Borja Brugés et al

Willingness to Pay and Preferences among Patients Undergoing Cystoscopies: Results from a Large Survey-based Study in Spain

Appendix 1

Questionnaire Translated

Consent

Dear Participant, this survey is being distributed with the aim of supporting a research project led by Cindy Borja as part of her Thesis for the Master's degree in Business Administration and Healthcare Innovation, at the Copenhagen Business School, CBS. Cindy Borja, in cooperation with the medical device company Ambu, will focus her research on patients' preference for cystoscopy procedures. We would like to invite you to participate in this study to determine patient preferences for single-use or reusable cystoscopes. The questionnaire is estimated to take approximately 8 minutes to complete. You only need to complete the survey once.

Only fully completed questionnaires will be included in the analysis. No patient identification data will be collected in this survey. Therefore, by clicking the "continue" button, button located at the bottom, you are giving your consent for the researchers to use your answers provided for the purposes described above anonymously.

The information collected will be used in a master's thesis that will be published and will support other research in the future, being shared anonymously with other researchers. Please answer the questions below from your perspective as a patient.

I am over 18 years of age and I understand that my participation in this survey is voluntary. I am free to leave the survey at any time I wish, without having to give any reason and without any consequences. If you have questions about the survey, please contact Cindy Borja at cibo13ab@student.cbs.dk. Thanks in advance for your participation and collaboration!

Part I

1. Please indicate your place of residence

1. Andalusia

2. Aragón

3. Asturias

4. Canary Islands

5. Cantabria

6. Castilla de la Mancha

7. Castile and Leon

8. Catalonia

9. Madrid

10. Valencian Community

11. Estremadura

12. Galicia

13. Balearic Islands

14. The Rioja

15. Basque Country

16. Murcia

17. Navarre

2, How old are you?

1. 18 - 29

2. 30 - 39

3. 40 - 49

4. 50 - 59

5. 60 - 69

6. 70 - 79

7. 80 - 89

8. 90 or greater

3. How were you born anatomically?

1. Female

2. Male

3. I don't want to answer this question

Part II: A Cystoscopy is a procedure that allows a visual inspection of the inside of the bladder and urethra. A cystoscopy can be used for therapeutic or diagnostic purposes. There are a number of reasons why a healthcare specialist may recommend a cystoscopy. For example: Hematuria, Recurrent urinary tract infections, abnormal urinary symptoms. The cystoscopy procedure can be carried out in different settings including an operating room or in a specialist's office.

1. How many cystoscopy procedures have you had previously?

1. 0

2. 1 - 3

3. 4 - 6

4. 7 - 9

5. 10 or more

2. Did you have your cystoscopy in a private or public hospital / clinic?

1. Private

2. Public

3. Both private and public

4. I do not remember

Cystoscopy is performed with a reusable cystoscope. For cleaning, the device requires reprocessing between each procedure to ensure disinfection and guarantee the safety of the next patient. Sanitary reprocessing of reusable cystoscopes includes the use of chemicals and trained personnel to carry it out. (Image reusable cystoscope)

A medical device company has developed a single-use sterile cystoscope. This single-use cystoscope comes sealed in sterile packaging ready for use on each patient, once the procedure is completed, the cystoscope is discarded. The performance of the single-use cystoscope is comparable to the performance of the reusable cystoscope in terms of image quality and bending capacity. (Image single use cystoscope)

3. If you had a choice, would you prefer to have a cystoscopy with a single-use or reusable device?

1. Single use cystoscope

2. Reusable Cystoscope

3. I do not have preferences

3.1 Please comment Why would you prefer to have the procedure done with a reusable cystoscope? :

4. In a hypothetical scenario, would you be willing to pay a small extra fee to perform your cystoscopy with a single-use device rather than a reusable one?

1. Yes

2. No

3. I do not know

5. Can you imagine asking your doctor if he / she could perform a cystoscopy with a single-use device instead of a reusable one?

1. Yes

2. No

3. I do not know

6. If the hospital / clinic you serve ONLY offers a cystoscopy with reusable and NOT single-use cystoscopes, Would you rather find another hospital / clinic where you can have your cystoscopy done with a single use cystoscope?

1. Yes

2. No

3. I do not know

7. Would you prefer to have the Consultation and the Cystoscopy on the same day or on two different days?

1. Consultation and cystoscopy on the same day

2. Consultation and Cystoscopy on two different days

3. I do not know

8.Are you concerned about the possibility of the cystoscope becoming contaminated (presence of infectious agents in the cystoscope), when you have a cystoscopy procedure?

1. Yes

2. No

3. I do not know

9.You have any of the following concerns regarding single-use or reusable cystoscopes: (Leave blank if the specific concern is not important to you, if you are not concerned with any concerns, please click "none"

	Single Use Cystoscopes	Reusable Cystoscopes
Contamination	❑	❑
Environmental impact	❑	❑
Urinary tract infection	❑	❑
Cystoscope rate of performance	❑	❑
None	❑	❑

10. In times of COVID-19, are you more concerned than before about exposure to contamination (presence of infectious agents in the cystoscope), regarding your cystoscopy procedure?

1. Yes

2. No

3. I do not know

11. During the cystoscopy process, the doctor connects the cystoscope to a screen. This allows the doctor and the patient to follow the procedure inside the bladder and urethra. Would it be valuable to you to get a photo or video of the findings of the cystoscopy procedure?

1. Yes

2. No

3. I do not know

Part III The preferred setting for treatment Thanks for your help so far! This is the last section of the questionnaire and you will now be introduced to what-if scenarios for your choice. In the next six questions you will be presented with two different scenarios in each question. Please select only one scenario which best suits your needs.

Attributes	Alternative 1	Alternative 2
Direct Cystoscopy	No	Yes
Risk of Cancellation	6%	6%
Contamination	0%	0%
Negative Environmental Impact	High	High
Cost per Procedure	165€	70€

Willingness to Pay and Preferences among Patients Undergoing Cystoscopies: Results from a Large Survey-based Study in Spain

Appendix 2

Discrete choice experiment

Lancaster's approach comes with the principle that all the goods have attributes which are demanded by the consumers. This opinion led to analyze the demand of a good in a different way as the traditional demand theory would not be able to forecast the fluctuation on demand when a specific attribute or characteristic was changed. Subsequently, Batten and Johansson (1987) prolonged the Lancaster's model to preferences groups, studying substitute products as well as technological changes. Later Trajtenberg (1990) outlined theoretically and econometrically the basis to measure social gains from innovation in goods changing characteristics like prices, production costs, attributes and preferences.[1] Choice experiments are better to measure passive values, as the utility connected with an alternative, is a function of the observed level and also the unobserved ones. In a Discreet Choice Experiment (DCE) method the attribute levels can be controlled in the experiment and through this technique a quantifiable result of the impact of the changes in attribute levels on the choices can be found by decompiling the choice. In other words, this quantitative technique helps to elicit individual preferences and uncover what is the value that a respondent will put into a specific attribute of a product or intervention when being exposed to hypothetical alternatives. Some of the reasons are that they can prove a more well-off description of the attribute tradeoffs that individuals are willing to make, that is to say, it helps to comprehend the relative valuations of various attributes, Additionally DCE have become popular in healthcare for example for healthcare provision and treatment characteristics, and it also allows to include non-health characteristics related to process utility like the contingent valuation method. The use of DCE may help to design future services or existing ones and consequently identify any opportunity cost, meaning potential benefits that one could miss when choosing one alternative over another.[2]

To carry a DCE, efforts are required to develop the scenarios with appropriate attributes which also has to have an attribute-based approach to measure the amount of other goods required for compensation. Thereafter the use of a proper statistical method is required.

The International Society for Pharmacoeconomics and Outcomes Research (ISPOR) created good research practices for conjoint analysis in health care applications, thanks to regular meetings with the aim to identify relevant steps and improve key criteria in conjoint analyses plus following the structure stablished from Drummond.[2] It resulted in a 10 items checklist: 1) research question; 2) attributes and levels; 3) construction of tasks; 4) experimental design; 5) preference elicitation; 6) instrument design; 7) data- collection plan; 8) statistical analyses; 9) results and conclusions; and 10) study presentation. This check list can facilitate to compare good research practice for applications of conjoint methods in healthcare studies.[3]

Another more specific approach from the article "Constructing Experimental Designs for Discrete-Choice Experiments: Report of the ISPOR Conjoint Analysis Experimental Design Good Research Practices Task Force" aims to be more explicit in regards the experimental design for Discreet Choice Experiments[4]. Instead of a checklist, it is illustrating the key stages for the experimental design of a DCE. The key stages are: 1) research question; 2) attributes and levels; 3) Choice Question Format; 4) Analysis Requirements 5) Experimental Design; 6) Data; 7) Statistical Analysis. For this study both of the previous artefacts will be used in a combined way to describe as appropriable as possible the DCE method used.

1) Research Question:

Is relevant for a good research question to specify a testable hypothesis (exploratory and/or descriptive) and to define the study perspective and reasons behind the study.[3]

Hypothesis: the main objective of this study is to be able to identify patients' preferences within cystoscopy procedures focused on the type of devices being used. The raised hypothesis aims to identify which are the most relevant attributes/characteristics for decision making. This is done by defining the correspondent features and finding which are statistically significant meaning the noteworthiness role in patients' choice. Additionally, significant relevance is concentrated on measuring their overall willingness to pay and per specific attribute when comparing the preference of patients within both devices, single-use and reusable ones.

Perspective: The study is being addressed from the patient's perspective. The results from this study aims to contribute in the enhancement of the patient's voice across public and private hospitals/clinics across Spain as well as in a general way for patients who may undergone a cystoscopy. Then the results may well contribute

partly for an overall HTA evaluation for decision making if these results could be complemented with the other relevant aspects for a health technology assessment. The definition of patients in this paper is defined as individuals over 18 years old who had undergone at least one cystoscopy procedure in the past. The reasons why a patient may have resulted undergoing cystoscopy procedures are not relevant to the inclusion criteria neither social economical level meaning aspects like income, level or education, employment etc., are not being classified. It also assumed that all the patients from the Spanish sample count with the same public coverage for healthcare interventions and social supports.

Rational for DCE: Among the variety of techniques, DCEs have become in recent years, the most commonly applied method in healthcare.[4] This study aims to contribute to the patient's aspect of a health technology assessment. To do so the DCE method to elicit preferences helps to assign utilities to the stated predilections denoted by the individuals. In contrast to extra-welfarist methods, where the weighting of the outcomes can be based on principles across individuals rather than based on preferences. In the same manner alternatively to ways defining outcomes like health status, life years, QALYs, DALYs, illness events avoid, among others, the method of discreet choice experiment has been chosen as it aims to clarify the individual perspective without any health outcome measure. Consequently, ranking and rating methods can help to assign values to some extent in an ordinal scale but still is difficult to assign absolute values to preferences.

The method allows to easily reveal the individual trade-off between the features of the intervention or service estimating the relative importance of the attribute to be able to predict a possible demand, in other words it can also help to measure the trade-offs which are often implicit in the questions of choice.[5]

2) Attributes and levels:

To be able to elicit preferences**,** is necessary to define the most probable profiles/alternatives to patients over the range of attributes or characteristics of a healthcare intervention. Inclusions and omission of the attributes may be explained.[3] This paper intent to stick to the health economic terminology by Carlsson and Martisson (2003) [6]:

Attribute Identification & Selection: According to Bridges et al 2011, there should be a balance in the selection of the attributes according to the perspective chosen, what is relevant for the respondents, in this case

the patients, and at the same time considering the decision making or policy setting. The identification, inclusion and/or exclusion of the attributes should be evidence-based.

Some of the sources to base the evidence could be by literature review, consultation with clinical experts, pilot testing, qualitative research or previous studies. The attributes chosen for the DCE were based on the electronic search of the databases PubMed and Embase, studies were considered after 2010. Several brief summaries were assed based on reusable cystoscopes in the aim to provide an overall understanding on issues like infection rates, contamination risks and outbreaks reports, cost on JJ stent removals, and also costs on office, environment and other adverse events like cancellation risks**.** Beside some references all the content was based on the data on file confidential document from Ambu®. The environmental attributes were also based on literature review plus the assessment of an internal Environmental and Climate Specialist from Ambu®. Extra inputs were picked from qualitative research like semi-structured interviews with the targeted population and clinical experts.

The number of attributes suggested by a literature review from Ryan and Gerard 2003, suggest an acceptable amount of four to six attributes for the choice task. Though according to Bridges et al. 2011 there is not a set rule, however a plausible profile or alternative to be chosen has to be assured. Likewise, a review by Marshall et al 2010, shows that a 70% of the studies use three to seven attributes and four levels. Is important to remember that there can be a challenge when incorporating DCE and WTP into economic evaluation as conferring with Brennan and Dixon 2013, there is not a standardization in the ranges of the attributes.[2] The mixture of attributes embraces scenarios that are selected from the all the possible scenarios. It is also important to be able to measure the effects of the attributes in an independent way.

Levels Selection: The levels of the attributes were collected and defined in the same way than the deduction of the attributes was done. Other cost prices were compared from the Spanish site “eSalud” (http://esalud.oblikue.com/index.asp) which is based on the Spanish Official Gazette, giving levels of a min of 64.78€, an average of 150.22€ and a max of 293.92€, however it was decided to choose values from a previous UK survey, to keep standardization as the ranges were very similar.

The levels can be categorical, continuous or a probability. Yet no best practices exist in terms of recoding subjective levels. Furthermore ranges (i.e., 50-100€) were avoided as this would give the chance to respondents to interpret the levels at their own way which can result in ambiguity among all respondents.

ISPOR in their checklist, recommends to limit the levels from 3 to 4, so the full span do not necessarily have to be included. The decision behind the choice of levels should be transparent and verified by the pilot testing.

3) Construction of Tasks / Choice Question Format:

The tasks are the mechanism by which the possible profiles are presented to the respondents, which could be presented in several ways fluctuating on the number of attributes, profile choices and other alternatives.

Full or Partial Profile: A full profile was presented to the respondents, meaning each profile/ alternative had all the attributes included. After a pilot testing it was confirmed that including all the attributes did not cause misunderstanding in the conception neither was shown any tendency to ignore attributes from the respondent's side or cognitive fatigue.

Number of Profiles / Alternatives: Although it can be considered that more profiles/alternatives may represent more efficiency, according to Bridges et al., 2011, little research shows the effect on respondents when increasing the profiles. There might be a tendency to lose interest and end in increased distraction if too many options are shown. In contrast to rating the profiles / alternatives, for this study two alternatives were presented and it was induced to choose among the two options using the format of the forced choice-elicitation. The two set of profiles were shown in six scenarios, or six choices situations. However, according to Potoglou et al. 2011 for respondents who do not have experience making choices between two or more profiles/alternatives a best-worst scale will be more suitable, likewise more attributes could be presented. Subsequently, studies shown that the preference weight are not statistically significant between the best-worst scale and the standard DCE.[2]

Opt-out or status-quo options: No Status-quo option was included in each task as this would limit the estimations of the patient's preferences by censoring the data. This could also affect the statistical efficiency measurements.[4] Other assumption done is that the inclusion criteria comprised patients who had undergone a cystoscopy, it could also be seen, as a patient who will be in a scenario where the cystoscopy procedure was necessary and unavoidable, so even if it was not a hypothetical scenario, a decision regarding the needed

cystoscopy procedure would have to be taken imperatively. Status quo option would have given the chance to the respondent to not choose among any of the profiles/ alternatives presented and this will also represent a challenge into figuring out the patients' preferences.

Two hypothetical alternatives were presented and the patients could answer on their preference. After the first answer of each respondent, the following scenario was personalized based on the previous answer. The significance of the results was assessed in a 95% confidence interval being statistically significant all the p-values lower than 0.05

4) Experimental Design

The experimental design helps analysts to chart the attributes and levels as set of alternatives to be chosen. The perspective and the identification of the respondents is highly important to define the alternatives. The following consideration were done and helped to systematically elicit the choice response.

• The relevant attributes of the cystoscopy procedure were identified and the stakeholder was identified. In this case the main stakeholder was a patient who had undergone a cystoscopy procedure.

• The specific value of levels and ranges were approved.

• The observations were presented in a D-optimal design, explained below.

• Strategy of how the data was modelled as a function of levels and attributes. In this case each attribute represented a variable.

A total of five attributes represented by A, and the levels represented by L, led to the calculation of the total combinations (# alternatives = #levels#attributes). Three levels for four attributes and two levels for one attribute will equal to the total combinations of 162. (81 x 2) also called a full factorial design. Nevertheless, this full factorial design would have meant too many combinations for each respondent therefore a D-optimal design was used to identify the most favorable combination set where the main effects and all the higher order interactions could still be estimated and assed, this is called a fractional factorial design, keeping the same properties as the full design in the best possible way capturing all the relevant effects to be assed.

Experimental Design Concepts

Identification: The parameters were identified as not being bias and coded accordingly. With this the effects were also acknowledged to be independent.

Efficiency: The statistical efficiency of the model may depend on the sample size and on the way the attributes were combined. For this study a sample of 300 respondents was selected. The combination of attributes was performed randomly by the survey tool QuestionPro, using the D-optimal design provided by a computer algorithm, according on the first answer of the respondent the following scenarios will be based on that first response. This means some scenarios could have been improbable but not illogical.

Implausible Scenarios: The study may have had implausible scenarios, nevertheless, all the levels chosen were realistic and this would be complemented with the D optimal design used to mitigated.

Interaction Effects: In the profiles showed neither type of symptom's severity nor duration was assed. This could have been an important aspect to estimate to assure the efficiency of the model and be able to help respondents to evaluate. According to Zwerina at al 1996 the majority of the studies use the fractional factorial design where the interactions between attributes are insignificant in a two-way or higher order of interactions. The interaction may take into account how the preference of a patient on one attribute or variable may depend on the level of another attribute. In contrast to this Louviere et al. 2000 express that the exclusion of the interactions may not necessarily lead to bias. [7]

Cognitive limitations of particular groups of respondents: the cognitive capacity of the respondents was not taken into account. People with Alzheimer, schizophrenia and other neurological conditions were not strictly excluded.

Labeled and constant alternatives: No specific labels or names were used during the DCE, though previous explanation of the single used device launched by Ambu® was described, without mentioning any brand.

Blocking: Due to the extended full factorial design a blockage of six scenarios were arranged. This this not mean blockage in questions but instead in the combinations. Thus, resulting in the increase of efficiency of the respondents.

Orthogonality: refers to the exposure of all elements or levels of the attributes in a balanced way. According to Huber and Zwerina there are three properties for efficient designs: 1) the Level Balance: levels of an attribute occur with equal frequency. 2) Orthogonality: the occurrence of any two levels of the different attributes are uncorrelated. And 3) The minimal Overlap: the cases where attribute levels do not change across a choice set should be minimized.[8] Nevertheless there are different ways to make an efficient design and one is the D-optimal Experimental design.[4] D-optimal design allow parameters to be estimated without bias and minimizing the variance. Reducing the number of runs to reach an equivalent precision.[9] This was also translated into a more economical option. However, D-optimal design matrices tend not to be orthogonal and effects appraised can be correlated. [10] The D optimal design is a computer-generated design which needs an algorithm, this helps to determine the best experiment making alternative versions for each level and that's how an alternative is presented to the respondent according to their first answer. This method consists of the best subset of the theoretically, practical and conceivable possible experiments which could be carried out maximizing the result of the matrix. It is also good to reduce the number of runs or choice sets required.

5) Preference elicitation

As there is not really a defined level of information to be include in the questionnaire, the amount and of information was assessed by a previous questionnaire in the UK showing satisfaction on patients to the extension and content of the survey. The discreet choice experiment was introduced with an explanatory text aiming to help the respondents to previously understand the set-up of the block of questions to answer but not practice example was performed. The force-choice format was induced.

6) Instrument design;

The survey questionnaire was divided into four sections: 1) Introduction and Informed Consent: a brief introduction on the aim of the study, the anonymity of personal information and an informed consent was described in this section, allowing the participant to reflect and agree voluntarily to the survey. All the participants had the choice to leave the survey without giving any reason, whenever they would like to 2) Part 1: Descriptive Information: general questions to find info on Age, Residence, Gender etc. 3) Part 2: Cystoscopy Procedures and Devices: an introduction to the cystoscopy procedure was provided here as well as a brief difference on the

type of cystoscopes of single use vs reusable, general questions were also stated in this section, like the context of the cystoscopy i.e. public/private hospital/clinic, overall willingness to pay, general concerns on the procedure, the cystoscope and overall concern on Covid-19 and lastly more specific characteristics of concern on the devices for cystoscopy. 4) Part 3: Discreet choice experiment: Finally, the experimental design on the DCE with a brief introduction for the task and the explanation on each of the attributes to evaluate.

7) Data - Collection Plan;

The Discreet choice experiment method was performed in an empirical way, collecting data based on a plan for sampling strategy, mode of administration and assessment of ethical considerations. According to Orme BK 2006 the minimum sample size recommended are 300 respondents with 200 resp. per group of subgroup analysis. The current study includes 300 respondents of a subgroup, in this case the subgroup are patients who had undergone a cystoscopy procedure.[3]

The mode to administer the survey was done through email. The company IQVIA was in charge of distributing the link survey to each patient and it was done through an extra QuestionPro account. IQVIA has a database, where they could look according to the criteria of several pathologies. The pathologies were suggested according to the previous knowledge on which a cystoscopy may be required. This was the starting point for IQVIA to select the possible respondents. From an approximate pool of 1512 individuals who had click to enter the survey, 300 were selected as the ones who complied the acceptance criteria to have at least one cystoscopy procedure before and completed the full survey.

The accuracy and the easiness for completion of the questionnaire was assured by simplifying as much as possible each question. The questionnaire was reviewed by three people having different inputs and agreeing in the most simplified way assessing readability, population suitability, and ethical aspects.

As this study did not collect any personal data, none ethical approval was necessary. However, the legal consent in the survey was requested. The approval of the consent was done after each respondent will click continue in the first part of the questionnaire. Besides © 2020 QuestionPro is ISO 27001:2013 certified company, which is globally recognized internationally for standard management of risks and conforms with the General Data Protection Regulation (GDPR) privacy and data security by the European Union-regulation.[11]

8) Statistical analyses;

A good understanding of the preference data and its characteristics is highly required for a well-made DCE. To be able to do this, is also relevant to understand the alternative methods, the assumptions made, strongpoints and restrictions of the method chosen.[12] The task force article by Hauber et al, 2016, aims to gather all the different methods used in DCE. Because there is some inconsistency in the methods used and a lack of knowledge in the common audience about other alternative methods, the article mentioned, can help to give a brief understanding of the statistical methods more used.

Coding of attributes: There are two ways of coding the attributes. It can be effects coding or dummy-variable coding. In either way, one level of each attribute is omitted. This means when one profile/alternative level was chosen it will be appearing as 1 while the non-chosen will be 0. In the approach of effect coding the non-omitted levels are represented as -1 when the omitted level is present. On the second approach of the dummy-variable, all non-omitted levels will be 0 when the omitted level is present. The decision of which approach to use may depend on the simplification for the understanding of the estimates.

For this study the dummy-variable coding method was used. This means that the coefficients estimated, represent a measure of the strength of the preference of that level making refence to the omitted level of that specific attribute. Therefore, the statistical significance will be reflecting the difference between that preference weight and the omitted category.

Data Generated: The date was imported from QuestionPro to Stata/SE 16.0. Each respondent was coded with an ID, each respondent then answered between 2 profiles/alternatives, 6 times. Generating 12 tasks. The Task is the number of the choice task. This generated 3600 observations or rows as there were 300 respondents. (12x300). All the levels of each attribute were classified into a variable with the dummy-variable approach. Where 1 was representing the chosen alternative and 0 the non-chosen. All the mid-levels were selected as the reference meaning the omitted ones. Thanks to the dummy-variable coding is possible to get meaningful interpretations between the changes in attribute levels, in contrast of having the absolute values of preference weights alone. Fourteen main variables were considered for the DCE experiment.

Model: Most of the commonly used methods to analyze DCE in health care are extensions of the Conditional Logit Model which according to Mc Fadden (1974) is consistent with the random utility theory, originally for transportation choices. Among these extensions it can be found the Random Parameter (mixed) Logit (RPL) or also called the "mixed-logit", the Hierarchical Bayes (HB) usually used to model responses from each individual and not all observations from a sample, allowing also preference heterogeneity, and the Latent-Class Finite-Mixture model (LCM) which are based on limited dependent-variable models consistent with random utility and assuming a probability of choosing an alternative from 0% to 100%.[12] Thanks to the random utility theory (RUT), the utility connected with an alternative, is a function of the observed level and also the unobserved ones. This assumes that a respondent will try to maximize his or her utility when making a choice between two or more alternatives. Being a utility maximizer assuming that the individual acts rationally when making a choice, looking for the highest utility.[7] The random utility comes from the probabilistic choice theory and this one is based on the assumption that there is some uncertainty in line to the choice of an individual, being hard to predict individual choices in a perfect way. This model assigns to each alternative a probability of being picked. The model can be classified into two families:[7]

1. Decision rule assumed to be random; Utility deterministic – Individual's behavior can change according to internal and external factors (Intrinsically probabilistic)

2. Decision rule assumed to be deterministic; Utility Random – Inability of the analyst to frame the individual's behavior.

For binary discreet choice models, where respondents are to answer more than one choice, the variation across **i**ndividuals, is not considered random, the following utility formula can be represented as:

$$Uin = Vin + \varepsilon in + \mu n$$

Where n is representing a respondent, V is the function defined by the attribute levels for alternative i in a set of alternatives up to J, and εi is the random error, which can be represented as the lack of information that the individual has for making a choice

Lastly, μn represents the random error across respondents and will be a constant for each respondent, in this case is null. In the aim to clarify, one example from the actual questionnaire in Table 1:

Table 1: Example of the choice sets

Attributes	Alternative 1 (i)	Alternative 2 (j)
Direct Cystoscopy	No (Xi_1)	Yes
Risk of Cancellation	6% (Xi_2)	6%
Contamination	0% (Xi_3)	0%
Negative Environmental Impact	High (Xi_4)	High
Cost per Procedure	165€ (Xi_5)	70€

Each alternative provides a utility and the RUT assumes that each respondent will choose the alternative with the highest utility. The utility is then decomposed into the observed factors, like attributes, alternatives and attributes of the decision maker that will affect the utility, plus the unobserved factors. Then V is a vector of the attributes of the alternative and also the vector of attributes of the decision maker. In this study, the decision maker attributes were generalized so we should only take into account the attributes of the alternative. Being denoted as:

$$Vnj = \beta' Xnj$$

where n represents a respondent. Assuming a linear utility function, as its usually done as a standard practice in DCE,[7] the utility formula could be represented as

$$Unj = \beta' Xnj + \varepsilon nj$$

Thereafter the coefficient β represents the relationship of the observed attributes with the unobserved utility. The respondent is to maximize the total utility but not the representative utility. The representative utility is part of the total utility. The representative utility will be then denoted as below:

(Note: in the study there is no intercept coefficient as they are cancelled in the likelihood and therefore do not have an identifiable role in the model (Ying So, 2010))

$$V_{n1} = \beta 0 + \beta 1 X_{ni1} + \beta 2 X_{ni2} + \beta 3 X_{ni3} + \beta 4 X_{ni4} + \beta 5 X_{ni5}$$

Vn1 is representing the representative utility of choosing alternative 1, β1 then represents the coefficient that will be calculated to show the marginal utility of the respondent relative to the attribute 'Direct Cystoscopy', thereafter calculated with Xni1 which is representing the level of 'Direct cystoscopy' in this case 'No'. The same will be with β2Xi2 which is then the coefficient to be calculated of the attribute 'Risk of cancellation' and Xi2 is one of the levels of the attribute in this case 6%. So on and so forth for the last three attributes. When one of the attributes is 'Cost' the marginal willingness to pay or part-worth's can be calculated. For the dummy coding effect, the values here will be 1 if chosen cero if not chosen Thereafter the use of the probability statements where it will be desired that the probability of the alternative chose will be close to 1 and the probability of the no chosen alternative will be close to 0. As the probability is on a cardinal scale, so are the utility scores therefore a meaningful wtp can be calculated thanks to the cardinal scale,[7] and 'all health interventions can be ranked according to their willingness to pay.[2] Is important to highlight that what is relevant is to compare the level of utility among the two alternatives. For this study the Conditional Logit model (CLM) was used.

Conditional Logit Model

Helps to relate the probability of choice between two or more alternatives relative to the levels/characteristics of the attributes. In other words, it can help to estimate the average preferences in a sample, though it may not explain for preference heterogeneity or variability and this could lead to some bias, unlike the RPL mixed logit model which is becoming more used. Heterogeneity suggest that preferences vary among individuals therefore the analyst cannot know which type of preference is hold by the individual and as a consequence cannot predict accurately the choices made.[7]

This study used a binary discrete choice showing to the respondent two alternatives, where the dependent variable had two states.

The CLM based on logistic regression or logit regression, and multinomial logit model (RPL) has the same statistical assumptions though multinomial logit usually aims to relate the choices made to the characteristics of the respondents while in contrast the CLM relates the choices to the elements defining the alternatives in between the respondents. Therefore, the conditional model was more relevant for this study where the general preferences of patients are willing to be discovered. Thereafter the multinomial logit is more suitable when the dependent variable has more than 2 alternatives. In this study, the dependent variable has only two options.

(selected or non-selected) therefore the CLM model was chosen. Thereafter, HB methods may also require updates in preferences estimates repetitively and this may be a concern for transparency compared to other methods. Lastly the LCM method requires a more specialized software in contrast to the conditional logit model which many software's have available. It also requires a deeper substantial judgment from the analyst side as well as a larger sample size to adjust for the increased number of parameters needed.

The conditional logit model is also called fixed effects logit models and is based on the logistic regression that is based on log(odds) as unit of measurement, which can be defined as a scale of the coefficients on the Y axis. The logistic regression is similar to linear regression but is predicted weather something is true or false or something continuous. Logistic regression tries to fit an "s" shape from 0 to 1 and it can be useful for discreet and continuous data. It is also useful to classify which variables are useful for organizing the samples.[13] Thereafter, conditional logistic regression is an extension of the logistic regression which takes into account stratification and matching. The main difference between logistic regression and conditional logistic regression is that the respondents are exposed to different situations before stating their choice. Then instead of having individual characteristics in the logistic regression, there will be characteristics of the different alternatives, in this case attributes, proposed to the respondents in the CLM. The observations are not independent within a block from the same respondent.[14] So, there is one row for each profile/alternative per choice task per respondent. In this study each respondent had 12 rows as there were 6 choice sets and 2 alternatives in each.

Differing from linear regression into how to measure the explanatory power of the model, the logistic regression uses the log likelihood.[12] After trying to find the best fit by picking a probability in the "s" line and observe the probability in another point, the curve with the maximum likelihood is selected. The algorithm with the maximum likelihood, find the optimal fit and that's the one chosen as the best fit. However, the log likelihood can-not be used alone to measure the model fit, it can be useful to calculate the "goodness-of-fit" like with the likelihood of the chi-square test and Mc Fadden's pseudo-R-squared.

In the conditional logistic regression, the likelihoods are computed relative to each stratum. Stata calculates the maximum likelihood of the dependent variable which in this case is dichotomous.[15] As some of the characteristics to consider for the conditional logistic regression, a few are named:

- There should be a panel data or longitudinal data, which can provide information about the behavior of the individuals.
- The dependent variable has to be binary or dichotomous.
- The same individuals are measured at two or more time points, in this case six scenarios/ choice sets were presented so they are considered to be measured six times.

Whilst several data assumptions have to be taken into account:

- Every respondent should have two or more measurements on the same dependent variable.
- For some of the individuals in the sample, the values of the independent variable must change over time. The change in the independent variable may create changes in the dependent variable, meaning if the dependent variable does not change over time those respondents should not be considered.

On one hand some of the benefits of the conditional logistic regression are that all the characteristics are controlled and therefore stable and out of the regression i.e., ethnicity and birth year, which may contribute to reducing bias. This model is better to estimate variables that can change over time, and another reason to attribute a plus is that it would be hard to attach behavioral interpretation when focusing only on the characteristics of the individual and not of the good.

The reason why CLM is a good fit for DCE experiment according to Hoffman and Duncan 1988, is due to the choice between alternatives is modeled as a function of the characteristics of the alternatives, in other words of the attributes and its levels, instead of the characteristics of the respondent who is making the choice. In this study the attributes have two or more levels describing to be chosen.

The DCE allows to control over some attributes and levels and at the same time helps to decompile the impact of the changes in the attribute levels in a quantifiable manner. The effects of the impact represent the strength for the preference also called “part-worth utility” or “preference weights”.

The estimated model in Stata/SE 16.0 was performed in the following way:

clogitselected, immediately to cystoscopy Risk of cancellation 0%, Risk of cancellation 12% Risk of contamination 0% Risk of contamination 12% Negative environmental impact Low Negative environmental impact High Cost per, group (taskid)

The command clogit, represents the CLM, in Stata, the second object "selected" is the dependent variable tailed by the regressors. As Stata needs extra info to identify which observations to compose the variable Taskid helps the software to recognize the observations group for a particular choice. From the conditional logit model the command WTP (willingness to pay) is used to estimate the willingness to pay for each attribute. Next, the estimated coefficients can be represented as a preference weight, representing the relative contribution or the so-called relative importance of the attribute levels or the utility that respondent's assign to an alternative.

The WTP of a respondent for a unit of change in the specified attribute can be found thanks to the ratio (marginal rate of substitution) of an attribute over the absolute cost.[2] by measuring the utility that each respondent put into attributes, explaining the relevance of importance of the choices attributes for decision making when being ask which device they will prefer. The understanding of the probabilities allows a cardinal utility scale which is needed to compare possible benefits, in this case monetary measures like the willingness to pay.[7]

References

1. Marcin TC. A Characteristics Model Approach to Demand Analysis for Wood Composites. *For Sect trade Environ impact Model theory Appl Proc an Int Symp*. 1992:119-124.
2. Drummond M, Sculper MJ, Claxon K, Stoddart G, Torrance G w. Methods for Economic Evaluation. 4 th ed. New York ; Oxford University Press; 2015.
3. Bridges JFP, Hauber AB, Marshall D, et al. Conjoint analysis applications in health - A checklist: A report of the ISPOR Good Research Practices for Conjoint Analysis Task Force. *Value Heal*. 2011;14(4):403-413. doi:10.1016/j.jval.2010.11.013
4. Johnson FR, Lancsar E, Marshall D, et al. Constructing experimental designs for discrete-choice experiments: Report of the ISPOR conjoint analysis experimental design good research practices task force. *Value Heal*. 2013;16(1):3-13. doi:10.1016/j.jval.2012.08.2223

5. Lemiere C. Discrete Choice Experiment (DCE): a methodology for eliciting health workers preferences. *World Bank Webminar*. January 2009:6. Available from: https://www.who.int/hrh/migration/hmr_expert_meeting_lemiere.pdf. Accessed: August 30, 2021.

6. Rose J. Workshop on Experimental Design and Data Analysis Experimental Design. University of Sydney Webinar, December 04,05,2012. Accessed August 30, 2021. file:///C:/Users/Cindy/Downloads/Workshop%20John%20Rose.pdf

7. Kjær T. A Review of the Discrete Choice Experiment - With Emphasis on its Application in Healthcare. *Heal Econ Pap*. University of Southern Denmark. 2005;(1):1-139. Available from: https://www.sdu.dk/-/media/files/om_sdu/centre/cohere/working+papers/20051pdf.pdf. Accessed: August 27, 2021

8. Zwerina K, Huber J, Kuhfeld W. A general method for constructing efficient choice designs. Durham, NC Fuqua Sch Business, Duke Univesrsity. 1996;(September):39-59. Available from: Zwerina Kuhfeld Huber.pdf (duke.edu). Accessed: August 30,2021.

9. QuestionPro. Conjoint Analysis- D-Optimal Design |. Published 2021. Available from: https://www.questionpro.com/help/conjoint-analysis-d-optimal-design.html.. Accessed August 20, 2021

10. Nist Sematech. 5.5.2.1. D-Optimal designs. Engineering Statistics Handbook. Published 2021. Available from: https://www.itl.nist.gov/div898/handbook/pri/section5/pri521.htm.. Accessed August 20, 2021.

11. QuestionPro. Security & Privacy | QuestionPro. Published 2020. Available from: https://www.questionpro.com/security/.. Accessed August 23, 2021.

12. Hauber AB, González JM, Groothuis-Oudshoorn CGM, et al. Statistical Methods for the Analysis of Discrete Choice Experiments: A Report of the ISPOR Conjoint Analysis Good Research Practices Task Force. *Value Heal*. 2016;19(4):300-315. doi:10.1016/j.jval.2016.04.004

13. Politzer-Ahles S. Interpreting logistic regression coefficients. Published 2016. Available from: https://www.polyu.edu.hk/cbs/sjpolit/logisticregression.html.. Accessed August 31, 2021.

14. Xlstat A. Conditional Logit model | Statistical Software for Excel. Published 2021. Available from: https://www.xlstat.com/en/solutions/features/conditional-logit-model.. Accessed August 31, 2021.

15. StataCorp. Logistic regression | Stata. Published 2021. Available from: https://www.stata.com/features/overview/logistic-regression/.. Accessed August 31, 2021.

579 **Willingness to Pay and Preferences among Patients Undergoing Cystoscopies: Results from a Large**
580 **Survey-based Study in Spain**

581 **Appendix A3**

582 **Variables**

Variable Name	Variable Description
ID	Respondents ID
TaskID	Particular value for each choice set.
Inst_Cystoscopy	The availability to go direct to cystoscopy in the same day, Yes/ No, qualitative variable.
Risk_C1	Risk for cancellation of the cystoscopy 0%, quantitative variable. Dummy-variable coded.
Risk_C2	Risk for cancellation of the cystoscopy 6% , quantitative variable. (Chose as reference, omitted). Dummy-variable coded.
Risk_C3	Risk for cancellation of the cystoscopy 12%, quantitative variable. Dummy-variable coded.
Cont_R1	Risk of Contamination during cystoscopy 0%, quantitative variable. Dummy-variable coded.
Cont_R2	Risk of Contamination during cystoscopy 6%, quantitative variable. (Chose as refence, omitted). Dummy-variable coded.
Cont_R3	Risk of Contamination during cystoscopy 12%, quantitative variable. Dummy-variable coded.

Environmental _Imp_Low1	Low Negative Environmental Impact, Qualitative Variable. Dummy-variable coded.
Environmental _Imp_Neutral2	Neutral Negative Environmental Impact, Qualitative Variable. (Chose as refence, omitted). Dummy-variable coded.
Environmental _Imp_High3	High Negative Environmental Impact, Qualitative Variable. Dummy-variable coded.
Cost	Cost per procedure of cystoscopy. Continuous Variable.
Selected	Dichotomous variable, representing the profile/alternative chosen by the respondent in each scenario. 1 for the selected and 0 for the non-selected. Dummy-variable coded.

583