance with Ethical Standards section of the manuscript:

Declaration: I certify that I have fully read and fully understood this form, and that the information that I have presented here is accurate and complete to the best of my knowledge.

Your name (please print): AMIRAM GAFNI

Signature (please **HAND-WRITE**): 

Date: 05/06/2018

AUTHOR DECLARATION FORM

At submission, **EVERY AUTHOR** listed in the manuscript must **READ** and **COMPLETE** the following statements on: (A) Authorship Responsibility, (B) Authorship Criteria, (C) Authorship Contribution, (D) Funding Disclosures, (E) Contributor Disclosures/Acknowledgments, and (F) Conflicts of Interest Disclosures.

It is important that you return this form as early as possible in the publication process. **EVERY AUTHOR MUST COMPLETE AN INDIVIDUAL COPY OF THE FORM, AND EVERY SECTION OF THE FORM MUST BE COMPLETED.** We will **NOT** consider your manuscript for publication until every author has completed the form and returned it to us.

Your name (please print): Nicolas KRUCIEN E-mail: nicolas.krucien@abdn.ac.uk

Journal name: PharmacoEconomics Corresponding author: Nicolas KRUCIEN

Manuscript title: Measuring public preferences for health outcomes and expenditures in a context of healthcare resources re-allocation

A. AUTHORSHIP RESPONSIBILITY

☒ I certify that **ALL** of the following statements are correct **(PLEASE CHECK THE BOX)**.

- The manuscript represents valid work; neither this manuscript nor one with substantially similar content under my authorship has been published or is being considered for publication elsewhere (except as described in the manuscript submission); and copies of any closely related manuscripts are enclosed in the manuscript submission; ***AND***
- For manuscripts with more than one author, I agree to allow the corresponding author to serve as the primary correspondent with the editorial office and to review and sign off on the final proofs prior to publication; or, if I am the only author, I will be the corresponding author and agree to serve in the roles described above.
- For manuscripts that are a report of a study, I confirm that this work is an accurate representation of the trial results.

B. AUTHORSHIP CRITERIA

To fulfil all of the criteria for authorship, every author of the manuscript must have made substantial contributions to **ALL** of the following aspects of the work:

- Conception and planning of the work that led to the manuscript or acquisition, analysis and interpretation of the data, or both; ***AND***
- Drafting and/or critical revision of the manuscript for important intellectual content; ***AND***
- Approval of the final submitted version of the manuscript.

☒ I certify that I fulfill **ALL** of the above criteria for authorship **(PLEASE CHECK THE BOX)**.

C. AUTHORSHIP CONTRIBUTION

I certify that I have participated sufficiently in the work to take public responsibility for **(PLEASE CHECK 1 OF THE 2 BOXES BELOW)**:

☐ Part of the content of the manuscript; ***OR***
☒ The entire content of the manuscript.

D. FUNDING DISCLOSURES

PLEASE CHECK 1 OF THE 2 BOXES BELOW:

☐ I certify that no funding has been received for the conduct of this study and/or preparation of this manuscript; ***OR***
☒ I certify that all financial and material support for the conduct of this study and/or preparation of this manuscript is clearly described in the Compliance with Ethical Standards section of the manuscript.

Some funding organizations require that authors of manuscripts reporting research deposit those manuscripts with an approved public repository.

☐ Please check here if you have received such funding.

E. CONTRIBUTOR DISCLOSURES

All persons who have made substantial contributions to the work reported in the manuscript (e.g. data collection, data analysis, or writing or editing assistance) but who do not fulfill the authorship criteria **MUST** be named with their specific contributions in the Acknowledgments section of the manuscript. Groups of persons who have contributed may be listed under a heading such as 'Clinical investigators' and their function described. Because readers may infer their endorsement of the manuscript, all persons named in the Acknowledgments section **MUST** give the authors their written permission to be named in the manuscript.

☒ I certify that all persons who have made substantial contributions to this manuscript but who do not fulfill the authorship criteria are listed with their specific contributions in the Acknowledgments section in the manuscript, and that all persons named in the Acknowledgments section have given me written permission to be named in the manuscript.

F. CONFLICT OF INTEREST DISCLOSURES

A conflict of interest exists when professional judgment concerning a primary interest (such as patients' welfare or the validity of research) may be influenced by a secondary interest (such as financial gain or personal rivalry). A conflict of interest may arise for authors when they have a financial interest that may influence – probably without their knowing – their interpretation of their results or those of others. We believe that to make the best decision on how to deal with a manuscript we should know about any such conflict of interest that the authors may have. We are not aiming to eradicate conflicts of interests – they are almost inevitable. We will not reject manuscripts simply because the authors have a conflict of interest, but we will publish a declaration in the manuscript as to whether or not the authors have conflicts of interests.

All authors **MUST** complete the following checklist:

Category of potential conflict of interest	If you have had any of the listed relationships with an entity that has a financial interest in the subject matter discussed in this manuscript, please check the appropriate "Yes" box below and provide details. If you do not have a listed relationship, please check the appropriate "No" box. When completing this section, please take into account the last 36 months through to the foreseeable future.		
	No (√)	**Yes (√)**	**Details**
Employment	√		
Grant received/grants pending	√		
Consulting fees or honorarium	√		
Support for travel to meetings for the study, manuscript preparation or other purposes	√		
Fees for participation in review activities such as data monitoring boards, etc	√		
Payment for writing or reviewing the manuscript	√		
Provision of writing assistance, medicines, equipment or administrative support	√		
Payment for lectures including service on speakers bureaus	√		
Stock/stock options	√		
Expert testimony	√		
Patents (planned, pending or issued)	√		
Royalties	√		
Other (err on the side of full disclosure)	√		

Every author **MUST** complete option 1 or option 2 as appropriate below. If you answered "Yes" to any of the questions relating to financial conflicts of interests in the table above (or if you wish to disclose a non-financial conflict of interest), you **MUST** write a suitable statement in the box below and include this statement in the Compliance with Ethical Standards section of the manuscript.

☒ I have no conflicts of interest to declare; ***OR***

☐ The following statement regarding conflicts of interest and financial support for conduct of this study and/or preparation of this manuscript is to be published in the Compliance with Ethical Standards section of the manuscript:

Declaration: I certify that I have fully read and fully understood this form, and that the information that I have presented here is accurate and complete to the best of my knowledge.

Your name (please print): Nicolas KRUCIEN

Signature (please **HAND-WRITE**):

Date: 05/06/2018

AUTHOR DECLARATION FORM

At submission, **EVERY AUTHOR** listed in the manuscript must **READ** and **COMPLETE** the following statements on: (A) Authorship Responsibility, (B) Authorship Criteria, (C) Authorship Contribution, (D) Funding Disclosures, (E) Contributor Disclosures/Acknowledgments, and (F) Conflicts of Interest Disclosures.

It is important that you return this form as early as possible in the publication process. **EVERY AUTHOR MUST COMPLETE AN INDIVIDUAL COPY OF THE FORM, AND EVERY SECTION OF THE FORM MUST BE COMPLETED.** We will **NOT** consider your manuscript for publication until every author has completed the form and returned it to us.

Your name (please print): Nathalie PELLETIER-FLEURY E-mail: nathalie.pelletier-fleury@inserm.fr

Journal name: PharmacoEconomics Corresponding author: Nicolas KRUCIEN

Manuscript title: Measuring public preferences for health outcomes and expenditures in a context of healthcare resources re-allocation

A. AUTHORSHIP RESPONSIBILITY

☒ I certify that **ALL** of the following statements are correct **(PLEASE CHECK THE BOX)**.

- The manuscript represents valid work; neither this manuscript nor one with substantially similar content under my authorship has been published or is being considered for publication elsewhere (except as described in the manuscript submission); and copies of any closely related manuscripts are enclosed in the manuscript submission; ***AND***
- For manuscripts with more than one author, I agree to allow the corresponding author to serve as the primary correspondent with the editorial office and to review and sign off on the final proofs prior to publication; or, if I am the only author, I will be the corresponding author and agree to serve in the roles described above.
- For manuscripts that are a report of a study, I confirm that this work is an accurate representation of the trial results.

B. AUTHORSHIP CRITERIA

To fulfil all of the criteria for authorship, every author of the manuscript must have made substantial contributions to **ALL** of the following aspects of the work:

- Conception and planning of the work that led to the manuscript or acquisition, analysis and interpretation of the data, or both; ***AND***
- Drafting and/or critical revision of the manuscript for important intellectual content; ***AND***
- Approval of the final submitted version of the manuscript.

☒ I certify that I fulfill **ALL** of the above criteria for authorship **(PLEASE CHECK THE BOX)**.

C. AUTHORSHIP CONTRIBUTION

I certify that I have participated sufficiently in the work to take public responsibility for **(PLEASE CHECK 1 OF THE 2 BOXES BELOW)**:

☐ Part of the content of the manuscript; ***OR***
☒ The entire content of the manuscript.

D. FUNDING DISCLOSURES

PLEASE CHECK 1 OF THE 2 BOXES BELOW:

☐ I certify that no funding has been received for the conduct of this study and/or preparation of this manuscript; ***OR***
☒ I certify that all financial and material support for the conduct of this study and/or preparation of this manuscript is clearly described in the Compliance with Ethical Standards section of the manuscript.

Some funding organizations require that authors of manuscripts reporting research deposit those manuscripts with an approved public repository.

☐ Please check here if you have received such funding.

E. CONTRIBUTOR DISCLOSURES

All persons who have made substantial contributions to the work reported in the manuscript (e.g. data collection, data analysis, or writing or editing assistance) but who do not fulfill the authorship criteria **MUST** be named with their specific contributions in the Acknowledgments section of the manuscript. Groups of persons who have contributed may be listed under a heading such as 'Clinical investigators' and their function described. Because readers may infer their endorsement of the manuscript, all persons named in the Acknowledgments section **MUST** give the authors their written permission to be named in the manuscript.

☒ I certify that all persons who have made substantial contributions to this manuscript but who do not fulfill the authorship criteria are listed with their specific contributions in the Acknowledgments section in the manuscript, and that all persons named in the Acknowledgments section have given me written permission to be named in the manuscript.

F. CONFLICT OF INTEREST DISCLOSURES

A conflict of interest exists when professional judgment concerning a primary interest (such as patients' welfare or the validity of research) may be influenced by a secondary interest (such as financial gain or personal rivalry). A conflict of interest may arise for authors when they have a financial interest that may influence – probably without their knowing – their interpretation of their results or those of others. We believe that to make the best decision on how to deal with a manuscript we should know about any such conflict of interest that the authors may have. We are not aiming to eradicate conflicts of interests – they are almost inevitable. We will not reject manuscripts simply because the authors have a conflict of interest, but we will publish a declaration in the manuscript as to whether or not the authors have conflicts of interests.

All authors **MUST** complete the following checklist:

Category of potential conflict of interest	If you have had any of the listed relationships with an entity that has a financial interest in the subject matter discussed in this manuscript, please check the appropriate "Yes" box below and provide details. If you do not have a listed relationship, please check the appropriate "No" box. When completing this section, please take into account the last 36 months through to the foreseeable future.		
	No (√)	**Yes (√)**	**Details**
Employment	√		
Grant received/grants pending	√		
Consulting fees or honorarium	√		
Support for travel to meetings for the study, manuscript preparation or other purposes	√		
Fees for participation in review activities such as data monitoring boards, etc	√		
Payment for writing or reviewing the manuscript	√		
Provision of writing assistance, medicines, equipment or administrative support	√		
Payment for lectures including service on speakers bureaus	√		
Stock/stock options	√		
Expert testimony	√		
Patents (planned, pending or issued)	√		
Royalties	√		
Other (err on the side of full disclosure)	√		

Every author **MUST** complete option 1 or option 2 as appropriate below. If you answered "Yes" to any of the questions relating to financial conflicts of interests in the table above (or if you wish to disclose a non-financial conflict of interest), you **MUST** write a suitable statement in the box below and include this statement in the Compliance with Ethical Standards section of the manuscript.

☒ I have no conflicts of interest to declare; ***OR***

☐ The following statement regarding conflicts of interest and financial support for conduct of this study and/or preparation of this manuscript is to be published in the Compliance with Ethical Standards section of the manuscript:

Declaration: I certify that I have fully read and fully understood this form, and that the information that I have presented here is accurate and complete to the best of my knowledge.

Your name (please print): Nathalie PELLETIER-FLEURY

Signature (please **HAND-WRITE**):

Date: 05/06/2018

FULL TITLE

Measuring public preferences for health outcomes and expenditures in a context of healthcare resources re-allocation

SHORT TITLE

Public preferences for health outcomes and expenditures

AUTHORS

Nicolas KRUCIEN, PhD

Health Economics Research Unit, Institute of Applied Health Sciences, University of Aberdeen, Aberdeen (AB25 2QN), United Kingdom; nicolas.krucien@abdn.ac.uk

Nathalie PELLETIER-FLEURY, PhD, MD

Centre de Recherche en Epidémiologie et Santé des Populations, Université Paris-Sud, UVSQ, INSERM, Université Paris-Saclay, Villejuif, France ; nathalie.pelletier-fleury@inserm.fr

Amiram GAFNI, PhD

Centre for Health Economics and Policy Analysis, Department of Health Research Methods, Evaluation and Impact, McMaster University, Hamilton, Canada; gafni@mcmaster.ca

CORRESPONDING AUTHOR

Nicolas KRUCIEN, PhD
Health Economics Research Unit
Institute of Applied Health Sciences
University of Aberdeen
Aberdeen, AB25 2QN
Tel: +44(0)1-224-437-892
Fax: +44(0)1-224-437-195
Email: nicolas.krucien@abdn.ac.uk

ACKNOWLEDGMENTS

We thank all participants who took part in the study.

Compliance with ethical standards

Financial support for this study was provided by the French National Institute of Health and Medical Research (INSERM). The funding agreement ensured the authors' independence in designing the study, interpreting the data, writing, and publishing the report. The authors (Nicolas KRUCIEN; Nathalie PELLETIER-FLEURY; Amiram GAFNI) have no conflict of interest to declare.

ABSTRACT

Backgound: The final outcome of any resource allocation decision in healthcare cannot be determined in advance. Thus decision makers, in deciding which new programme to implement (or not), need to accommodate the uncertainty of different potential outcomes (i.e., change in both health and costs) that can occur, the size and nature (i.e., "bad" or "good") of these outcomes and how they are being valued. Using the decision making plane (DMP), which explicitly incorporates opportunity costs and relaxes the assumptions of perfect divisibility and constant returns to scale of the cost-effectiveness plane, all the potential outcomes of each resource allocation decision can be described.

Objective: In this study we describe the development and testing of an instrument, using a discrete choice experiment methodology, allowing the measurement of public preferences for potential outcomes falling in different quadrants of the DMP.

Method: In a sample of 200 participants providing 4,200 observations we compared four versions of the preferences-elicitation instrument using a range of indicators.

Results: We identified one version that was well accepted by the participants and with good measurement properties.

Conclusion:This validated instrument can now be used in a larger representative sample to study the preferences of the public for potential outcomes stemming from re-allocation of healthcare resources.

Key points for decision makers

1. A validated preferences-elicitation instrument was developed that can be used in a representative sample of the general population
2. This study found preliminary evidence of non-linearities in public valuation of outcomes stemming from reallocation of healthcare resources
3. This study describes a comprehensive methodology for testing and comparing the properties of preferences-elicitation instruments, which can be applied to other dicrete choice experiments

1: INTRODUCTION

If healthcare budgets were unlimited, all most effective treatments could be adopted. However, resources allocated to healthcare are scarce, hence health policy decision-makers (HPDMs) need to decide how to best allocate them. Resources scarcity can occur in different contexts (i.e., fixed budgets, shrinking budget with less resources allocated to healthcare, growing budget with more resources allocated to healthcare) as long as the total amount of resources available is not sufficient to support the implementation of all the most effective treatments. As a result of scarcity, HPDMs need to determine where the resources should come from to fund the implementation of new treatments to replace or complement existing treatments. For example, in the case of a fixed budget, HPDMs may decide to cancel existing treatment(s) in order to free up resources to implement the new treatment(s). In the case of a growing budget, because not all new treatments can be implemented, HPDMs would still need to decide which new treatment(s) to implement and which ones to abandon.

This resources allocation decision (RAD) is challenging because it typically requires to trade off potential health gains for patients who will benefit from the new treatment against potential health losses for those who will see their current treatment being cancelled or replaced (or potential new treatment not adopted). In this context, HPDMs need to consider the opportunity costs of their decisions to "*ensure that the value of what is gained from an activity* [e.g., implementing the new treatment] *outweighs the value of what has to be sacrificed* [e.g., cancelling an existing treatment]"[1].

Cost–Effectiveness Analysis (CEA) is widely advocated as a tool to help HPDMs to allocate the resources available in a way that maximizes the health benefits produced to the population. The analytical tool of CEA is the incremental cost-effectiveness ratio (ICER), which is then compared with a threshold ICER to determine whether the new treatment should be implemented. Assuming that healthcare resources are efficiently used, this threshold ICER should in principle correspond to the ICER of the last treatment adopted [2] and would then correspond to the shadow price of the budget constraint. However, it has been shown that this approach would lead to an optimal use of healthcare resources only under the strong assumptions of perfect divisibility and constant returns to scale of all treatments [3]–[5]. Birch & Gafni (B&G) have suggested an alternative approach relaxing these two questionable assumptions [2],[3]. Other studies also questioned the validity of the "*ICER of the last treatment adopted*" as threshold to guide the RAD [6]. Eckerman & Pekarsky showed that the shadow price is impacted by different factors such as type of financing (i.e., expansion of healthcare budget vs. displacement of existing resources) and whether existing resources were optimally allocated or not. For instance, when funding is done by expansion of the healthcare budget in an economically efficient system, the shadow price should correspond to the ICER of the "best" (i.e., most cost-effective) existing programme. When the funding is done by displacement of existing resources in an economically inefficient system, the shadow price should depend on the ICERs of the "best" programme, of the "worst" (i.e., least cost-effective) programme and of the displaced programme [6]. The B&G approach is based on the key concept of opportunity cost. The differences are: (i) It does not require the use of underlying unrealistic assumptions and their consequences [3], and (ii) it does not require the use of an ICER and a ICER threshold. The B&G approach identifies the source of the additional resource requirements of the new program and makes recommendation regarding the adoption of the new program based on a direct comparison of the total additional benefits produced from the new program with the total benefits forgone. In doing so it ensures, that if followed and under conditions of certainty, "*the value of what is gained from an activity outweighs the value of what has to be sacrificed"* [1].

In the past two decades it has been recognized that both costs and effects of all programs are stochastic, and then the B&G approach has been extended to account for the uncertainty in costs and effects of re-allocating resources [7], [8]. Visually, it takes the form of a Decision Making Plane (DMP) allowing to describe all the possible outcomes stemming from resource reallocation due to the uncertainty [7]. The cost-effectiveness plane (CEP) describes only the difference in health outcomes (E) and costs (C) of a candidate treatment for implementation (A1) with a reference one (A0) using measures of incremental effectiveness ($\Delta E_A = E_{A1} - E_{A0}$) and increment

costs ($\Delta C_A = C_{A1} - C_{A0}$). Those measures are used to compute the ICER and compare it to the ICER threshold. The DMP "extends" the CEP by also comparing a candidate treatment(s) for cancellation (B1) with another reference treatment(s) (B0)[1] (i.e., the explicit consideration of the source of additional resources), leading thus to another set of incremental effects ($\Delta E_B = E_{B1} - E_{B0}$) and costs ($\Delta C_B = C_{B1} - C_{B0}$). All these incremental measures are used to compute net changes in health outcomes ($\Delta E = \Delta E_A - \Delta E_B$) and costs ($\Delta C = \Delta C_A - \Delta C_B$) which are then mapped into the DMP (**Figure 1A**). The DMP is divided into four quadrants which will affect the RAD. Quadrant I (QI) describes situations where the joint decision to replace A0 by A1 and B1 by B0 allows improving the population health (i.e., $\Delta E > 0$) for an overall lower level of medical expenditures (i.e., $\Delta C < 0$). At the opposite, quadrant III (Q_{III}) describes situations where population health is decreased (i.e., $\Delta E < 0$) and medical expenditures increased (i.e., $\Delta C > 0$). Quadrant II (Q_{II}) describes situations where both the population health and level of medical expenditures are decreased (i.e., $\Delta E < 0$; $\Delta C < 0$). Quadrant IV (Q_{IV}) describes situations where both the population health and level of medical expenditures are increased (i.e., $\Delta E > 0$; $\Delta C > 0$).

In terms of health policy decision-making, the decision to replace existing treatments in order to free up resources for the implementation of a new treatment should be made, ideally, only if the final outcome will be located in QI. However, this cannot be guaranteed because, as explained, net changes in population health and medical expenditures are uncertain, such that RAD becomes a risky decision. Each proposed way of allocating healthcare resources might have a non-null probability to end up in each of the four quadrants of the DMP. This uncertainty can be represented by a joint distribution of net changes over the DMP (**Figure 1B**). Given this element of risk, the RAD will depend not only on the probabilities of falling in the four DMP quadrants, but also on how HPDMs value each possible situations. It would be too restrictive to assume that HPDMs view all potential situations as being equally desirable [8]. The valuation of each (ΔE; ΔC) situation is likely to depend on the specific quadrant that it falls in and the exact location within the quadrant. Assuming, for example, that HPDMs positively value an improvement in population health and a decrease in the level of medical expenditures, a situation falling in Q_{III} should be perceived as "bad" (i.e., to have a negative value), and likewise a Q_I situation should be seen as "good" (i.e., to have a positive value). But this description of HPDMs' preferences for changes in population health and medical expenditures remains largely incomplete. It is unknown whether: (i) HPDMs would be more concerned by a "bad" situation rather than a comparable (same-size) "good" situation; (ii) all situations falling in Q_I (Q_{III}) should be seen as equally "good" ("bad"). Also (ΔE; ΔC) situations can fall in quadrants II and IV where one outcome is "good" and the other is "bad". The answers to these questions depend on both the sign and size of HPDMs' preferences for net changes in population health and medical expenditures. To the best of our knowledge such valuation function, that describes the preferences (or value attributed) for every potential outcome in each quadrant of the DMP, does not exist. In order to measure such function one first needs to develop and validate a tool which will provide reliable measures of preferences for net changes in population health and medical expenditures. This is the objective of this study. In this study we report the development and testing of a preference-elicitation instrument (PEI) which can be used to measure preferences for changes in population health and medical expenditures in a context of resources scarcity.

2: METHODS

2.1: Developing a preference elicitation instrument (PEI)

2.1.1: Choice experiment

The discrete choice experiment (DCE) methodology was used to measure preferences for net changes in health outcomes and costs. DCEs are commonly used in health for eliciting preferences for a wide range of policy

[1] The DMP can also be extended to the case where more than one existing treatments have to be replaced in order to free up resources for the implement of the new treatment.

questions [9], [10]. We use the concept of healthy year equivalent (HYE)[2] [11] to describe net changes in health outcomes (ΔE) and amount of Euros for net changes in the medical expenditures (ΔC). In our study, we used a modified DCE format known as best-worst scaling (BWS) case III. Unlike the standard DCE approach which only asks participants to identified their most preferred choice option (i.e., BEST choice), the BWS approach also asks them to identify their least preferred option (i.e., WORST choice). This approach allows for a full rank ordering of the situations and then provide more information about individuals' preferences for the same number of choice tasks [12].

Whilst there are studies in the DCE literature explaining how to identify relevant attributes, to the best of our knowledge there is no comparable evidence regarding the selection of attributes' levels. Furthermore, methodological research on the designing of DCEs showed that individuals' preferences were not invariant to changes in the range of attributes' levels [13]. Unfortunately this issue has been overlooked in the DCE literature and there is no validated approach to identify the "best set" of attributes' levels. In our study we addressed this issue by testing and comparing four different versions of the BW-DCE questionnaire that only differ in terms of attributes' levels: The ranges of possible values for the ΔE attribute were {-4; -2; 0; +2; +4} and {-8; -4; 0; +4; +8} for versions 1 & 2 ($V_{1\text{-}2}$) and versions 3 & 4 ($V_{3\text{-}4}$) respectively; The ranges of possible values for the ΔC attribute were {-60,000; -30,000; 0; +30,000; +60,000} and {-120,000; -60,000; 0; +60,000; +120,000} for V_1 and $V_{2\text{-}4}$ respectively. All four versions were based on a D-Efficient design [14], [15] allowing for the estimation of all main effects and one continuous interaction effect between ΔE and ΔC. However, we used non-informative (i.e., null) priors about participants' preferences to generate the list of choice tasks for $V_{1\text{-}3}$ and used results from V_3 as informative (i.e., non-null) priors to design the tasks for V_4. This last version also included 12 experimental tasks (instead of 10) to allow for the estimation of two alternative-specific constants in addition to the other effects[3].

In every choice task we included three generic situations (i.e., scenario 1; scenario 2; scenario 3) (**Figure 2**) to reflect the uncertainty in the consequences of reallocating healthcare resources. In the information sheet of the questionnaire, the participants were told that deciding to reallocate resources would have uncertain consequences which are represented by the different scenarios. One of these generic situations was designed to correspond to the origin point of the DMP, hereafter *neutral changes situation* (NCS). The specification of the BW-DCE was completed by manually designing two quality checks. For the estimation of individuals' preferences, only answers to the experimentally designed choice tasks were considered. The order of the choice tasks within the questionnaire and the order of the alternatives within the choice tasks were randomised across participants to control for potential order effects (e.g., left-to-right, learning/fatigue).

2.1.2: Sampling

The development of our PEI took place in France. In 2016, we contacted a market research company to recruit a total of 200 participants from the general population. V_4 of the instrument was administered two months after the first three versions[4]. Participants were randomly allocated to V_1-V_3. We used the same recruitment procedure for

[2] Those who prefer to use quality-adjusted life years (QALYs) as a measure of health outcome can use the methodology described in this paper but will need to change the description of the health outcome in the instrument.

[3] We used the same experimental design for V1-3 because we specified null preferences for the ΔE and ΔC attributes, making thus the D-efficiency measure insensitive to changes in the magnitude only of the attributes' levels. The purpose of V4 was to investigate whether a "better" (i.e., statistically more efficient) design would allow building a better PEI. The gain in statistical efficiency was obtained by relaxing the assumption of null preferences for ΔE and ΔC, using V3 as non-null priors for the designing of V4.

[4] The V4 was administered two months after the three other versions because we first needed to analyse data obtained from V3 before being able to improve the statistical efficiency of the V4 design (by using V3 results as non-null priors).

all four versions. Following Louviere et al (2010) [16] formulae[5], we needed to recruit a minimum of 44 participants per version (rounded up at 50).

2.2: Empirical testing of the preferences-elicitation instrument

As there is no validated approach regarding the selection of the "best set" of attributes' levels, we used different criteria to compare the four DCE versions in terms of statistical performance (predictive validity), behavioural realism (properties of participants' choices) and acceptability by the participants. This multi-criteria analysis should increase our chance to identify the best PEI.

2.2.1: Debriefing questions

The participants were asked to rate the interest in and difficulty of the questionnaire on a 5-points scale. Then we asked them to answer questions related to how they made their choices: (i) Decision objective (i.e., random choice; decision to minimise ΔC; decision to maximise ΔE; to find a compromise between ΔE and ΔC); (ii) Minimum acceptable level of ΔE, and maximum acceptable level of ΔC; (iii) Importance of ΔE and ΔC. Differences between the four choice experiments were investigated using Chi-2 tests.

2.2.2: Properties of participants' choices

In addition to the experimental choice tasks, we also included two additional tasks to control for the quality of participants' choices. One task was used to check the monotonicity of participants' choices. In the monotonicity task, one option was the best and one was the worst in terms of both ΔE and ΔC. Participants' were expected to choose the best option as "most preferred" and the worst one as "least preferred". The second quality check consisted was a stability task. We tested the stability of choices by repeating task #2 as 2nd last task. Participants were expected to pass the stability test when at least one of their choices was repeated.

We also defined two other quality measures based on serial non-participation (i.e., participants who systematically select either the left, middle or right-located option) and response time (RT). A participant was classified as serial non-participant when s/he selected the NCS situation as BEST in more than 80% of the choice tasks. We recorded the RTs at the task level to identify "speedsters" (i.e., participants who tended to answer the choice tasks "too quickly"). A "quick decision" was defined as a choice with a RT falling in the 1st quintile of the corresponding RT distribution. A participant was considered as a "speedster" when s/he made *quick* decisions in at least 80% of the choice tasks. We compared the proportions of participants who pass/fail the quality checks using Chi-2 tests.

2.2.3: Behavioural realism

Multi-attribute choices are typically analysed using the random utility maximisation (RUM) framework [17], [18].

The "base case" model can be written:

$$U_{ntj} = \beta_1 \Delta E_{ntj} + \beta_2 \Delta C_{ntj} + \varepsilon_{ntj} \quad \textbf{(Eq. 1)}$$

Where U_{ntj} corresponds to the utility (U) derived by respondent (n) at task (t) from the option (j), ε_{ntj} are modelling errors, and the (β_1, β_2) parameters capture the main effect of a 1-unit change in ΔE attribute (i.e., +1 HYE) and in ΔC (i.e., +10,000 euros) on participants' choices respectively. We expect to find positive preferences for ΔE, meaning that on average participants would positively respond to better health outcomes, and negative preferences for ΔC, meaning that participants would negatively respond to increase in level of medical expenditures.

However, as suggested by Gafni et al., there is *a priori* no guarantee that this reference specification provides the best account of participants' decisions [8]. For each version of the choice experiment, we estimated 32 different

[5] The formulae is for choice proportions and it allows testing whether observed proportions significantly differ from proportions that would be obtained by chance (In our case, 33% as there is three choice options per task): H_0: proportion = 33%; H_1: proportion ≠ 33%.

specifications allowing for more flexible choice behaviours: (i) Non-linear preferences (in ΔE and/or ΔC); (ii) Interaction effect between ΔE and ΔC; (iii) NCS bias; (iv) Choices inconsistency. As the number of parameters differs across the models, we used the Bayesian Information Criterion (BIC) to identify the best performing specification for each version (i.e., the specification associated with the smallest BIC value).

The more sophisticated specification of the choice model would take the following form:

$$U_{ntj} = \exp(\beta_1 TYPE_{ntj})[\beta_2 NCS_{ntj} + [\beta_3 \Delta E\{\text{Max loss}\} + \beta_4 \Delta E\{\text{Min loss}\} + \beta_5 \Delta E\{\text{Min gain}\} + \beta_6 \Delta E\{\text{Max gain}\}] + [\beta_7 \Delta C\{\text{Max loss}\} + \beta_8 \Delta C\{\text{Min loss}\} + \beta_9 \Delta C\{\text{Min gain}\} + \beta_{10} \Delta C\{\text{Max gain}\}] + \beta_{11}(\Delta E_{njt} \times \Delta C_{ntj})] + \varepsilon_{ntj} \quad \textbf{(Eq. 2)}$$

Where (β_1) captures an effect of the type of choices (i.e., BEST vs. WORST) on the errors variance, (β_2) a preference for the NS above and beyond the preferences for ΔE and ΔC, (β_3-β_6) are parameters capturing the preferences for the different ΔE values relative to a null change, (β_7-β_{10}) are similar parameters for the ΔC attribute, and finally (β_{11}) is a parameter capturing an interaction effect between ΔE and ΔC.

In terms of *behavioural realism* of the choice model, the best version is expected to be the one associated with: (i) Non-linear preferences for at least one attribute; (ii) Non-significant bias towards NCS; (iii) Significant interaction effect between preferences for ΔE and ΔC ([19], [20], [21]).

2.2.4: Predictive validity

The relative performance of choice models across the four versions of the choice experiment can be compared in terms of ability to predict individuals' choices. We use a cross-validation (CV) procedure to determine the level of predictive validity of each model on its corresponding version. The CV procedure consists in randomly splitting the sample into two groups, namely an estimation sample and a validation sample. The observations from the estimation sample are used to estimate the choice model, and the estimates are then used to predict choices observed in validation sample. The predictive validity corresponds to the % of correct matches between predicted and observed choices. Because sample sizes are limited, we proceed to a 75%-25% repartition of the respondents between the estimation and validation samples respectively. The CV procedure was repeated 10,000 times to compute mean score of predictive validity and associated 95% confidence interval.

3: RESULTS

3.1: Samples of respondents

The proportion of men was approximately 50% in all four samples (44%-50%). About a quarter of the respondents reported a less than good health status (20%-30%) and a third declared at least one chronic condition (28%-44%). The samples mainly included respondents with a higher level of education (either University or college degree) (62%-84%). In overall the differences in sample characteristics across the four versions don't reach significance, thus suggesting there is no sample selection bias (i.e., some profiles of participants are not significantly more represented in one group compared to the others).

3.2: Debriefing questions

Results are reported in **Table 1**. The descriptive analysis of debriefing questions indicates that overall the participants considered the questionnaire as being interesting (≈ 70%) and easy to answer (≈ 50%). A majority of participants were willing to trade net changes in health outcomes (ΔE) against net changes in medical expenditures (ΔC). V_4 appeared to be more interesting (+ 20 points) (P = 0.004) than the other ones. This version also has an impact on the individuals' perception of the ΔE and ΔC attributes. In this version, 48% of the participants declare not being willing to accept a net change in health outcomes below +8 HYE (which also corresponds to the maximum value presented to the participants). However we don't find a similar effect for net changes in medical

expenditures with only 16% of the participants not being willing to accept an option offering a level of change below the maximum value.

3.3: Properties of participants' choices

Results are presented in **Table 2**. Regarding the monotonicity of preferences, all four versions of the choice experiment were associated with high levels of performance. However, V_4 appeared to perform better than the other ones (+10-16 points) but this difference did not reach significance ($P = 0.147$). V4 outperformed the other versions in terms of stability ($P < 0.001$); it achieved better performance in terms of serial non-participation and response time (+6-8 points) but the difference did not reach significance ($P = 0.283$ and $P = 0.580$, respectively).

Regarding the results of the RTs analysis, we found a similar pattern of RTs across the four versions of the choice experiment. The 1st choice task (task #1) is associated with significantly longer RTs, and then RTs tend to slightly decrease over the sequence of tasks. Although most of RT differences between the four versions did not reach significance, V_4 appeared to be systematically associated with longer RT at every task. This last result might indicate that participants might have been more engaged in the completion of the choice tasks [22].

3.4: Behavioural realism

Results are presented in **Table 3**[6]. Regarding V_1 and V_3, the best fitting choice model appeared to be a model allowing for linear preferences for both ΔE and ΔC. At the opposite final model for V_2 and V_4 allowed for non-linearities in preferences for ΔE and/or ΔC. In V_1 and V_3, there was evidence of a NCS bias. In all versions but V_4 we found a significant interaction effect between preferences for ΔE and ΔC. Overall the results verify our *a priori* assumptions regarding the nature of respondents' preferences for ΔE and ΔC attributes (i.e., positive effect of gains; negative effect of losses; monotonic preferences for changes in $\Delta E/\Delta C$).

3.5: Predictive validity

Results are presented in **Figure 3**. With a level of predictive validity close to 78%, the version V_4 appeared to perform significantly better than the other ones. This high level of predictive validity indicates that most of participants made choices that can be well explained by the RUM hypothesis, providing thus evidence that participants were actually making trade-offs between the ΔE and ΔC attributes.

4: DISCUSSION

The objective of this study was to develop and test a preferences-elicitation instrument (PEI) that can then be used in large representative sample of the general population to identify the outcome valuation function needed to help the decision makers to decide whether a given distribution of potential situations is acceptable or not, which will affect the decision to reallocate (or not) resources. After having compared four different versions of the PEI in terms of data quality, behavioural realism and predictive validity, version 4 (V_4) was identified as best. To the best of our knowledge, it is the first time that such an instrument is being developed. A copy of the instrument is provided as online supplementary material.

Commented [A1]: Comments from the Editor: Please cite the copy of the final version of the questionnaire in English language (online supplementary material) in the main text.

We find that public preferences were sensitive both to the range of values for medical expenditures and health outcomes, and to the design of the PEI. The two versions based on "extended" sets of attributes' values (i.e., V_3 and V_4) outperformed V_1 which was based on a "narrow" set of values. Increasing the level of statistical efficiency, and thus presumably making the choice tasks more difficult, also had a positive effect on the preferences elicitation. In the DCE literature, it is usually argued that participants would respond to an increase in task difficulty by adopting simplifying decision rules or making more random decisions [23]. In our case making the choice tasks more statistically efficient (and presumably more difficult) was not correlated with an increase in perceived difficulty (as reported by the participants). This result suggests that an increase in the statistical efficiency of the tasks could also have made the choice situations more realistic and more engaging for the

[6] Summary information about all model specifications can be found in online supplementary material.

participants, leading this to better quality data. Overall our study results indicate that *small* changes in the design of the choice tasks can have significant effects on stated preferences. This result is consistent with the hypothesis of ill-defined/malleable preferences [24], [25] following which individuals would not know *a priori* how much they value the different product attributes and therefore would "build their preferences on the fly". Previous studies have investigated the effect of manipulating some experimental features, such as level of statistical efficiency, on respondents' choices [26], [27] and provide mixed evidence.

In the past, studies have already used the DCE methodology to investigate public preferences for the allocation of healthcare resources [28]–[36]. We cannot compare our results to this literature because our study differs in one central feature: the way the resource constraint was described and incorporated. Whilst the previous studies were motivated by the context of resources scarcity (i.e., because resources are limited it becomes important to understand public preferences regarding how these scarce resource should be used), they did not explicitly incorporate the resource constraint in the decision problem (i.e., the question asked). Our study is different because it forces the respondent to deal with the consequences of taking into account the opportunity costs of implementing a new program (e.g., having to make a decision where a programme(s) have to be cancelled to free up resources to implement a new programme and what would be the outcome of such decision). We explained, carefully, to all participants the meaning of the concepts of resources scarcity and opportunity costs and their implications when making a decision about reallocation of resources.

Our study is not exempt from limitations. First, due to sample size limitation we were unable to explore the impact of respondents' characteristics on their preferences for health outcomes and medical expenditures. Erdem & Thompson (2014) used a latent class approach to investigate preferences heterogeneity and found the existence of three different classes of preferences [29]. As this flexible modelling of respondents' choices requires a large amount of data, we plan to repeat this analysis in a bigger sample. Second, the recruitment of the participants was done at two different points in time (i.e., participants to the V_4 were recruited two months after those for V_1-V_3) which might have introduced a potential bias in our comparison. However, this seems unlikely as (i) the recruitment procedure was the same for all four versions, (ii) the main samples socio-demographic characteristics did not significantly differ between the four samples, and (iii) the time lag was relatively short (i.e., only two months). Our instrument measures preferences for two core elements of the decision making process, namely changes in health outcomes and medical expenditures. In real life situations, decision makers are likely to take into account more factors (e.g., profile of the patients who will benefit from the new intervention and those who will lose, etc.). However, a priori we had no guarantee that members of the general population would be willing to make such difficult decisions and then we decided to focus on the core dimensions of resources reallocation. There is no point making the the instrument more complex by including other factors in the decision making process if public already struggles to trade medical expenditures against health outcomes. Building on this work, future studies could further improve the quality of our instrument by including additional factors in the decision making problem. While we had no formal way of controlling whether participants did consider other factors when making their decisions, we tried to prevent/decrease a potential omission bias by explicitly prompting them to only consider the information about health outcomes and medical expenditures.

This validated PEI can now be used in a larger sample of participants to measure their preferences for net changes in health outcomes and medical expenditures. This information about public preferences can then be used in combination with the information described in the DMP to inform the resources allocation decision (RAD). Knowing the joint distribution of net changes in health outcomes and medical expenditures, the preferences information can be used, for example, to compute the expected value of implementing the new treatment. The HPMs may decide to adopt the new intervention only if, for example, the expected value is positive. By following this approach the decision-making process would become more transparent and would allow public to have a say in the management of healthcare system by letting their preferences for population health and medical expenditures influence the RAD.

Data Availability Statement

The datasets generated during and/or analysed during the current study are available from the corresponding author on request.

Author Contributions

All three authors (Nicolas KRUCIEN, Nathalie PELLETIER-FLEURY, Amiram GAFNI) were involved in the designing of the study and the writing up the article. NK was in charge of the data analysis.

5: References

[1] A. Williams, "The economic role of health indicators," in *Measuring the Social Benefits of Medicine*, London: Office of Health Economic: G. Teeling Smith, 1983, pp. 63–67.

[2] M. Weinstein and R. Zeckhauser, "Critical ratios and efficient allocation," *J. Public Econ.*, vol. 2, no. 2, pp. 147–157, Apr. 1973.

[3] A. Gafni and S. Birch, "Incremental cost-effectiveness ratios (ICERs): The silence of the lambda," *Soc. Sci. Med.*, vol. 62, no. 9, pp. 2091–2100, May 2006.

[4] S. Birch and A. Gafni, "Cost effectiveness/utility analyses. Do current decision rules lead us to where we want to be?," *J. Health Econ.*, vol. 11, no. 3, pp. 279–296, Oct. 1992.

[5] A. Gafni and S. Birch, "Guidelines for the adoption of new technologies: a prescription for uncontrolled growth in expenditures and how to avoid the problem," *Can. Med. Assoc. J.*, vol. 148, no. 6, pp. 913–917, 1993.

[6] S. Eckermann and B. Pekarsky, "Can the Real Opportunity Cost Stand Up: Displaced Services, the Straw Man Outside the Room," *PharmacoEconomics*, vol. 32, no. 4, pp. 319–325, Apr. 2014.

[7] P. Sendi, A. Gafni, and S. Birch, "Opportunity costs and uncertainty in the economic evaluation of health care interventions," *Health Econ.*, vol. 11, no. 1, pp. 23–31, Jan. 2002.

[8] A. Gafni, S. Walter, and S. Birch, "UNCERTAINTY AND THE DECISION MAKER: ASSESSING AND MANAGING THE RISK OF UNDESIRABLE OUTCOMES: UNCERTAINTY AND THE DECISION MAKER," *Health Econ.*, vol. 22, no. 11, pp. 1287–1294, Nov. 2013.

[9] E. W. de Bekker-Grob, M. Ryan, and K. Gerard, "Discrete choice experiments in health economics: a review of the literature," *Health Econ.*, vol. 21, no. 2, pp. 145–172, 2012.

[10] M. D. Clark, D. Determann, S. Petrou, D. Moro, and E. W. de Bekker-Grob, "Discrete Choice Experiments in Health Economics: A Review of the Literature," *PharmacoEconomics*, vol. 32, no. 9, pp. 883–902, Sep. 2014.

[11] A. Gafni and S. Birch, "QALYs and HYEs Spotting the differences," *J. Health Econ.*, vol. 16, no. 5, pp. 601–608, Oct. 1997.

[12] J. J. Louviere, D. Street, L. Burgess, N. Wasi, T. Islam, and A. A. J. Marley, "Modeling the choices of individual decision-makers by combining efficient choice experiment designs with extra preference information," *J. Choice Model.*, vol. 1, no. 1, pp. 128–164, Jan. 2008.

[13] M. R. Mørkbak, T. Christensen, and D. Gyrd-Hansen, "Choke Price Bias in Choice Experiments," *Environ. Resour. Econ.*, vol. 45, no. 4, pp. 537–551, Apr. 2010.

[14] J. M. Rose and M. C. J. Bliemer, "Constructing Efficient Stated Choice Experimental Designs," *Transp. Rev.*, vol. 29, no. 5, pp. 587–617, Sep. 2009.

[15] F. Reed Johnson *et al.*, "Constructing Experimental Designs for Discrete-Choice Experiments: Report of the ISPOR Conjoint Analysis Experimental Design Good Research Practices Task Force," *Value Health*, vol. 16, no. 1, pp. 3–13, Jan. 2013.

[16] J. J. Louviere, D. A. Hensher, J. D. Swait, and W. Adamowicz, *Stated choice methods: analysis and applications*, 7. printing. Cambridge: Cambridge Univ. Press, 2010.

[17] D. McFadden, "Conditional logit analysis of qualitative choice behavior," in *Frontier in econometrics*, New York: Academic Press, 1974, pp. 105–142.

[18] K. Train, *Discrete choice methods with simulation*, 2nd ed. Cambridge ; New York: Cambridge University Press, 2009.

[19] D. Kahneman and A. Tversky, "Prospect Theory: An Analysis of Decision under Risk," *Econometrica*, vol. 47, no. 2, p. 263, Mar. 1979.

[20] W. Samuelson and R. Zeckhauser, "Status Quo Bias in Decision Making," *J. Risk Uncertain.*, vol. 1, no. 1, pp. 7–59, 1988.

[21] D. Kahneman, J. L. Knetsch, and R. H. Thaler, "Anomalies: The Endowment Effect, Loss Aversion, and Status Quo Bias," *J. Econ. Perspect.*, vol. 5, no. 1, pp. 193–206, 1991.

[22] P. Bonsall and B. Lythgoe, "Factors affecting the amount of effort expended in responding to questions in behavioural choice experiments," *J. Choice Model.*, vol. 2, no. 2, pp. 216–236, 2009.

[23] J. J. Louviere, T. Islam, N. Wasi, D. Street, and L. Burgess, "Designing Discrete Choice Experiments: Do Optimal Designs Come at a Price?," *J. Consum. Res.*, vol. 35, no. 2, pp. 360–375, Aug. 2008.

[24] I. J. Bateman, D. Burgess, W. G. Hutchinson, and D. I. Matthews, "Learning design contingent valuation (LDCV): NOAA guidelines, preference learning and coherent arbitrariness," *J. Environ. Econ. Manag.*, vol. 55, no. 2, pp. 127–141, 2008.

[25] B. Day *et al.*, "Ordering effects and choice set awareness in repeat-response stated preference studies," *J. Environ. Econ. Manag.*, vol. 63, no. 1, pp. 73–91, Jan. 2012.

[26] R. T. Yao, R. Scarpa, J. M. Rose, and J. A. Turner, "Experimental Design Criteria and Their Behavioural Efficiency: An Evaluation in the Field," *Environ. Resour. Econ.*, vol. 62, no. 3, pp. 433–455, Nov. 2015.

[27] R. Viney, E. Savage, and J. Louviere, “Empirical investigation of experimental design properties of discrete choice experiments in health care,” *Health Econ.*, vol. 14, no. 4, pp. 349–362, Apr. 2005.
[28] A. Diederich, J. Swait, and N. Wirsik, “Citizen Participation in Patient Prioritization Policy Decisions: An Empirical and Experimental Study on Patients’ Characteristics,” *PLoS ONE*, vol. 7, no. 5, p. e36824, May 2012.
[29] S. Erdem and C. Thompson, “Prioritising health service innovation investments using public preferences: a discrete choice experiment,” *BMC Health Serv. Res.*, vol. 14, no. 1, Dec. 2014.
[30] M. K. Lim, E. Y. Bae, S.-E. Choi, E. K. Lee, and T.-J. Lee, “Eliciting Public Preference for Health-Care Resource Allocation in South Korea,” *Value Health*, vol. 15, no. 1, pp. S91–S94, Jan. 2012.
[31] P. A. Scuffham *et al.*, “Engaging the public in healthcare decision-making: quantifying preferences for healthcare through citizens’ juries,” *BMJ Open*, vol. 4, no. 5, p. e005437, Apr. 2014.
[32] D. L. B. Schwappach and T. J. Strasmann, “‘Quick and dirty numbers’?,” *J. Health Econ.*, vol. 25, no. 3, pp. 432–448, May 2006.
[33] D. L. B. Schwappach, “Does it matter who you are or what you gain? an experimental study of preferences for resource allocation,” *Health Econ.*, vol. 12, no. 4, pp. 255–267, Apr. 2003.
[34] C. Green and K. Gerard, “Exploring the social value of health-care interventions: a stated preference discrete choice experiment,” *Health Econ.*, vol. 18, no. 8, pp. 951–976, Aug. 2009.
[35] C. D. Skedgel, A. J. Wailoo, and R. L. Akehurst, “Choosing vs. allocating: discrete choice experiments and constant-sum paired comparisons for the elicitation of societal preferences,” *Health Expect.*, vol. 18, no. 5, pp. 1227–1240, Oct. 2015.
[36] C. Skedgel, A. Wailoo, and R. Akehurst, “Societal Preferences for Distributive Justice in the Allocation of Health Care Resources: A Latent Class Discrete Choice Experiment,” *Med. Decis. Making*, vol. 35, no. 1, pp. 94–105, Jan. 2015.