Individualized Outcome Measures in Prostate Cancer Screening

Comparison of three instruments for the elicitation of prostate cancer screening preferences in a Portuguese population

Erasmus School of Health Policy and Management
Master Thesis in Health Economics, Policy & Law
Word Count: 13838

Francisco Pais Baptista
573331

Supervisors:
dr. W.K. van Deen MD
S.P.I. Huls, MSc

Reading Committee:
Prof. dr. M. Varkevisser

Rotterdam
June 2021

Abstract

Background

Individualized outcome measures are used to elicit the individual's goals for care and their respective prioritization, thus acknowledging the patient as the central element of care. Such measures acquire particular importance in medical areas where the risk-benefit relation of a procedure is not consensual. Prostate cancer screening is an example of such a procedure. Given the multiplicity of instruments available for eliciting patient preferences for prostate cancer screening, it is important to know how different instruments are used and perceived by patients. In this study, a ranking, a rating, and a point distribution instruments were compared for the elicitation and prioritization of individual preferences for PSA screening.

Methods

A sample of Portuguese men was recruited and asked to fill in a web-based questionnaire consisting of three different instruments – rating, ranking, and point distribution – to elicit their preferences for prostate cancer screening. In each instrument, respondents were asked to reflect on four outcomes of prostate cancer screening as well as on that instrument's usefulness and ease of use. The instruments were compared in terms of their outcome valuation patterns, usefulness and ease of use scores and time to completion by means of descriptive statistics, Friedman tests, Wilcoxon matched-pairs signed-rank test and ordinal logistic regressions.

Results

119 respondents were eligible for analysis. The pattern of outcome valuation regarding the frequency each outcome was considered to be important was identical across instruments. The valuations of the most important outcome as defined in the ranking instrument were more similar to the valuations of the rating instrument than to the point distribution instrument, with 78% of respondents saying that their number one outcome as defined in the ranking instrument was very important in rating instrument. 51% of men attributed the highest point score in the point distribution instrument to their number one outcome as defined in the ranking instrument. In addition, differences in the perceived usefulness ($p = 0.006$ and $p = 0.001$ for clarification and communication usefulness items, respectively) and ease of use ($p = 0.018$ and $p = 0.009$ for clearness and ease to complete items, respectively) between ranking, rating and point distributions instruments were found. The rating instrument was perceived as being more useful and easier to use than the ranking ($p = 0.025$, $p = 0.007$, and $p = 0.036$ for the clarification usefulness, communication usefulness and ease to complete items, respectively) and point distribution instruments ($p = 0.04$, $p = 0.005$, $p = 0.012$, and $p = 0.001$ for the clarification usefulness, communication usefulness, clearness and ease to complete items, respectively), when adjusted for the demographic variables. A single statistically significant difference was found between the ranking and the point distribution instruments in the clearness item ($p = 0.008$). There was a statistically significant difference in the instruments' time to completion ($p < 0.001$), with respondents being quicker to complete the rating instrument and slower to complete the point distribution instrument. Furthermore, individuals with college degrees were less likely to attribute higher usefulness and ease of use scores to the instruments when compared with individuals with lower educational status ($p < 0.05$). Family history of prostate cancer

and previous PSA testing did not seem to be related with the likelihood of perceiving instruments as more useful and/or easy to use.

Conclusion

The mechanisms underlying each instrument influence the instruments' output in terms of outcome valuation**.** The rating instrument seems to be perceived by this sample as the most useful and easy to use instrument when compared to the ranking and point distribution instruments. It is also the instrument that respondents are quicker to complete. These results suggest that the rating instrument may be preferred to implement in daily clinical practice. Further research studying a more representative sample of the Portuguese population is needed.

Table of Contents

1 Introduction

Every healthcare system strives for delivering better quality care to the population it serves. Health outcomes measurement and patient-centered care are two concepts that have been frequently mentioned as means for achieving greater quality levels in healthcare systems (1–4). The combination of both concepts produces a goal-oriented framework based on patient empowerment, thus replacing the traditional medical model with a more individualized approach, in which the individual's goals and needs are the driving force of care (5). The patient's values provide the basis for this framework, which then involves the elicitation, operationalization and measurement of patient-identified goals (6). As not all individuals may share the same priorities, individualized outcome measures (IOMs) are used to obtain personal goals and their respective prioritization (7). The identification and hierarchization of individual goals are complex processes laden with difficulties (8-10). To ease the aforementioned processes, instruments of distinct natures are at patients' and physicians' disposal. As different tools reflect preferences differently (11), it is of paramount importance to understand what the impact of each instrument is on the outcomes of the process. At this point, a semantic distinction should be made regarding the meaning of "values" and "preferences". The former concept can be defined as "what matters to an individual in a given health decision", whereas the latter is associated with "the extent to which a decision option or health state is desirable" (12). That said, the individual's values can be considered an intrinsic factor that supports the goal-oriented framework and will be ultimately reflected on his/her preferences. Henceforth, these are the definitions adopted in this dissertation.

Being the most frequently diagnosed cancer among men in Europe (13), prostate cancer (PC) is an example par excellence of the importance of an individualized approach to care, even in more upstream decisions such as screening. Inasmuch as PC is characterized by slow progression and little clinical signs and symptoms for most men, careful weighting of the harms and benefits of PC screening is needed (14–16), namely the trade-off between a potential mortality reduction due to an early detection of PC and a considerably high risk of overdiagnosis and overtreatment with the associated side effects, such as impotence and incontinence (17). While the implementation of universalistic prostate-specific antigen (PSA) screening programmes remains a controversial topic among the medical community (18,19), the necessity of considering how individuals weight these harms and benefits is already contemplated in the American guidelines (20,21). On the other hand, the current European state-of-the-art guidelines mention the importance of adequate information flows, but do not explicitly include patients' values in the shared decision-making processes (13,22). Such regional differences may justify the fact that a greater share of work done on eliciting patients' screening preferences is performed in the former region (23–26). Along with differences in medical practices, there is evidence that individuals' perception of health, disease, and screening differs according to social and cultural circumstances (24,27).

Given the harms and benefits associated with PC screening, the decision to be screened lies beyond purely medical criteria. Hence, IOMs can assist this "preference sensitive decision" (17) by aiding patients to actively reflect on and prioritize the trade-offs involved with PC screening and communicate their preferences in daily clinical practice. Due to the multiplicity of IOMs strategies available it is crucial that providers know what to expect from each instrument. Therefore, studying how different instruments are used

and perceived by men in their preference elicitation process for PC screening is essential for the success of patient-centered decisions. Moreover, by investigating this apparently under-researched topic on a European population, the present study hopes to contribute to the current lively scientific discussion on PC screening and the role of patients in the decision process.

The objective of this study is to compare three different instruments – Rating, Ranking and Point Distribution – for the elicitation and prioritization of individual preferences for PSA screening in a Portuguese population. Through an online-based survey, a fourfold comparison is performed. Firstly, the extent to which different instruments result in different outcome valuation patterns is explored. Secondly, the differences between the instruments' perceived usefulness to clarify individual preferences and communicate those preferences and the perceived ease of use are analysed as well as differences in instruments' time to completion. Finally, the extent to which there is a relation between individuals' characteristics and their perception of instruments is studied.

For gathering insights on the aforementioned topic, the main question of this thesis is:

"What are the differences in the outcome valuation patterns, user perception and time to completion between Rating, Ranking and Point Distribution instruments for eliciting the preferences for prostate cancer screening among a general population of Portuguese men aged 40-75 years?"

This thesis' main research question can be branched into the following sub-questions:

- "What are the differences between the three instruments in the outcome valuation patterns?"
- "What are the differences between the three instruments in perceived usefulness?"
- "What are the differences between the three instruments in perceived ease of use?"
- "What is the difference in the time spent completing each instrument?"
- "How are individual characteristics related to the men's perception of the instruments?"

The remainder of the thesis is structured as follows. In the next chapter, the theoretical framework will explain the concepts used throughout this study. In chapter 3, the research methods will describe the experimental design and the data analysis procedures. The results of this study will be presented in chapter 4. These findings are then discussed in chapter 5. Additionally, a reflection on this study, its limitations and recommendations for future research and practice will be given.

2 Theoretical Framework

In this section, patient-centered medicine and individualized medical decisions are first addressed from a theoretical perspective. The concepts and frameworks mentioned below informed the present study design and clarified the positioning of the patient outside the traditional medical decision-making model. The role and characteristics of different types of individualized outcome measures are reviewed as well as the empirical evidence on this matter. The last section of this framework relates to PC screening, namely the current evidence that has been fuelling the discussion on the risk-benefit relation of this procedure and the studies assessing instruments for preferences elicitation within this therapeutic area. Information on the current Portuguese paradigm regarding PC screening is also presented.

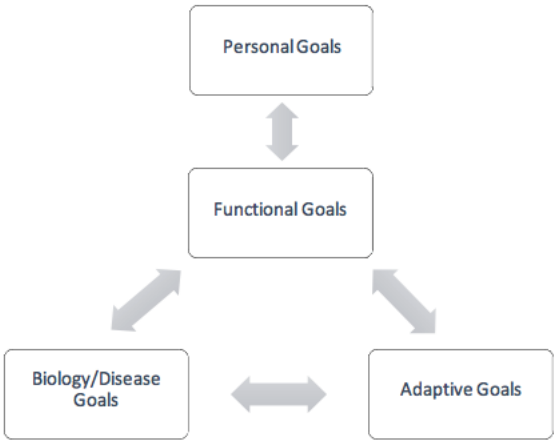
2.1 Individualized Medical Decisions

Each personalized decision that includes patients' values is part of cyclical patient-centered process, consisting of four different steps: 1) goal identification, 2) care planning, 3) care delivery and 4) goal evaluation (8). The sequence of cycles for a given patient is called the "individualized patient pathway" (8). As Berntsen et al. (9) pragmatically state, individualized care goals are "operationalizations of the general goal of promoting, restoring, and maintaining health". These operationalizations have different natures which ought to be considered throughout the process of care delivery:

1. Personal goals – individual constructions of health
2. Functional goals – desired functional abilities in social contexts
3. Biological goals – absence of biological malfunctions
4. Adaptive goals – social constructs of health

Following this logic, Berntsen et al. (9) designed a goal framework (Figure 1) based on ethics of authenticity, i.e., emphasis is put on patients' autonomy and singularity. According to this view, personal goals are at the top of the goal hierarchy and are supported by functional, biological and adaptive dimensions. Moreover, the individual is considered to be a self-determined actor capable of providing a "legitimate vision for future desired health states" (9). In PC, the decision to accept or decline screening will most likely be associated with the value men place in overarching personal goals, such as ruling out a potential PC diagnosis or avoiding complications from biopsies or subsequent treatments (28).

Figure 1 - Goal framework based on ethics of authenticity (9)



Similarly, Vermunt et al. (29) developed a three-goal model for patients. Using a physician's perspective three overarching goals were identified: disease- or symptom-specific, functional and fundamental. While the first two goals overlap with the aforementioned goal framework, the Dutch study broadened the scope of personal goals defined by Berntsen et al. (9) and included fundamental goals as the most comprehensive set of personal objectives, including values, core relationships and priorities in life (29).

By acknowledging the primacy of individual constructions of health and life and clarifying the inherent values, patients' preferences will be better reflected in the health care pathway decisions. This patient-centered approach in which the individual's awareness of his/her preferences exists has been shown to be associated with less decisional regret, improved communication flows with providers, and increased satisfaction (30).

Individual preferences are the result of the process of weighting the harms and benefits of a given intervention, which can be framed from a search for dominance perspective (31–33). According to Montgomery's Search for Dominance Structure (SDS) model, after assessing the respective advantages and disadvantages, an individual reaches dominance when they perceive a given intervention as being superior (and by inherence not inferior) in its attributes (31,32,34). The mechanisms through which these preferences can be clarified range from more passive approaches to more interactive approaches (32,35). The former set of instruments involve presenting the pros and cons associated with a given option and let patients progressively realize what their preferred option is (17,35). With explicit instruments patients are actively involved and asked to compare the relative importance of a given set of attributes (17). Discrete choice experiments, rating scales, ranking and point distribution instruments are examples of explicit instruments. The first instrument involves asking individuals to make choices from a set of alternatives, constituted by predefined attributes and levels (36). A rating exercise involves asking the individual to specify the importance of an attribute on a visual analogue scale or Likert scale (36), whereas a ranking instrument involves asking the individual to order attributes in terms of their relative importance (36). In a point distribution instrument, the respondent is asked to distribute a limited number of points by a set of attributes, according to their importance (36). Despite having been used in prior studies from different specialties (7,23,25,26,37–40), evidence on their use and respective patient perception is still scarce (41)

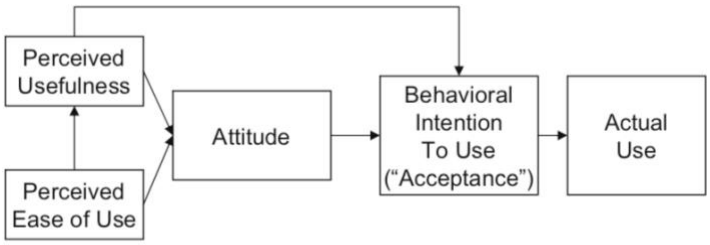
The extent to which a given instrument is applicable to a clinical practice environment is dependent on a number of factors. Both the complexity of the instrument and the added time required to implement it are considered to be the most limiting aspects which hamper the execution of a more personalized approach to care (42). That being said, discrete choice experiments are time-consuming and may not measure values and preferences of an individual patient as accurately as other explicit methods due to the need of larger number of choice tasks from an individual patient, which renders implementation in daily clinical practice more difficult (36).

The qualitative study by van Deen et al. (7) reveals that patients acknowledge usability differences between instruments. The ranking instrument was perceived to be easy to complete as it allowed a straightforward prioritization of outcomes, although it did not allow participants to attribute the same weight to outcomes of equal importance (7). Conversely, the rating instrument does not impose a relative prioritization of outcomes,

allowing patients to rate outcomes identically (7,35). This instrument was also considered to be easy to complete by the study participants (7). The point distribution was reported as being the most challenging instrument, with some patients referring to it as "cognitively difficult" (7). However, it allowed patients to discern with precision the personal importance of each outcome (7). Different perceptions of the instruments can arise according to the facility with which patients can use those instruments as well as the extent to which they find them useful for clarifying values and communicating their preferences to the provider. These practical characteristics of IOMs can be translated into two overarching concepts derived from the technology acceptance model (TAM): usefulness and easiness of use (43,44).

Originally devised in the 1980's, the TAM built on the Theory of Reasoned Action (45) and aimed at increasing the adoption of information technologies (ITs) within the workplace, by assessing individuals' future intentions to use these technologies (43,44). Since its creation, the model has been subject to several transformations, with variables removed or added in order to capture different determinants of use (46–48). According to the original model (Figure 2), attitude, i.e., the individual's set of positive or negative feelings towards a behaviour, is determined by his/her salient beliefs about the instrument (45). These beliefs, also called constructs, include the perceived usefulness and perceived ease of use and are the most distal determinants of the actual use of an instrument. Usefulness relates to the individual's perception that using a given instrument will enhance job performance, and ease of use is related to the perception that using the instrument will be effortless (43,44). Both have been shown to be related with the attitude towards a given instrument and its acceptance and subsequent use (49).

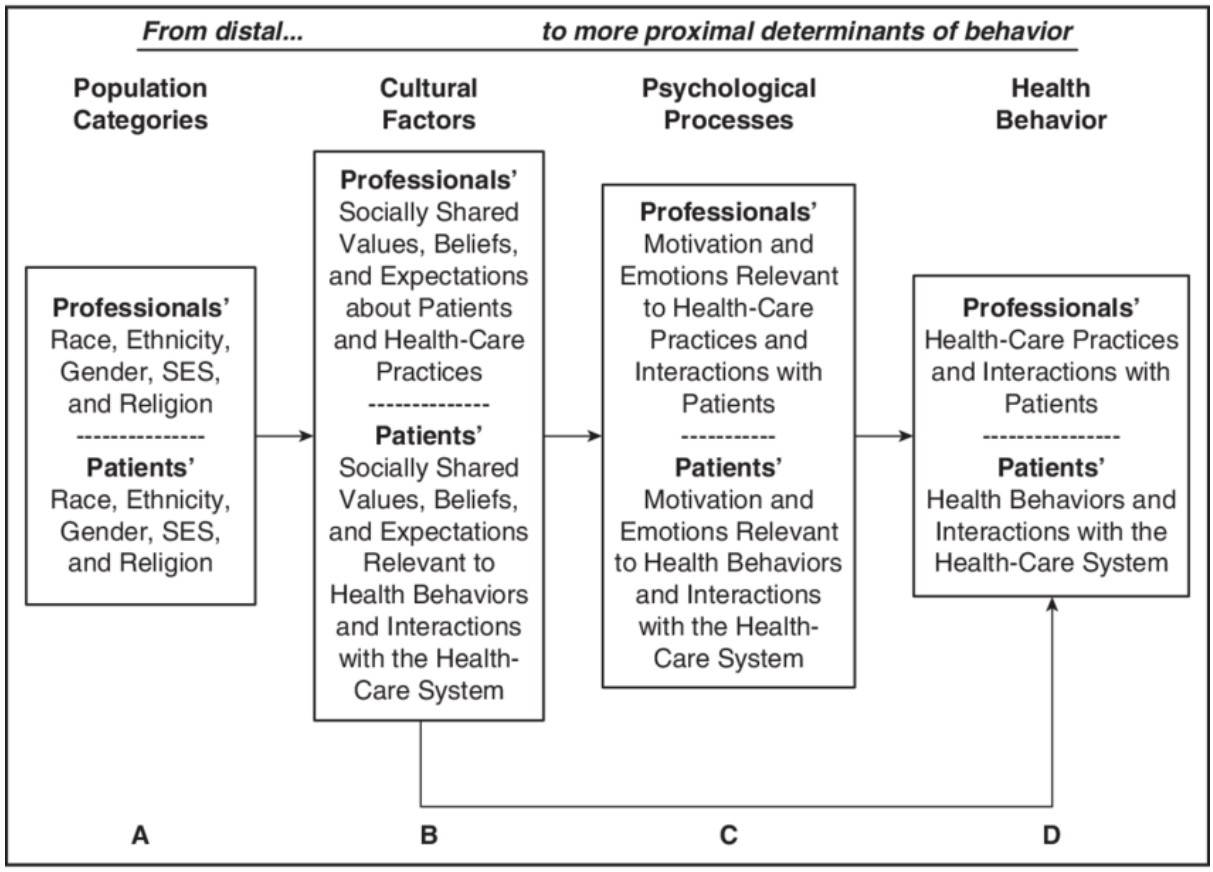
Figure 2 - Technology Acceptance Model [Adapted from (44)]



The TAM and its derivations have been widely used in healthcare as the theoretical base for assessing health ITs acceptance. Holden & Karsh (49) provide a comprehensive review of studies using TAM in healthcare, discerning different definitions and measurement dimensions of the models' constructs used.

Acknowledging the individual as a self-determined actor also implies considering his/her circumstance. The different determinants of health behaviour and their relations were studied by Betancourt & Flynn (37) and brought together in a conceptual model (Figure 3).

Figure 3 – Betancourt's model of culture, psychological processes, and behaviour adapted for the study of health behaviour [adapted from (37)]



The determinants of health identified in Figure 3 are organized from more distal to more proximal according to their potential impact on health behaviour. Psychological processes, such as emotions and motivations, are deemed as the most influential factors of health behaviour. Culture, defined as "value orientations, beliefs, and norms that are socially shared among individuals", is considered to play a role on behaviour either directly or indirectly through psychological processes. These cultural aspects are in turn influenced by social characteristics such as race or socioeconomic status. These relations have been empirically studied in a cancer screening context, either in their impact on the individual's decision or in the instruments used in the decision process itself (24,27).

While the work by van Deen et al. (7) provides useful insights into the patients' perception of instruments, at the time of writing of this dissertation, no quantitative work exploring the relation between the individual and his/her perception of the instrument was found.

2.2 Prostate Cancer Screening

PSA testing for PC screening is a controversial topic within the medical community. A recent European study associated PSA testing with reductions in advanced disease and PC-related mortality and advocated for a risk-stratified strategy for PC screening (52). However, the fact that PC often is an indolent disease which tendentially does not result in complications in the majority of cases is mentioned as the main reason against universalistic PSA screening programmes (16). Moreover, the harms associated with overdiagnosis (i.e., unnecessary biopsies) and subsequent treatment (i.e., sexual and

urinary disfunctions) counterbalance the reported mortality reductions (28). Tikinnen et al. (28) report the results of studies comparing the harms and benefits of PC screening vs. "no screening". Screening was associated with a reduction of PC mortality (reduction of 1 case per 1000 people) and a reduction of the incidence of advanced cancer (reduction of 3 cases per 1000 people) throughout a 10 year-period. On the other hand, "no screening" was associated with a reduction of the incidence of cancer (any stage) as well as localized cancer (reduction of 18 events and 14 events per 1000 people, respectively). Naturally, "no screening" was also associated with fewer biopsy-related complications, such as blood in semen, pain, fever or blood in urine as well as fewer cancer-treatment complications, such incontinence or impotence. Hoffman & Del Mar (53) report that men tend to be misinformed about PC screening and overestimate the benefits of the procedure. The uncertainty surrounding the necessity for PC screening together with the existent variability among men's values and preferences render shared-decision making necessary for men considering screening (28).

With an estimated incidence rate for PC of 135.7 cases per 100,000 population, Portugal is the 15th European country with the lowest incidence rate of PC, lower than the EU average (150.5 cases per 100,000 population) (54). The latest recommendation issued by the Portuguese Directorate-General for Health (*Direção Geral da Saúde*), the governmental regulatory body of technical aspects of healthcare, state that the determination of PSA should not have a universalistic character. Moreover, opportunistic screening is recommended only for a population of men between 50 and 75 years (55). A cross-sectional study has shown that, among Portuguese men, 67.3% believe they should undergo a PSA evaluation every 14.7 months (56). As far as general practitioners (GPs) are concerned, 61% of Portuguese GPs order PSA tests annually (57). Braga et al. (58) have estimated a prevalence of prostate cancer screening in Portuguese men between 40 and 79 years of 44.2% [95% confidence interval (CI): 37.5–51.0]. The authors speculate that this figure may be explained by the fact that nine out of every ten men perceive PC as a cancer for which screening is recommended (59) and by the high prevalence of PC, resulting in an increased sharing of experiences between PC survivors and their acquaintances as well as high attention from media (60). In light of the increasing awareness of the importance of the preference sensitive character of a PC screening decision in Portugal, Baptista et al. (61) developed the first translated and culturally adapted decision aid to support Portuguese men's decision concerning PC screening.

Empirical research has been performed on the area of individual values elicitation for PC screening (16,23,24,57–59). In Portugal, Baptista et al. (17) are conducting a study in a Portuguese men population to compare the perceived clarity of values using a decision aid with an explicit instrument vs. an implicit instrument vs. no instrument. Recruitment for this study is ongoing (65). In line with the aforementioned harms and benefits, Pignone et al. (25) studied the effects of three different instruments on the preference for PC screening in an American and Australian population. To this end, a set of four attributes that the authors considered relevant for the topic were derived from the literature and own clinical experience, namely: effect on prostate cancer mortality, risk of biopsy, risk of being diagnosed with prostate cancer, and risk of becoming impotent or incontinent as a result of treatment.

3 Research Methods

In this section, the data collection and analysis methods are explained. The section starts with a description of the study target population. Thereafter, the survey design and structure are outlined. Lastly, the statistical methods used for the data analysis are described.

3.1 Study Population

The latest recommendation issued by the Portuguese Directorate-General for Health (*Direção Geral da Saúde*) on prostate cancer screening recommends screening of Portuguese men aged 50 to 75 years (55). Given the lockdown imposed in Portugal at time of recruitment, a fallback plan was devised if the recruitment of sufficient respondents was not feasible in the time window available. That said, the age criterion was broadened, and Portuguese men aged 40 to 75 were eligible to participate in the project. Excluded from participating were men with previous diagnosis of PC and men unable to understand written Portuguese.

3.2 Data Collection

A quantitative exploratory within-subject study was performed through an online web-based questionnaire constructed using Qualtrics® (66). The questionnaire was first tested with a group of 7 men, eligible to participate, in three different rounds of approximately 2 persons each. Feedback was processed between the rounds to optimize the pilot phase. A Portuguese physician with research experience (67) was consulted throughout the pilot phase to ensure the survey's clinical correctness. After the pilot, the survey was distributed using snowball and convenience sampling techniques. Through direct contacts and social media, the author's acquaintances were approached to fill in the questionnaire and were asked to distribute the study to their acquaintances.

The questionnaire used for this study consisted of six parts (see Appendix I for a translated version of the survey):

1. Introduction
2. Informed Consent
3. Demographic Characteristics
4. Information on PC Screening
5. Value Elicitation Instruments & Usefulness and Ease of Use Assessments
6. Box to give the respondents an opportunity to leave a comment

All questions were close ended apart from the optional comment questions. No forced responses were required to proceed with the survey. However, the system would flag any incomplete sections before proceeding.

3.2.1 Introduction & Informed Consent

The questionnaire started with a short introduction and two questions to exclude respondents who did not fit the age criterion and/or who had or have had PC. After this

section, the respondents eligible to participate were shown the informed consent form. Respondents were informed about the survey's purpose and the anonymity of their data. Only respondents who gave their informed consent were able to proceed with the questionnaire.

3.2.2 Demographic Characteristics

This section referred to the respondents' personal information. In line with the theoretical framework (27) and previous studies on this topic (17,18,24,25), the following demographics were included: age, ethnicity, employment status, educational level, income, family history of cancer, and prior PSA testing. The categories present in the ethnicity question were set in accordance with the latest recommendations issued by the Portuguese 2021 national census workgroup (68).

3.2.3 Information on Prostate Cancer Screening

In this section, general information on PC as well as the risks and benefits of PC screening were presented. The content of this section was derived from an approved and validated PC screening decision aid tool (61), so to ensure the applicability and adequacy of the information provided to respondents. The authors were contacted and permission to use this information on the questionnaire was granted. All respondents were provided with the same information.

3.2.4 Value Elicitation Instruments & Usefulness and Ease of Use Assessments

Value Elicitation Instruments

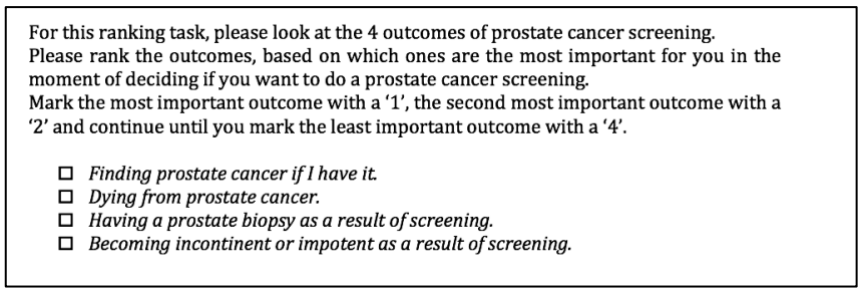
In order to answer the first sub-question regarding the differences between the three instruments in the outcome valuation patterns, respondents were presented with three different value elicitation instruments comprising a ranking (Figure 3), a rating (Figure 4) and a point distribution (Figure 5) instruments. All respondents completed the three instruments. The instruments have been previously validated and used with the same purpose in research from different medical specialties (7,23–25). In each instrument, respondents were asked to reflect on four outcomes of PC screening:

- *Finding prostate cancer if I have it.*
- *Dying from prostate cancer.*
- *Having a prostate biopsy as a result of screening.*
- *Becoming impotent or incontinent as a result of an eventual prostate cancer treatment.*

These outcomes are overarching personal goals derived and adapted from previous similar studies performed in the United States of America, Australia and the Netherlands (24,25,64). Additionally, to assure that the outcomes were also applicable to a Portuguese population, a conversation with Dr. Sofia Baptista, a Portuguese specialist on shared decision making in the field of PC screening in Portugal (17,61,69), occurred. Both the order by which rating, ranking and point distribution instruments appeared in the questionnaire and the outcomes in each instrument were randomized to limit any potential order bias. In order to answer the fourth sub-question regarding the differences

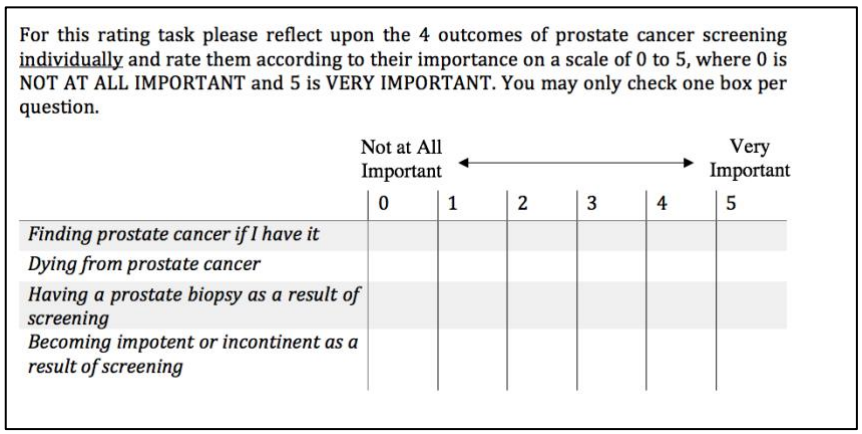
between instruments in time to completion the time spent on each instrument was also registered.

Figure 3 - Ranking Instrument


For this ranking task, please look at the 4 outcomes of prostate cancer screening.
Please rank the outcomes, based on which ones are the most important for you in the moment of deciding if you want to do a prostate cancer screening.
Mark the most important outcome with a '1', the second most important outcome with a '2' and continue until you mark the least important outcome with a '4'.

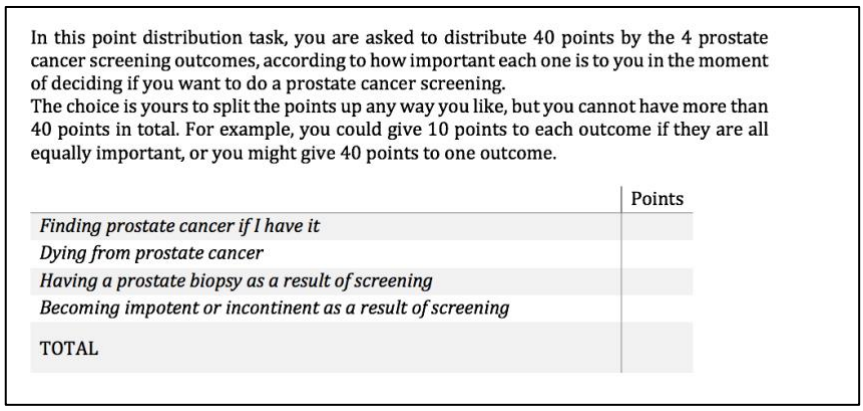
- ☐ *Finding prostate cancer if I have it.*
- ☐ *Dying from prostate cancer.*
- ☐ *Having a prostate biopsy as a result of screening.*
- ☐ *Becoming incontinent or impotent as a result of screening.*

Figure 4 - Rating Instrument


For this rating task please reflect upon the 4 outcomes of prostate cancer screening individually and rate them according to their importance on a scale of 0 to 5, where 0 is NOT AT ALL IMPORTANT and 5 is VERY IMPORTANT. You may only check one box per question.

	Not at All Important ← 0	1	2	3	4	5 → Very Important
Finding prostate cancer if I have it						
Dying from prostate cancer						
Having a prostate biopsy as a result of screening						
Becoming impotent or incontinent as a result of screening						

Figure 5 - Point Distribution Instrument


In this point distribution task, you are asked to distribute 40 points by the 4 prostate cancer screening outcomes, according to how important each one is to you in the moment of deciding if you want to do a prostate cancer screening.
The choice is yours to split the points up any way you like, but you cannot have more than 40 points in total. For example, you could give 10 points to each outcome if they are all equally important, or you might give 40 points to one outcome.

	Points
Finding prostate cancer if I have it	
Dying from prostate cancer	
Having a prostate biopsy as a result of screening	
Becoming impotent or incontinent as a result of screening	
TOTAL	

Usefulness and Ease of Use Assessments

In order to answer the second and third sub-questions regarding the differences between instruments in perceived usefulness and ease of use, respondents were asked to assess the instrument in two constructs: usefulness and ease of use. After completing each instrument, respondents rated four items (two for each construct) on a seven-point Likert scale from "*Strongly Disagree*" to "*Strongly Agree"*. The items of each construct were designed based on the TAM (43,44) as well as on conversations with Dr. Welmoed van Deen and Dr. Sofia Baptista (Table 1). The usefulness construct was defined based on the instrument's self-clarification and communication purposes, i.e., the extent to which

the instrument helped the individual clarifying his preferences for participating in prostate cancer screening as well as communicating those preferences to his physician. The ease of use construct referred to the usability of the instrument itself, namely the intelligibility of the instrument and the ease to complete it.

Table 1 - Usefulness and Ease of Use Items

Construct	Item
Usefulness	*This task would help me to clarify whether I would like to participate in prostate cancer screening.*
	This task would help me to communicate my preferences for prostate cancer screening to my doctor.
Ease of Use	*The task is clear and understandable.*
	The task is easy to complete.

Each usefulness and ease of use assessment ended with an optional open-ended question allowing respondents to leave a comment about the instrument concerned.

3.3 Data Analysis

The research data was processed with IBM® SPSS® Statistics 27.0.

A population-level descriptive analysis was performed to answer the first sub-question about the differences in outcome valuation patterns. Firstly, these patterns were studied per instrument. Thereafter, the patterns were studied between the most important outcome as defined in the ranking instrument and the remainder of the instruments, thus using the ranking task as the anchor point for comparison. Regarding the rating task, the proportion of men who rated their number one outcome with a "*5 – Very Important*" was analysed per outcome and overall. For the point distribution task, the patterns in point distribution to the most important outcome were investigated, namely the proportion of men who attributed the highest point score to their most important outcome.

To analyse the respondents' perception of instruments as well each instrument's mean time to completion, basic descriptive statistics were used. Cronbach's Alpha (α) was calculated to measure the internal consistency of the Likert-scales measuring the same construct within an instrument. In order to answer the second and third sub-questions regarding the differences between the three instruments in the usefulness and ease of use ratings, scores were assessed using the non-parametric Friedman test. For each Friedman test performed, descriptive statistics and the Friedman test statistics were registered. An overall statistically significant difference in the usefulness and ease of use scores between instruments was accepted if $p < 0.05$, while recognizing that tests of significance are approximations that serve to aid interpretation and inference. The Friedman test only reveals whether there is an overall statistically significant difference between instruments (70). Hence, when the result of the Friedman test pointed towards an overall statistically significant difference in the usefulness and ease of use scores between the three instruments, a post-hoc analysis was performed to examine where the

differences occurred. This analysis entailed running three separate Wilcoxon matched-pairs signed-rank test on the different combinations of the related groups. Strictly speaking, the Wilcoxon matched-pairs signed-rank test was originally designed for dealing with numerical data. However, it has been frequently used with ordinal data as well (71). As such analysis involved performing multiple comparisons, a Bonferroni adjustment was performed, in order to diminish the likelihood of declaring a result as statistically significant when it was not (Type I error). Any difference between instruments in the post-hoc analysis was then considered to be statistically significant when $p < 0.017$, i.e., the original statistical significance threshold was adjusted by dividing the original statistical significance level of 0.05 by 3, the number of comparisons being made. The same procedure was used to answer the fourth sub-question about the differences between tests in the time to completion, measured in seconds. Normality was assessed with the Shapiro-Wilk test.

Ordinal logistic regressions were used to answer the sub-question about the relation between individual characteristics, the type of instrument and men's perception of instruments. Given the within-subject nature of the present study, repeated measures of the same constructs had to be taken into account when estimating regression parameters. That said, the generalized estimation equation method (GEE) was used to obtain the regressions estimates. GEE is an extension of generalized linear models that accounts for the correlation of within-subject responses for the response variables, producing more unbiased regression estimates (72,73). By not considering correlation of responses, one would be at risk of incorrectly estimating regression model parameters, leading to potentially incorrect conclusions. The fact that GEE is also applicable to the analysis of non-normally distributed response variables is a noteworthy advantage of this method (72,73). The operationalization of GEE firstly consisted in defining a subject variable, i.e., a variable that defines subjects within the dataset. Given that the survey was completely anonymized, the response ID generated by Qualtrics® for each survey submitted was used as means to identify respondents. As every respondent completed and rated each instrument in terms of its usefulness and ease of use, the instrument (ranking, rating, or point distribution) was used as a within-subject variable, thus allowing the identification of each subject's measurement of each instrument. An ordinal logistic model was selected to estimate the associations between demographic variables, type of instrument and each item of the usefulness and ease of use constructs. Five independent variables were included: age, educational level, family history of prostate cancer, previous PSA testing, and the type of instrument. Because few respondents were aged between 55 and 75 years, these age categories were combined in one [55-75] years category. Following a similar rationale, educational levels comprising *high school* graduate and *lower than high school graduate* were combined in one *lower than college graduate* category. Hence, educational level was described as *college graduate* or *lower than college graduate.* Family history of cancer was described as *Yes* or *No.* The last PSA testing variable described in the demographics was transformed into previous PSA testing. The categories *More than 1 year ago* and *Less than 1 year ago* were combined into a *Yes* category and the categories *Never* and *Don't know* were combined into a *No* category. The model fitting information is reported using the likelihood-ratio test. This test compares the log-likelihoods of the full model and the intercept-only model, i.e., the model in which no independent variables are included. If the difference between the log-likelihood of both models (represented by the chi-square statistic) is statistically significant, then the final model is said to fit better the data than the intercept-only model (74,75). Odds Ratios

(ORs), significance levels and 95% confidence intervals (CIs) are reported. Statistical significance was set at $p < 0.05$. In order to enable a full comparison between instruments, a second regression model was performed for each usefulness and ease of use item using a different instrument as the reference category. The remainder of the variables and all regression specifications and procedures were maintained.

The present study also had qualitative sections in which respondents could comment on the different instruments or the survey. A thematic analysis (76) was performed, using an inductive approach, where the observations made by respondents were generalized based on previous knowledge. Hence, comments were grouped in key themes and, when applicable, sub-themes, most of them overlapping with topics described in the literature and mentioned in the introduction and theoretical framework of the present dissertation.

4 Results

In this section, the results of the analysis are presented. Firstly, a description of the respondents' demographic profile is outlined. Secondly, the valuation patterns are described through a descriptive analysis using the ranking task as the anchor point for comparison. Thirdly, the differences in the usefulness and ease of use items are presented. Thereafter, the results of the logistical ordinal regressions aiming at analysing the association between individual characteristics, type of instrument and the perception of instruments are detailed per item. Finally, the comments left on the open-ended questions are presented by theme and sub-theme, when applicable.

4.1 Descriptive Statistics

A total of 170 respondents submitted the questionnaire. Because no forced answers existed, there was the possibility that incomplete questionnaires were submitted. This resulted in the exclusion of 49 respondents (28.8%) whose questionnaires were incomplete. 2 respondents (1.1%) answered negatively to the informed consent section and were also excluded. The final sample for analysis therefore consisted of 119 observations (70%). The demographic characteristics of the sample are described in Table 2. All respondents were white men. 82% of respondents were aged between 40 and 54 years. The majority of respondents were employed (93.3%). There is also a preponderance of highly educated respondents, as almost 81% of men had a college degree. 85.2% of respondents earned an above average salary. An identic share of respondents did not have any first-degree family history of PC. As far as PSA testing is concerned, 47.1% of respondents reported never have performed any PSA screening exam. 23.5% of respondents reported that their last PSA test was performed in the past year.

Table 2. Demographic characteristics of the study population (n=119).

		N (%)
Age		
	[40-44]	17 (14.3)
	[45-49]	48 (40.3)
	[50-54]	33 (27.7)
	[55-59]	9 (7.6)
	[60-64]	7 (5.9)
	[65-69]	3 (2.5)
	[70-75]	2 (1.7)
Ethnicity		
	White / European origin	119 (100)
	Black / Afro descendent	-
	Asian	-
	Other	-
Employment Status		
	Employed	111 (93.3)
	Unemployed	2 (1.7)
	Retired	6 (5)
Educational Level		
	College graduate	96 (80.7)
	High school graduate	19 (16)
	Less than high school graduate	4 (3.4)
Income		
	Below average (less than 970€ / month)	11 (9.2)
	Above average (more than 970€ / month)	102 (85.7)
	Prefer not to answer	6 (5)
Family History of Prostate Cancer		
	Yes	17 (14.3)
	No	102 (85.7)
Last PSA Testing		
	Less than 1 year ago	28 (23.5)
	More than 1 year ago	29 (24.4)
	Never	56 (47.1)
	Don't know	6 (5)

4.2 Descriptive Analysis of Outcomes Valuation Patterns

This section contains the results of the descriptive analysis performed to understand the differences between instruments in the outcomes valuation patterns. Firstly, a general analysis per instrument is presented. Thereafter, a comparative analysis using the most important outcome as defined in the ranking instrument is presented.

In the ranking task, men were asked to explicitly prioritize the PC screening outcomes. As shown in Table 3, dying from prostate cancer was the most frequently number one ranked outcome, with almost half of the men (47.8%) stating that this was the most

important outcome in the moment of deciding whether to undergo PC screening. "*Finding prostate cancer if I have it*" was reported by almost 22% of men as being the most important outcome to consider for a PC screening decision. "*Having a prostate biopsy as a result of screening*" and "*Becoming incontinent or impotent as a result of a potential treatment*" were deemed as the most important outcome in the ranking task by the same proportion of men (15.1%).

In the ranking task, men were asked to rate each outcome from "*0 – Not At All Important*" to "*5 – Very Important*". No man rated every outcome with the same number. 47 men (39.4%) rated the outcomes as *"4 – Important"* or *"5 – Very Important"*. "*Dying from prostate cancer"* was the outcome which was more frequently rated as very important (73.9%), while "*Having a prostate biopsy as a result of screening*" was only perceived as very important by 31% of men. Almost 60% of men rated the outcome ""*Finding prostate cancer if I have it*" as very important. The outcome "*Becoming incontinent or impotent as a result of a potential treatment*" was perceived as very important by 51.2% of respondents.

In the point distribution task, men were asked to distribute 40 points by the four outcomes according to the importance each outcome had to them. 14% (117/119) of men attributed 10 points to all four outcomes, thus considering them equally important in the moment of deciding whether to undergo PC screening. In a similar fashion to the other instruments, "*Dying from prostate cancer"* was the outcome considered to be the most important in the point distribution instrument - 35.3% of respondents attributed the highest point score to this outcome. "*Finding prostate cancer if I have it*" was attributed the highest point score by 19.3% of men, whereas that figure decreased to 9.2% for the outcome "*Becoming incontinent or impotent as a result of a potential treatment*". "*Having a prostate biopsy as a result of screening"* was explicitly considered the most important outcome in the point distribution instrument by 6.7% of men.

Overall, Table 3 reveals that respondents produced an identical outcome valuation pattern across the three instruments, when considering the frequency each outcome was considered to be very important or the most important. The outcome "*Dying from prostate cancer*" was more frequently perceived as being the most important outcome in the ranking, rating and point distribution instruments, followed by "*Finding prostate cancer if I have it".* The possibility of having a biopsy following a positive result of the PSA exam was the outcome that was less frequently considered as being the most important in the moment of the PC screening decision.

Table 3 – Patterns of valuation in the Ranking, Rating, and Point Distribution instruments

	Instrument		
	Ranking	Rating	Point Distribution
	Number of times ranked as the most important outcome, N (%)	Number of times rated as a very important outcome, N (%)	Number of times with the highest point score, N (%)
Finding prostate cancer if I have it.	26 (21.8)	71 (59.6)	23 (19.3)
Dying from prostate cancer.	57 (47.9)	88 (73.9)	42 (35.3)
Having a prostate biopsy as a result of screening.	18 (15.1)	37 (31)	8 (6.7)
Becoming incontinent or impotent as a result of a potential treatment.	18 (15.1)	61 (51.2)	11 (9.2)

A comparison of the outcome valuation patterns of the ranking vs. rating instruments and ranking vs. point distribution instruments is depicted in Table 4.

For the comparison of the ranking instrument vs. rating instrument, the proportion of respondents who rated their number 1 outcome (as defined in the ranking task) with a "*5 – Very Important*" is analysed. 89.5 % (51/57) of men who ranked "*Dying from prostate cancer*" as the most important outcome, rated it with a "5" in the rating task, i.e., considered it a very important outcome. Almost 77.8% (14/18) of men who considered the outcome "*Becoming incontinent or impotent as a result of a potential treatment"* to be the most important in the ranking task, also stated that this outcome was very important in the rating task. The outcome "*Finding prostate cancer if I have it*" was rated as very important by 65.4% (17/26) of the men who ranked it first, while only 61.1% (11/18) of men that ranked "*Having a prostate biopsy as a result of screening*" first had the same behaviour. Overall, 78% of men rated their number one outcome with a "*5 – Very Important*", thus revealing that the majority of men acknowledge the outcome that is the most important to them as being very important.

The proportion of respondents who allocated the highest point score to the most important outcome (as defined in the ranking task) is described as means of comparing the ranking instrument with the point distribution instrument. 61.4% (35/57) of men who ranked "*Dying from prostate cancer*" first, attributed it the highest point score. For the outcome "*Having a prostate biopsy as a result of screening*", 27.7% (5/18) of men who ranked it first attributed it the highest score. Half of the men (13/26) who reported the outcome "*Finding prostate cancer if I have it*" as the most important outcome on the ranking task also attributed it the highest point score. 44.4% (8/18) of respondents that considered "*Becoming incontinent or impotent as a result of a potential treatment*" the most important outcome attributed it the highest point score. Overall, 51% of respondents attributed the highest score to their most important outcome. Such figure reveals an apparent incongruence between these two instruments, as barely half of respondents explicitly considered their number one outcome as the most important outcome in the point distribution instrument.

In general, the data depicted in Table 4 shows that the valuation of the most important outcome as defined in the ranking instrument appears to be more similar in the rating instrument than in the point distribution instrument. This difference is particularly noticeable for three of the four outcomes assessed.

Table 4 - Comparison of outcome valuation patterns
Ranking vs. Rating Task reports the proportion of respondents who rated their number one outcome (as defined in the ranking task) with a "*5 – Very Important*".
Ranking vs. Point Distribution reports the proportion of respondents who allocated the highest point score to their most important outcome (as defined in the ranking task).

	Ranking vs. Rating	**Ranking vs. Point Distribution**
	Most important outcome rated with a "*5 - Very Importa*nt", N (%)	Highest point score to the most important outcome, N (%)
Finding prostate cancer if I have it.	17 (65.4)	13 (50)
Dying from prostate cancer.	51 (89.5)	35 (61.4)
Having a prostate biopsy as a result of screening.	11 (61.1)	5 (27.7)
Becoming incontinent or impotent as a result of a potential treatment.	14 (77.8)	8 (44.4)

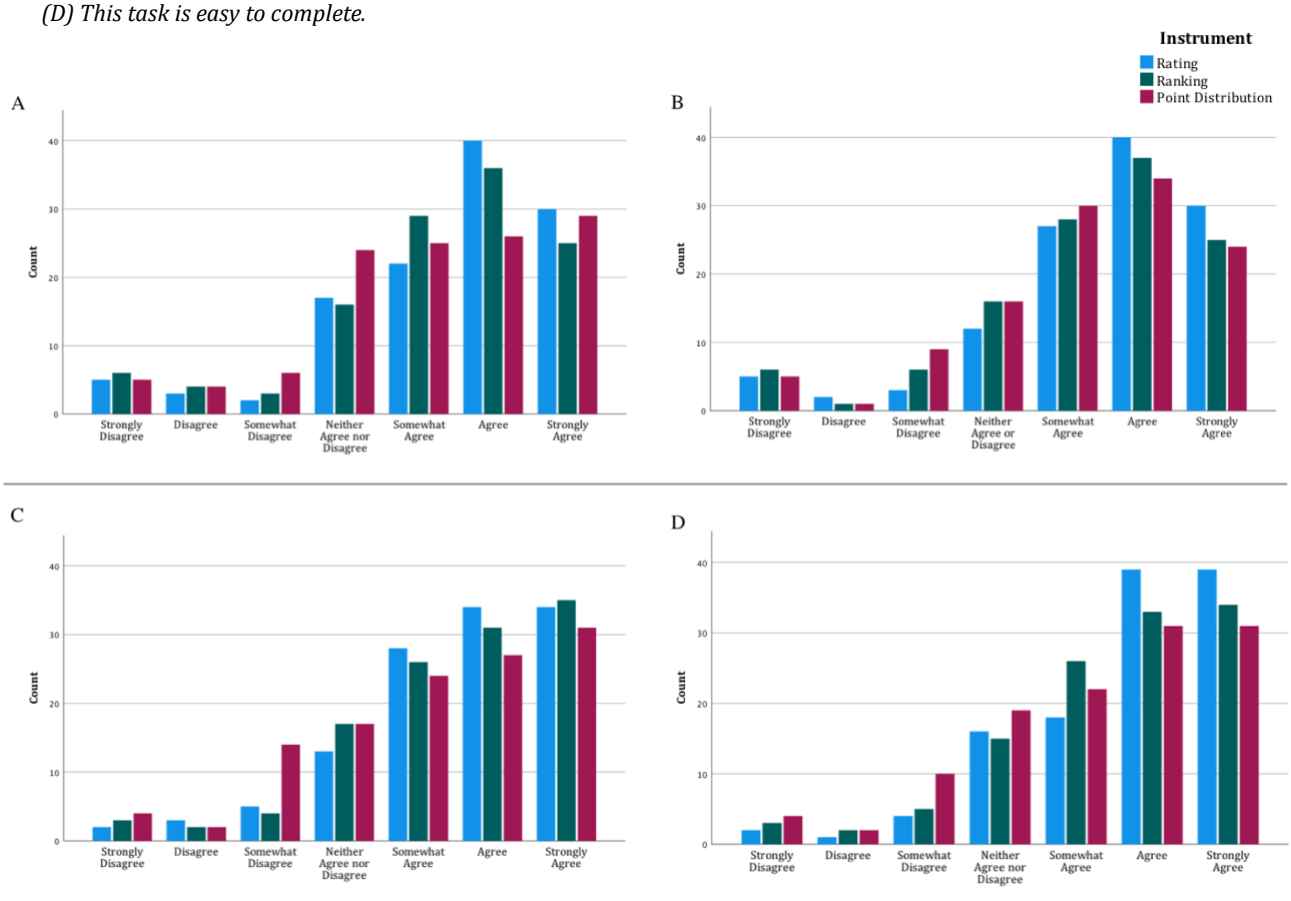
4.3 User Perception of Instruments

This section contains the results of the comparative analysis performed to understand whether there are differences between instruments in their usefulness and ease of use. Firstly, a descriptive analysis of the distribution of the usefulness and ease of use items per instrument is performed. Thereafter, instruments are compared in terms of their usefulness, namely their clarification and communication usefulness. Finally, the instruments are compared in terms of their clearness and easiness to complete.

After completing a given instrument, respondents were asked to rate the instrument in terms of its usefulness and ease of use on a seven-point Likert scale from "*1 – Strongly Disagree*" to "*7 – Strongly Agree*". The frequencies of answers for each item for the rating, ranking and point distribution instruments can be found, respectively, in Appendix II, III and IV.

Both constructs – 'usefulness' and 'ease of use' – consisted of two items. Both scales had high levels of internal consistency within each of the three studied instruments, as determined by Cronbach's alpha values > 0.9.

Figure 6 - Acceptability scores distribution.
(A) This task would help me to clarify whether I would like to participate in prostate cancer screening.
(B) This task would help me to communicate my preferences for prostate cancer screening to my doctor.
(C) This task is clear and understandable.
(D) This task is easy to complete.



As depicted in Figure 6, the majority of answers for all items is concentrated in the upper part of the seven-point Likert scale, i.e., above the middle point ("*4 – Neither Agree nor Disagree"*). Few respondents rated the instruments' perceived usefulness and ease of use negatively.

For the clarification usefulness item, represented by A in figure 6, 58.8% of respondents agreed, "6", or strongly agreed, "7", that the rating instrument was useful for clarifying their preferences for prostate cancer screening decision, whereas 51% of men attributed the same ratings to the ranking instrument. The median score for both tasks is "*6 – Agree*". While one quarter of men (25.2%) strongly agreed with the clarification usefulness of the point distribution instrument, the scores for this instrument were more dispersed. In fact, 20.2% of men neither agreed nor disagreed with the statement (vs. 14.3% and 13.4% in the rating and ranking instruments, respectively). The median score for the clarification usefulness of the point distribution instrument is "*5 – Somewhat Agree*".

The communication usefulness item, represented by B in figure 6, followed a similar trend to item A, with option "*6 – Agree*" being the most frequently chosen option. Both the rating and ranking instruments had a similar distribution of scores, with most participants agreeing or strongly agreeing with the instruments' communication usefulness (58.8% for the rating instrument vs. 52.1% for the ranking instrument). The same scores ("6" or "7") were attributed to the point distribution instrument by 48.8% of men. The median score for both instruments is "*6 – Agree*". Again, the point distribution instrument scored a median score of "*5 – Somewhat Agree*", with 38.6% of men attributing a score of "*5 – Somewhat Agree*" or "*4 – Neither Agree nor Disagree*" to the communication usefulness of this task.

With regards to the clearness item, represented by C in Figure 6, the higher share of scores tended to the extreme upper side of the Likert-scale. Most men rated the clearness of the rating and ranking instruments with a "6" or "7" (51.3% for the rating task vs. 56% for the ranking task). Almost 30% of men strongly agreed that the latter instrument was clear and understandable, being the instrument with the highest number of "7" ratings. The median clearness score for the rating and the ranking instruments was "*6 – Agree*". In a similar fashion to the aforementioned items, the median clearness score point distribution instrument was "*5 – Somewhat Agree*". 48.1% of men agreed or strongly agreed that this instrument was clear and understandable, while 14.3% of men neither agreed nor disagreed with the statement.

Similarly to the previous item, most men agreed or strongly agreed with the statement that the instruments were easy to complete of the instruments, as depicted in Figure 6D. The number of "6" and "7" ratings was identical between instruments, with 32.8% of respondents strongly agreed that the completing rating task was easy vs. 28.8% and 26.1% for ranking and point distribution task, respectively. The median easiness to complete score for the three instruments is "*6 – Agree*".

4.3.1 Usefulness

This task would help me to clarify whether I would like to participate in prostate cancer screening

The Friedman test showed that there was a statistically significant difference in the perceived clarification usefulness between the rating, the ranking and the point distribution instruments ($\chi^2(2) = 10.344$, $p = 0.006$). Post-hoc analysis using a Wilcoxon's matched-pairs signed-rank test were conducted with a Bonferroni correction applied, resulting in a significance level set at $p < 0.017$. Median (interquartile range) perceived clarification usefulness for the rating, ranking, and point distribution instruments were 6 (5 to 7), 6 (5 to 6) and 5 (5 to 6), respectively. There were no statistically significant differences between the point distribution and ranking instruments ($Z = -1.407$, $p = 0.160$) as well as between the ranking and rating instruments ($Z = -2.09$, $p = 0.037$). However, there was a statistically significant difference in the perceived clarification usefulness of the point distribution instrument vs. the rating instrument ($Z = -2.884$, $p = 0.004$). These results reveal that the rating instrument was perceived as being more useful for clarifying men's preferences than the point distribution instrument.

This task would help me to communicate my preferences for prostate cancer screening to my doctor

The Friedman test showed that there was a statistically significant difference in the perceived communication usefulness between the rating, the ranking and the point distribution instruments ($\chi^2(2) = 13.587$, $p = 0.001$). Post-hoc analysis using a Wilcoxon's matched-pairs signed-rank test were conducted with a Bonferroni correction applied, resulting in a significance level set at $p < 0.017$. Median (interquartile range) perceived communication usefulness for rating, ranking, and point distribution instruments were 6 (5 to 7), 6 (5 to 6) and 5 (4 to 6), respectively. There were no statistically significant differences between the communication usefulness of the point distribution and ranking instruments ($Z = -0.960$, $p = 0.337$). On the other hand, there was a statistically significant difference in the perceived communication usefulness of the point distribution instrument vs. the rating instrument ($Z = -2.929$, $p = 0.003$) as well as the rating and ranking instruments ($Z = -2.712$, $p = 0.007$). Therefore, these results reveal that the rating instrument was perceived as being more useful for communicating men's preferences than the point distribution and ranking instruments.

4.3.2 Ease of Use

This task is clear and understandable

The Friedman test showed that there was a statistically significant difference in the perceived clearness for the ease of use item between the rating, the ranking and the point distribution instruments ($\chi^2(2) = 8.010$, $p = 0.018$). Post-hoc analysis using a Wilcoxon's matched-pairs signed-rank test were conducted with a Bonferroni correction applied, resulting in a significance level set at $p < 0.017$. Median (interquartile range) perceived clarification usefulness for rating, ranking, and point distribution instruments were 6 (5 to 7), 6 (5 to 6) and 5 (4 to 7), respectively. There were no statistically significant

differences between the perceived clearness of ranking and rating instruments (Z = -0.391, p = 0.696). There was a statistically significant difference between the perceived clearness of the point distribution and rating instruments (Z=-2.811, p=0.005) and the rating and point distribution instruments (Z=-2.780, p=0.005). Hence, as far as the perceived clearness of the instruments is concerned, both the rating and ranking instruments registered higher scores than the point distribution instrument.

This task is easy to complete

The Friedman test showed that there was a statistically significant difference in the perceived easiness for the ease of use item between the rating, the ranking and the point distribution instruments ($\chi^2(2) = 9.396$, p = 0.009). Post-hoc analysis using a Wilcoxon's matched-pairs signed-rank test were conducted with a Bonferroni correction applied, resulting in a significance level set at $p < 0.017$. Median (interquartile range) perceived easiness for rating, ranking, and point distribution instruments were 6 (5 to 7), 6 (5 to 6) and 6 (4 to 7), respectively. There were no statistically significant differences between the perceived easiness of ranking and rating instruments (Z=-2.173, p=0.030) or between the ranking and point distribution instruments (Z=-1.730, p=0.084). On the other hand, there was a statistically significant difference between the perceived easiness of the point distribution and rating instruments (Z = -3.380, p = 0.001). This difference means that the rating instrument was perceived by the respondents as being easier to complete than the point distribution instrument.

4.4 Time to Completion

This section contains the results of the analysis performed to understand the differences between instruments in the time to completion, in seconds. The descriptive statistics for the time to completion are shown in Table 5.

Table 5 - Descriptive statistics for time to completion (in seconds)

		Percentiles		
	Mean (SD)	25th	50th (Median)	75th
Rating	56.06 (36.4)	35.38	47.29	66.21
Ranking	66.18 (47.0)	36.48	55.20	80.05
Point Distribution	110.63 (103.6)	57.57	82.54	130.74

There was a statistically significant difference in the time to completion depending on which instrument was performed ($\chi^2(2) = 71.210$, $p < 0.001$). Post-hoc analysis using a Wilcoxon matched-pairs signed-ranked tests were conducted with a Bonferroni correction applied, resulting in a significance level set at $p < 0.017$. Respondents were fastest to complete the rating instrument, taking an average of 56.06 seconds to finish it. The ranking instrument was the second quicker task to be completed in average (66.18 seconds) and the point distribution was the instrument that took the longest to be completed in average (110.63 seconds). There were statistically significant differences between the time to completion of the three tasks (Table 6).

Table 6 - Differences between instrument in time to completion

	Ranking vs. Rating	Point Distribution vs. Rating	Ranking vs. Point Distribution
Z	-2.713	-7.285	-6.094
Significance Level	0.007	< 0.001	< 0.001

4.5 Individual Characteristics, Type of Instrument and Perception of Instruments

This section contains the results of the ordinal logistic regressions performed to understand how individual characteristics and the type of instrument were related to the the perception of those instruments. An ordinal logistic regression was performed for each item of the usefulness and ease of use constructs. The results of these regressions are shown in Table 7, where the odds ratios, significance levels and 95% confidence intervals are reported per item. Effects are analysed at the significance levels of $p < 0.05$ and $p < 0.01$.

In order to enable a full comparison between instruments, a second regression model was performed for each item. In these second regression models, the ranking instrument was the reference category, whereas for the first models it had been the point distribution instrument. Such change allows the comparison between the ranking and rating instruments. The remainder of the variables as well as the regression specifications and procedures were maintained. The results of the regressions using the ranking instrument as the reference category are also shown in Table 7.

4.5.1 Usefulness

This task would help me to clarify whether I would like to participate in prostate cancer screening

An ordinal logistic regression was performed to study how individual characteristics and type of instrument were associated with the perceived clarification usefulness. The final model statistically significantly predicted the clarification usefulness variable, as shown by the model fitting information ($\chi^2(8) = 48.107$, $p < 0.001$). It can be seen that men in the [45-49] and [50-54] years groups are less likely to attribute a higher score for the clarification usefulness item compared to men in the [55-75] years group. These results are significant for both age groups at $p < 0.05$ and at $p < 0.01$, respectively. On the other hand, younger respondents i.e., the [40-44] years group, did not have a statistically significant different likelihood of attributing higher scores to instruments compared to the older group, i.e., [55-75] years group. Regarding educational status, men with higher education (i.e., college graduates) are less likely to attribute higher clarification usefulness scores than men who do not hold a university degree. In addition, neither family history of PC nor previous PSA testing were associated with statistically significant differences in the perception of clarification usefulness. Furthermore, in terms of instruments, rating is associated with higher clarification scores compared to point distribution. This result is significant at $p < 0.05$. On the contrary, no statistically

significant association was obtained for the comparison between the ranking and point distribution instruments. The rating instrument was more likely to be perceived as more useful for clarifying preferences when compared to the ranking instrument. This result is significant at $p < 0.05$.

This task would help me to communicate my preferences for prostate cancer screening to my doctor

An ordinal logistic regression was performed to study how individual characteristics and type of instrument were associated with the perceived communication usefulness. The final model statistically significantly predicted the communication usefulness variable, as shown by the model fitting information ($\chi^2(8) = 42.340$, $p < 0.001$). At $p < 0.05$, men in the [50-54] years group are less likely to attribute a higher score for the communication usefulness item compared to men in the [55-75] years group, whereas the remainder of age groups do not have a statistically significant difference compared to the reference group. The direction and statistical significance levels of the remainder of the independent variables are identical to the results reported for the clarification usefulness item.

4.5.2 Ease of Use

This task is clear and understandable

An ordinal logistic regression was performed to study how individual characteristics and type of instrument were associated with the perceived clearness item. The final model statistically significantly predicted the clearness variable, as shown by the model fitting information ($\chi^2(8) = 49.283$, $p < 0.001$). At $p < 0.05$, only men with ages between 50 and 54 years were less likely to attribute higher clearness scores to the instruments when compared to men aged between 55 and 75 years. The results obtained for the other age groups were not statistically significant at conventional significance levels. Regarding educational status, men with higher education were less likely to perceive the instruments as clear and understandable when compared to other respondents who did not have a college degree. Men with a familiar history of prostate cancer were more likely to attribute higher clearness scores to the instruments than men without such history. On the contrary, previous PSA testing was not associated with any statistically significant relation with the clearness ratings in this survey. Finally, both the rating and ranking instruments were associated with an increased likelihood of having higher clearness scores when compared with the point distribution instrument. These effects are significant at $p < 0.05$ and $p < 0.01$, respectively. No statistically significant difference between the likelihood of the rating and ranking instruments being perceived as clear and understandable was obtained.

This task is easy to complete

An ordinal logistic regression was performed to study how individual characteristics and type of instrument were associated with the perceived easiness variable. The final model statistically significantly predicted the easiness variable, as shown by the model fitting information ($\chi^2(8) = 41.549$, $p < 0.001$). Since the odds ratios of the age group do not

provide significant results at any conventional significance levels, no relation between the easiness to complete scores of the older age group with other groups can be concluded. It can be seen that men with higher education are less likely to perceive instruments as easy to complete when compared to men that do not hold a university degree. This effect is statistically significant at $p < 0.05$. In a similar fashion to what was reported for the clearness item, no statistically significant relation with perceived easiness to complete was found within the familiar history of PC and previous PSA testing. As far as the instruments are concerned, the rating instrument is more likely to be perceived as being easier to complete when compared to the point distribution instrument, at $p < 0.01$. On the other hand, the comparison between ranking and point distribution instruments does not provide statistically significant results.

Table 7 - Ordinal logistic regression of usefulness and ease of use items

	Usefulness						Ease of Use					
	This task would help me to clarify whether I would like to participate in prostate cancer screening			*This task would help me to communicate my preferences for prostate cancer screening to my doctor*			*This task is clear and understandable*			*This task is easy to complete*		
	OR	P value	95% CI	OR	P value	95% CI	OR	P value	95% CI	OR	P value	95% CI
Age												
[40-44]	0.493	0.233	[0.154; 1.576]	0.637	0.452	[0.197; 2.064]	1.079	0.884	[0.389; 2.988]	1.156	0.792	[0.394; 3.392]
[45-49]	0.367*	0.042	[0.14; 0.966]	0.418	0.083	[0.156; 0.966]	0.808	0.634	[0.335; 1.947]	1.394	0.470	[0.566; 3.437]
[50-54]	0.269**	0.009	[0.1; 0.721]	0.313*	0.022	[0.116; 0.843]	0.337*	0.019	[0.136; 0.835]	0.525	0.147	[0.220; 1.254]
[55-75]	1			1			1			1		
Educational Status												
College Graduate	0.389*	0.026	[0.17; 0.892]	0.414*	0.039	[0.179; 0.955]	0.363*	0.016	[0.160; 0.825]	0.397*	0.03	[0.160; 0.825]
Lower than College Graduate	1			1			1			1		
Familiar History of PC												
Yes	1.487	0.438	[0.545; 4.059]	1.196	0.725	[0.442; 3.238]	2.517*	0.047	[1.011; 6.264]	2.375	0.055	[0.982; 5.745]
No	1			1			1			1		
Previous PSA Testing												
Yes	0.782	0.484	[0.392; 1.558]	0.893	0.747	[0.448; 1.780]	0.763	0.423	[0.393; 1.480]	1.139	0.707	[0.579; 2.239]
No	1			1			1			1		
Instrument												
Rating	1.448*	0.04	[1.123; 1.866]	1.481**	0.005	[1.125; 1.950]	1.469*	0.012	[1.088; 1.983]	1.665**	0.001	[1.239; 2.237]
Ranking	1.128	0.338	[0.882; 1.443]	1.103	0.406	[0.876; 1.388]	1.469**	0.008	[1.104; 1.953]	1.469	0.104	[0.948; 1.762]
Point Distribution	1			1			1			1		
Instrument[a]												
Rating	1.284*	0.025	[1.031; 1.598]	1.343**	0.007	[1.083; 1.665]	1	1	[0.772; 1.295]	1.288*	0.036	[1.017; 1.630]
Ranking	1			1			1			1		
Point Distribution	0.887	0.338	[0.693; 1.134]	0.907	0.406	[0.720; 1.142]	0.681**	0.008	[0.512; 0.906]	0.773	0.104	[0.567; 1.054]

*significant at p<0.05
** significant at p<0.01
[a]: regression model performed with the ranking instrument as the reference category
OR: odds ratio
CI: confidence interval

4.6 Open-Ended Questions

After assessing a given instrument's usefulness and ease of use, respondents were invited to leave a comment on that instrument. Additionally, a text box was presented at the end of the survey for more general observations. The content and thematic analysis of the comments revealed three main themes. The relevant quotes are organized in Table 8.

Values Elicitation

A fraction of the observations was related to the values elicitation instruments. Three respondents reported being difficult to reflect on the outcomes either in relative terms (Quotes 1 & 2) or in absolute terms (Quote 3). One respondent stated that the point distribution was not useful for clarifying his preferences for PC screening (Quote 4).

Prostate Cancer Screening Decision

Some respondents commented on the decision to undergo PC screening. For some men the decision to do PC screening seemed clear (Quotes 5 & 6), while another respondent advocated for the demystification of the procedure (Quote 7).

Patient-Physician Relationship

Three respondents left comments on the role of the physician and the patient in the PC screening decision-making process. Two men acknowledged the preponderant role of the former actor on the decision-making process, highlighting the existing asymmetries of information between parties (Quote 9 & 10). Other respondent reported never having had reflect on these topics, mentioning that patients are often not adequately informed by physicians on the harms and benefits of PC screening (Quote 11).

Table 8 - Quotes that emerged from the survey and respective (sub-)theme

#	Quote	Sub-theme	Theme
1	*"It is confusing to say if I prefer to die or become incontinent."*	Difficult to reflect on the outcomes	Values elicitation
2	*"Despite being interesting to think about these questions, it is very difficult to think about dying from PC without having the disease."*		
3	*"How can one classify dying in a 1-5 scale."*	Rating instrument	
4	*"This task did not help me at all. I was more confused after doing it."*	Point distribution instrument	
5	*"I would always do screening."*	Prostate cancer screening decision	
6	*"The benefits outweigh the risks. Screening is very important."*		
7	*"PC screening should be demystified."*		

8	*"PC screening decision should be made with objective data and without any emotional considerations."*	
9	*"Physicians were trained to decide this sort of things. I was not."*	Patient-physician relationship
10	*"I do PSA exams annually and trust what my physician tells me."*	
11	*"Normally patients are not adequately informed by physicians about the consequences of the screening. I have never reflected on this matter."*	

5 Discussion and Conclusion

The aim of this research has been to compare Rating, Ranking and Point Distribution instruments for the elicitation and prioritization of individual preferences for PSA screening in a Portuguese population. An initial literature research revealed the scarcity of evidence on the utilization and comparison of these instruments. This holds true especially for European populations. Based on this gap, the aforementioned instruments were compared in terms of their outcome valuation patterns, usefulness, ease of use and time to completion in a Portuguese sample. The extent to which individual characteristics were associated with instruments' perception was also studied. The valuation patterns were firstly analysed per instrument and then by means of comparing how men assessed the most important outcome of PC screening, as defined in the ranking instrument, in the rating and point distribution instruments. The usefulness construct comprised the usefulness of the instrument to clarify preferences for PC screening and communicate those preferences to the physician. The ease of use construct comprised the clearness of the instrument and the ease to complete it. To study the differences between instruments, an experiment was designed in which Portuguese men with ages between 40 and 75 years and with no previous diagnosis of PC were invited to participate in a survey. When agreeing to participate, respondents were asked to use Rating, Ranking, and Point Distribution instruments to reflect on their preferences for PC screening and to rate each instrument according to its usefulness and ease of use. The within-subject design allowed a transversal analysis on how a given individual uses and perceives the three instruments. Methods included a descriptive analysis of the PC screening outcomes valuation, Friedman test, Wilcoxon matched-pairs signed-rank test, and ordinal logistic regressions.

Firstly, the pattern of outcome valuation regarding the frequency each outcome was considered to be important was identical across instruments. The comparative analysis of the patterns of valuation using the ranking instrument as the anchor point revealed that the valuations of the most important outcome as defined in the ranking instrument are more similar to the valuations of the rating instrument than to the point distribution instrument. In fact, around 78% of men perceived their most important outcome as being very important, while barely half of respondents allocated the highest point score to their most important outcome as defined in the ranking instrument.

Secondly, the right-skewed distribution of answers in all four items shows that the study population generally perceived instruments positively both in their usefulness and ease of use. However, differences between instruments on these constructs were found. The rating instrument was perceived as being more useful and easier to use than the point distribution instrument. Significant differences were also found between the rating and ranking instruments in their communication usefulness and between the clearness of the ranking and point distribution instruments.

Thirdly, all instruments were also associated with statistically significant different times to completion. Respondents were quicker to complete the rating instrument and slower to complete the point distribution instrument.

Finally, when isolated from other variables in the regression analyses, there were still differences in the perception of the instruments. In line with the aforementioned findings,

the rating instrument was perceived as being more useful and easier to use than the point distribution and ranking instruments. When comparing the ranking and point distribution instruments, a significant result favouring the former instrument was found only for the clearness item. The rating instrument was significantly associated with higher scores in three out of the four items when compared with the ranking instrument. Additionally, the regression analyses showed that, in this sample, few individual characteristics influence the perception of instruments. There was a general trend of higher educated respondents being less likely to attribute higher scores both to the usefulness and ease of use items. Generally, no significant relation was found between the scores attributed to usefulness and ease of use constructs and familiar history of PC and previous PSA testing.

Considering the abovementioned aspects, the findings of this project add new insights into the realm of individualized outcome measures and into the realm of individualized decision making within the PC screening domain. At the time of writing, no quantitative research on different methods to elicit patient preferences for PC screening in Portugal is published. Therefore, the finding that the rating instrument is more likely to be perceived by this sample of Portuguese men as more useful and easier to use than the ranking and point distribution instruments is an interesting insight that helps building an evidence base in this topic. Although, to the best of my knowledge, no similar experiment was conducted, the existence of such differences was expected based on the work by van Deen et al. (7), particularly the lower usefulness and ease of use scores of the point distribution instrument when compared to the rating instrument. Distributing points can be perceived as being a cognitively difficult task as it involves simultaneously prioritizing outcomes and ensuring that the total point limit is not exceeded. This increased cognitive load was previously described (7) and may help explaining this result. Such observation is also corroborated by one of the comments on the survey which states that the point distribution instrument was not useful as it left that respondent more confused rather than clarified. In terms of perceived usefulness and ease of use, the rating instrument was more likely to be perceived as more useful and easier to use than the ranking instrument, when adjusted for the demographic variables. While previous qualitative work has reported both instruments as being easy to complete (7), the present work cannot determine which are the reasons behind this difference. Additional studies should be performed in order to further explore this aspect. Together with the time to completion differences favouring the rating instrument, these results suggest that using the rating instrument is more advisable for eliciting individual preferences for PC screening as it apparently better performs on the clarification and communication purposes of IOMs while being clear and easy to complete.

The fact that the pattern of outcome valuation pertaining to the frequency each outcome was considered to be important was identical across instruments is aligned with the findings reported by Pignone et al. (63). The instruments studied in the present project may have led men to rely on simple heuristics, thus not involving more deliberative processes as far as the PC screening outcomes valuation is concerned. That may justify the fact that, as reported by Pignone et al. (63), men tended to focus on the most "accessible" outcomes first, such as death from cancer and having cancer itself, across the three instruments. By providing respondents with different mechanisms to reflect on the importance of the same set of outcomes, this research project also shows that the mechanisms underlying each instrument influence the output of that instrument in terms

of the valuation of the most important outcome as defined in the ranking instrument. The rating instrument explicitly asks respondents to reflect on the importance of each outcome individually, i.e., no outcome prioritization is needed. Respondents may consider all outcomes as very important or not, but the importance assessment is independent. While no man rated every outcome with the same number, a considerable proportion of respondents rated the outcomes as either being important or very important. This observation is in line with the work by van Deen et al. (7). On the other hand, the ranking and point distribution instruments involve a relative importance assessment of the outcomes, making the comparisons between these two instruments potentially more insightful. The fact that respondents are obliged to explicitly prioritize outcomes in the ranking instrument, while the point distribution instrument allows them to attribute the same importance to two or more outcomes may justify the aforementioned difference between instruments. The ranking instrument can thus be considered more restrictive as far as the outcome prioritization process is concerned.

Although this is a quantitative research project by nature, its qualitative component also provokes interesting reflections on the role of individualized outcome measures within PC screening decision-making. It seems that the traditional medical model, where the patient is a physician-abiding actor compliant with the decisions of the medical professional (5), still prevails. This long-established patient-physician relationship, based on the large information asymmetries between parties, was clear in many comments to the survey. The fact that this more individualized approach to care is still not prominent among the daily clinical practice may also justify the fact that some respondents either did not understand the purpose of the instruments or found it difficult to reflect on and prioritize the outcomes. This may reveal that men in this sample are still not sufficiently participating on decisions pertaining to their health care. Additionally, some comments also reveal trends on PC screening that were previously described in the literature, such as the lack of awareness of the harms associated with PC screening and, as such, an overestimation of its benefits (53,69,77). Physicians may also accentuate this wrong perception of screening in general (69,78).

At this point, a number of choices and limitations of this research should be highlighted as they necessarily warrant a cautious interpretation of the reported findings.

First of all, due to the lockdown in force in Portugal at the time of recruitment, respondents were recruited through convenience and snowball sampling techniques and, as such, are not a representative sample of the total Portuguese population. Hence, the present results are not generalizable. The population of this study is white, similarly educated and relatively the same age. Considering that a significant proportion of the Portuguese population is African or Afro-descendant together with the higher prevalence of PC in Africans and Afro-descendants (28) as well as the impact of cultural factors on screening decisions and perception of instruments, the predominance of white men in this sample is a major limiting factor of the generalizability of data. Previous work has reported that black men are less likely to participate in PC screening programmes due to a greater fear of having a positive result and the chance of becoming emasculated with the screening procedure and an eventual cancer treatment (79,80). That said, it would be expectable that Afro-Portuguese men reflected on the PC screening outcomes differently. Given the reported information gap on PC screening between black and white men (79,81), differences on the perception of the usefulness of instruments would also be

expectable. Future research including a representative sample of the Portuguese population that allows studying and comparing the instruments' perception of Afro-Portuguese men is thus needed. Additionally, the age inclusion criterion was extended in order to increase the recruitment of respondents. Given that surveys were distributed through online platforms, most respondents are from younger age intervals, thus not completely overlapping with the recommended age range for undergoing PC screening in Portugal.

A within-subject design was chosen for this study, so all respondents were presented and used the three instruments. The major reason behind this choice rests on the fact that recruiting a sufficient number of respondents to each instrument as in a between-subjects design did not seem feasible given the circumstances and time available for recruitment. Given its characteristics, a within-subject design is associated with order effects that carry a bias potential. In order to diminish these effects, the order through which the instruments appeared in the survey was randomized. Additionally, the order through which the outcomes appeared in each instrument was also randomized. Hence, instead of being randomly assigned to a given instrument as it would happen in a between-subjects design, respondents were randomly assigned to different orders of instruments (and outcomes).

The extent to which respondents were adequately informed on PC, screening and the associated trade-offs while completing the survey is also debatable. The information included in the survey was extremely condensed in order to reduce the survey length and, consequently, the dropout rate. Moreover, this project only dealt with PC screening through the PSA test, albeit other screening methods exist, namely the digital rectum examination (DRE). Given the knowledge deficits concerning PC screening reported in the literature (69,77,82–84) and the relative unawareness of the general population on individualized outcome measures and shared-decision making, it is likely that a significant share of respondents could have benefited from a more intensive information flow on these topics, albeit what the impact of more informed respondents on the perception of the different instruments would be is unknown.

Finally, respondents used instruments and reflected on their preferences for PC screening based on a hypothetical scenario. That said, the extent to which different results would have been observed in a real-case scenario is unknown. As in Pignone et al. (63), such effect was attenuated by enrolling men eligible for PC screening and by asking respondents to reflect on their preferences as if they were to make a decision.

Given the reported limitations, the results of the present comparison of a ranking, a rating, and a point distribution instruments for the elicitation of individual preferences for PC screening in a Portuguese population have to be interpreted with caution. As such, further research needs to create a holistic understanding of the role and effects of these (and other) IOMs in conditions that mimic daily clinical practice. This means designing a randomized experiment in a clinical practice setting with a representative sample of the Portuguese population where different instruments and preferences are compared over time. Fortunately, an approved and validated PC screening decision aid tool already exists and can be used in future research in Portuguese men population (61). This setting would also allow men to be adequately informed by providers on the harms and benefits of PC as well as on the aims of a more individualized approach to care. Furthermore, having a

sample that reflects the cultural diversity of the Portuguese population would enable the analysis of the preferred instrument per ethnic group. Such analysis would be of particular interest given the impact of culture and ethnicity on the individual preferences for PC screening. Moreover, it would be also interesting to study what the impact of these instruments in the final decision is by assessing whether the intention to be screened for PC changes after the completion of an instrument.

Final Remarks

The research performed confirms that men use and perceive instruments differently and suggests that the rating instrument seems to be preferred over the ranking and point distribution instruments for the elicitation and prioritization of preferences for PC screening.

Although still undervalued, IOMs may play an important role for physicians, patients, and policymakers. By acknowledging patients as an integrant part of the decision-making process, one is empowering the ultimate beneficiary of care: the patient. Such individualized care framework will ultimately overlap with patients' goals, leading to increasing satisfaction levels. However, it should be stressed that the pursuit of shared decision making, and patient empowerment should not be a blind one. In these situations, the information asymmetries that exist between the patient and his/her provider should be used in favour of the former party, with the physician tailoring care not only in accordance with the individual's goals but also in accordance with good medical practices.

PC screening is an example par excellence of a grey area, where risk-benefit relations are not unanimous and clinical guidelines are not consensual, which in turn renders the decision process ambiguous. That said, it might very well be the case that resources may have been being deployed into procedures/treatments which would not have been needed in the first place, had patients been involved. This is the case in Portugal, where the overuse of preventive healthcare services has been identified (56,57). The resource optimization, particularly relevant for these grey areas, provides a strong incentive for policy makers to invest in more personalized approaches to care, such as the implementation of IOMs in daily clinical practice. While representing a departure from a model of health care delivery that has been long established, IOMs present a unique opportunity to aid in the transformation of health care systems into even more dynamic platforms where healthcare value is maximized.

References

1. Porter ME, Lee TH. The strategy that will fix health care: Providers must lead the way in making value the overarching goal. Harv Bus Rev. 2013;91(10):50–69.
2. Baker A. Crossing the Quality Chasm [Internet]. Vol. 323, Bmj. Washington, D.C.: National Academies Press; 2001. 1192–1192 p. Available from: http://www.nap.edu/catalog/10027
3. Cormier JN, Cromwell KD, Pollock RE. Value-Based Health Care. A Surgical Oncologist's Perspective. Surg Oncol Clin N Am [Internet]. 2012;21(3):497–506. Available from: http://dx.doi.org/10.1016/j.soc.2012.03.001
4. International Aliance of Patients' Organizations. Declaration on: Patient-Centred Healthcare. World Health [Internet]. 2006;(March):2006. Available from: https://www.iapo.org.uk/sites/default/files/files/IAPO_declaration_ENG_2016.pdf%5Cnhttps://www.iapo.org.uk/patient-centred-healthcare
5. Lines LM, Lepore M, Wiener JM. Patient-centered, Person-centered, and Person-directed Care. Med Care [Internet]. 2015 Jul;53(7):561–3. Available from: https://journals.lww.com/00005650-201507000-00001
6. Reuben DB, Jennings LA. Putting Goal-Oriented Patient Care Into Practice. J Am Geriatr Soc [Internet]. 2019 Jul 18;67(7):1342–4. Available from: https://onlinelibrary.wiley.com/doi/abs/10.1111/jgs.15885
7. van Deen WK, Kiaei B, Weaver SA, Crate DJ, Kwon MH, Shah SA, et al. A qualitative inquiry into patients' perspectives on individualized priorities for treatment outcomes in inflammatory bowel diseases. Qual Life Res [Internet]. 2020;29(9):2403–14. Available from: https://doi.org/10.1007/s11136-020-02492-8
8. Berntsen G, Høyem A, Lettrem I, Ruland C, Rumpsfeld M, Gammon D. A person-centered integrated care quality framework, based on a qualitative study of patients' evaluation of care in light of chronic care ideals. BMC Health Serv Res. 2018;18(1):1–15.
9. Berntsen GKR, Gammon D, Steinsbekk A, Salamonsen A, Foss N, Ruland C, et al. How do we deal with multiple goals for care within an individual patient trajectory? A document content analysis of health service research papers on goals for care. BMJ Open. 2015;5(12):1–8.
10. Boeckxstaens P, Willems S, Lanssens M, Decuypere C, Brusselle G, Kühlein T, et al. A Qualitative Interpretation of Challenges Associated with Helping Patients with Multiple Chronic Diseases Identify Their Goals. J Comorbidity [Internet]. 2016 Jan 14;6(2):120–6. Available from: http://journals.sagepub.com/doi/10.15256/joc.2016.6.64
11. Donnelly C, Carswell A. Individualized Outcome Measures: A Review of the Literature. Can J Occup Ther [Internet]. 2002 Apr 22;69(2):84–94. Available from: http://journals.sagepub.com/doi/10.1177/000841740206900204
12. Rocque R, Chipenda Dansokho S, Grad R, Witteman HO. What Matters to Patients and Families: A Content and Process Framework for Clarifying Preferences, Concerns, and Values. Med Decis Mak [Internet]. 2020 Aug 22;40(6):722–34. Available from: http://journals.sagepub.com/doi/10.1177/0272989X20940660
13. Van Poppel H, Hogenhout R, Albers P, van den Bergh RCN, Barentsz JO, Roobol MJ. Early Detection of Prostate Cancer in 2020 and Beyond: Facts and Recommendations for the European Union and the European Commission. European Urology. 2020.

14. Holman CDJ, Wisniewski ZS, Semmens JB, Rouse IL, Bass AJ. Mortality and prostate cancer risk in 19,598 men after surgery for benign prostatic hyperplasia. BJU Int. 1999;84(1):37–42.
15. Berry SJ, Coffey DS, Walsh PC, Ewing LL. The development of human benign prostatic hyperplasia with age. J Urol [Internet]. 1984;132(3):474–9. Available from: http://dx.doi.org/10.1016/S0022-5347(17)49698-4
16. Ilic D, Neuberger MM, Djulbegovic M, Dahm P. Screening for prostate cancer. Cochrane Database Syst Rev [Internet]. 2013 Jan 31;341(7773):563. Available from: http://doi.wiley.com/10.1002/14651858.CD004720.pub3
17. Baptista S, Heleno B, Teixeira A, Taylor KL, Martins C. Comparison of explicit values clarification method (VCM), implicit VCM and no VCM decision aids for men considering prostate cancer screening: Protocol of a randomized trial. BMC Med Inform Decis Mak. 2020;20(1):1–7.
18. Kotwal AA, Walter LC. Cancer Screening in Older Adults: Individualized Decision-Making and Communication Strategies. Med Clin North Am [Internet]. 2020;104(6):989–1006. Available from: https://doi.org/10.1016/j.mcna.2020.08.002
19. Holt JD, Gerayli F. Prostate Cancer Screening. Prim Care - Clin Off Pract [Internet]. 2019;46(2):257–63. Available from: https://doi.org/10.1016/j.pop.2019.02.007
20. Grossman DC, Curry SJ, Owens DK, Bibbins-Domingo K, Caughey AB, Davidson KW, et al. Screening for prostate cancer USPreventive servicestaskforcerecommendation statement. JAMA - J Am Med Assoc. 2018;319(18):1901–13.
21. Carter HB, Albertsen PC, Barry MJ, Etzioni R, Freedland SJ, Greene KL, et al. Early detection of prostate cancer: AUA guideline. J Urol [Internet]. 2013;190(2):419–26. Available from: http://dx.doi.org/10.1016/j.juro.2013.04.119
22. Remmers S, Roobol MJ. Personalized strategies in population screening for prostate cancer. Int J Cancer [Internet]. 2020 Dec 3;147(11):2977–87. Available from: https://onlinelibrary.wiley.com/doi/10.1002/ijc.33045
23. Brenner A, Howard K, Lewis C, Sheridan S, Crutchfield T, Hawley S, et al. Comparing 3 values clarification methods for colorectal cancer screening decision-making: A randomized trial in the US and Australia. J Gen Intern Med. 2014;29(3):507–13.
24. Howard K, Brenner AT, Lewis C, Sheridan S, Crutchfield T, Hawley S, et al. A comparison of US and Australian men's values and preferences for PSA screening. BMC Health Serv Res. 2013;13(1).
25. Pignone MP, Howard K, Brenner AT, Crutchfield TM, Hawley ST, Lewis CL, et al. Comparing 3 techniques for eliciting patient values for decision making about prostate-specific antigen screening: a randomized controlled trial. JAMA Intern Med [Internet]. 2013 Mar 11;173(5):362–8. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23400279
26. Vernooij RWM, Lytvyn L, Pardo-Hernandez H, Albarqouni L, Canelo-Aybar C, Campbell K, et al. Values and preferences of men for undergoing prostate-specific antigen screening for prostate cancer: a systematic review. BMJ Open [Internet]. 2018 Sep 5;8(9):e025470. Available from: https://bmjopen.bmj.com/lookup/doi/10.1136/bmjopen-2018-025470
27. Flynn PM, Betancourt H, Ormseth SR. Culture, emotion, and cancer screening: An integrative framework for investigating health behavior. Ann Behav Med. 2011;42(1):79–90.

28. Tikkinen KAO, Dahm P, Lytvyn L, Heen AF, Vernooij RWM, Siemieniuk RAC, et al. Prostate cancer screening with prostate-specific antigen (PSA) test: A clinical practice guideline. BMJ. 2018;362.
29. Vermunt NP, Harmsen M, Elwyn G, Westert GP, Burgers JS, Olde Rikkert MG, et al. A three-goal model for patients with multimorbidity: A qualitative approach. Heal Expect. 2018;21(2):528–38.
30. Stacey D, Légaré F, Lewis K, Barry MJ, Bennett CL, Eden KB, et al. Decision aids for people facing health treatment or screening decisions. Cochrane Database Syst Rev [Internet]. 2017 Apr 12; Available from: http://doi.wiley.com/10.1002/14651858.CD001431.pub5
31. Montgomery H. Decision Rules and the Search for a Dominance Structure: Towards a Process Model of Decision Making. In: European Journal of Operational Research [Internet]. 1983. p. 343–69. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0166411508622438
32. Fagerlin A, Pignone M, Abhyankar P, Col N, Feldman-Stewart D, Gavaruzzi T, et al. Clarifying values: An updated review. BMC Med Inform Decis Mak [Internet]. 2013;13(SUPPL. 2):S8. Available from: http://www.biomedcentral.com/1472-6947/13/S2/S8
33. Craig BM, Mitchell SA. Examining the Value of Menopausal Symptom Relief Among US Women. Value Heal [Internet]. 2016 Mar;19(2):158–66. Available from: https://linkinghub.elsevier.com/retrieve/pii/S1098301515050883
34. Pignone MP, Fagerlin a, Abhyankar P, Col N, Feldman-Stewart D, Gavaruzzi T, et al. Chapter D: Clarifying and Expressing Values. 2012 Updat Int Patient Decis Aids Stand Collab Backgr Doc Chapter D [Internet]. 2012;1–35. Available from: www.ipdas.ohri.ca
35. Llewellyn-Thomas HA, Crump RT. Decision Support for Patients. Med Care Res Rev [Internet]. 2013 Feb;70(1_suppl):50S-79S. Available from: http://journals.sagepub.com/doi/10.1177/1077558712461182
36. Witterman H, Ndjaboue R, Vaisson G, Dansokho S, Arnold B, Bridges J, et al. Clarifying Values : An Updated and Expanded Systematic Review and. 2021;
37. Pignone MP, Brenner AT, Hawley S, Sheridan SL, Lewis CL, Jonas DE, et al. Conjoint analysis versus rating and ranking for values elicitation and clarification in colorectal cancer screening. J Gen Intern Med. 2012;27(1):45–50.
38. Windon MJ, D'Souza G, Faraji F, Troy T, Koch WM, Gourin CG, et al. Priorities, concerns, and regret among patients with head and neck cancer. Cancer. 2019;125(8):1281–9.
39. Shapiro LM, Eppler SL, Kamal RN. The Feasibility and Usability of a Ranking Tool to Elicit Patient Preferences for the Treatment of Trigger Finger. J Hand Surg Am [Internet]. 2019 Jun;44(6):480-486.e1. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0363502318305343
40. Bradley C, Todd C, Gorton T, Symonds E, Martin A, Plowright R. The development of an individualized questionnaire measure of perceived impact of diabetes on quality of life: the ADDQoL. Qual Life Res [Internet]. 1999;8(1/2):79–91. Available from: http://link.springer.com/10.1023/A:1026485130100
41. Witteman HO, Julien A-S, Ndjaboue R, Exe NL, Kahn VC, (Angie) Fagerlin A, et al. What Helps People Make Values-Congruent Medical Decisions? Eleven Strategies Tested across 6 Studies. Med Decis Mak [Internet]. 2020 Apr 19;40(3):266–78. Available from: http://journals.sagepub.com/doi/10.1177/0272989X20904955
42. Whitty JA, Fraenkel L, Saigal CS, Groothuis-Oudshoorn CGM, Regier DA, Marshall

DA. Assessment of Individual Patient Preferences to Inform Clinical Practice. Patient. 2017;10(4):519–21.
43. Davis FD. Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. MIS Q [Internet]. 1989 Sep;13(3):319. Available from: https://www.jstor.org/stable/249008?origin=crossref
44. Davis FD, Bagozzi RP, Warshaw PR. User Acceptance of Computer Technology: A Comparison of Two Theoretical Models. Manage Sci [Internet]. 1989 Aug;35(8):982–1003. Available from: http://www.jstor.org/stable/2632151.
45. Fishbein, M., & Ajzen I. Belief, Attitude, Intention, and Behavior: An Introduction to Theory and Research. Addison-Wesley, editor. Reading, MA; 1975.
46. Venkatesh V, Davis FD. A Theoretical Extension of the Technology Acceptance Model: Four Longitudinal Field Studies. Manage Sci [Internet]. 2000 Feb;46(2):186–204. Available from: http://pubsonline.informs.org/doi/abs/10.1287/mnsc.46.2.186.11926
47. Venkatesh V, Morris MG, Davis GB DF. User acceptance of information technology: toward a unified view. MIS Q. 2003;27:425–78.
48. Taylor S, Todd PA. Understanding Information Technology Usage: A Test of Competing Models. Inf Syst Res [Internet]. 1995 Jun;6(2):144–76. Available from: http://pubsonline.informs.org/doi/abs/10.1287/isre.6.2.144
49. Holden RJ, Karsh B. The Technology Acceptance Model : Its past and its future in health care. J Biomed Inform [Internet]. 2010;43(1):159–72. Available from: http://dx.doi.org/10.1016/j.jbi.2009.07.002
50. Diop EB, Zhao S, Duy T Van. An extension of the technology acceptance model for understanding travelers' adoption of variable message signs. Dragan D, editor. PLoS One [Internet]. 2019 Apr 25;14(4):e0216007. Available from: https://dx.plos.org/10.1371/journal.pone.0216007
51. Betancourt, H., Flynn P. The psychology of health: physical health and the role of culture in behavior. In: F V, G C, J CG, et al., editors. Handbook of US Latino Psychology [Internet]. Thousand Oaks, CA: Sage Publications; 2009. p. 347–61. Available from: https://www.researchgate.net/publication/264623379_The_psychology_of_health_Physical_health_and_the_role_of_culture_and_behavior
52. Hugosson J, Roobol MJ, Månsson M, Tammela TLJ, Zappa M, Nelen V, et al. A 16-yr Follow-up of the European Randomized study of Screening for Prostate Cancer(Figure presented.). Eur Urol. 2019;76(1):43–51.
53. Hoffmann TC, Del Mar C. Patients' Expectations of the Benefits and Harms of Treatments, Screening, and Tests. JAMA Intern Med [Internet]. 2015 Feb 1;175(2):274. Available from: http://archinte.jamanetwork.com/article.aspx?doi=10.1001/jamainternmed.2014.6016
54. Hofmarcher T, Brådvik G, Svedman C, Lindgren P, Jönsson B, Wilking N. COMPARATOR REPORT ON CANCER IN EUROPE 2019 –DISEASE BURDEN, COSTS AND ACCESS TO MEDICINES [Internet]. 2019. Available from: https://www.efpia.eu/media/580501/comparator-report-on-cancer.pdf
55. Direção Geral da Saúde. Norma nº 060/2011 Prescrição e Determinação do PSA. Serviço Nac Saíude. 2017;1–18.
56. Martins C, Azevedo LF, Ribeiro O, Sá L, Santos P, Couto L, et al. A population-based nationwide cross-sectional study on preventive health services utilization in portugal - What services (and frequencies) are deemed necessary by patients?

PLoS One. 2013;8(11):1–9.
57. Martins C, Azevedo LF, Santos C, Sá L, Santos P, Couto M, et al. Preventive health services implemented by family physicians in Portugal - A cross-sectional study based on two clinical scenarios. BMJ Open. 2014;4(5).
58. Braga R, Costa AR, Pina F, Moura-Ferreira P, Lunet N. Prostate cancer screening in Portugal: Prevalence and perception of potential benefits and adverse effects. Eur J Cancer Prev. 2020;248–51.
59. Costa AR, Silva S, Moura-Ferreira P, Villaverde-Cabral M, Santos O, Carmo I do, et al. Cancer screening in Portugal: sex differences in prevalence, awareness of organized programmes and perception of benefits and adverse effects. Heal Expect. 2017;20(2):211–20.
60. Slater MD, Long M, Bettinghaus EP, Reineke JB. News Coverage of Cancer in the United States: A National Sample of Newspapers, Television, and Magazines. J Health Commun [Internet]. 2008 Aug 28;13(6):523–37. Available from: http://www.tandfonline.com/doi/abs/10.1080/10810730802279571
61. Baptista S, Heleno B, Pinto M, Guimarães B, China D, Ramos JP, et al. Translation and cultural adaptation of a prostate cancer screening decision aid: A qualitative study in Portugal. BMJ Open. 2020;10(3):1–9.
62. Howard K, Salkeld GP, Mann GJ, Patel MI, Cunich M, Pignone MP. The COMPASs study: Community Preferences for Prostate Cancer Screening. Protocol for a quantitative preference study. BMJ Open. 2012;2(1):1–6.
63. Pignone MP, Howard K, Brenner AT, Crutchfield TM, Hawley ST, Lewis CL, et al. Comparing 3 techniques for eliciting patient values for decision making about prostate-specific antigen screening: A randomized controlled trial. JAMA Intern Med. 2013;173(5):362–8.
64. De Bekker-Grob EW, Rose JM, Donkers B, Essink-Bot ML, Bangma CH, Steyerberg EW. Men's preferences for prostate cancer screening: A discrete choice experiment. Br J Cancer. 2013;108(3):533–41.
65. Baptista S, Heleno B, Teixeira A, Taylor KL, Martins C. Comparison of Explicit, Implicit and no Values Clarification Decision Aids for Men Considering Prostate Cancer Screening [Internet]. [cited 2021 May 22]. Available from: https://clinicaltrials.gov/ct2/show/NCT03988673
66. Qualtrics, Homepage [Internet]. 2021. Available from: https://www.qualtrics.com/
67. Tadeu ACR, e Silva Caetano IRC, de Figueiredo IJ, Santiago LM. Multimorbidity and consultation time: a systematic review. BMC Fam Pract [Internet]. 2020 Dec 28;21(1):152. Available from: https://bmcfampract.biomedcentral.com/articles/10.1186/s12875-020-01219-5
68. 2021 G de TC, (GT) – Questões "Étnico-Raciais." SUMÁRIO DO TRABALHO DO GRUPO DE TRABALHO (GT) CENSOS 2021 – QUESTÕES "ÉTNICO-RACIAIS" [Internet]. 2019. Available from: https://www.acm.gov.pt/documents/10181/167771/Sumário+Trabalho+GT+Censos+2021+Questões+Étnico-raciais.pdf/6ba40214-9a39-4a88-96b4-5c2919da14d9
69. Baptista AS. Shared Decision-Making in Prostate Cancer Screening. Faculty of Medicine, University of Porto; 2020.
70. Sheldon MR, Fillyaw MJ, Thompson WD. The use and interpretation of the Friedman test in the analysis of ordinal-scale data in repeated measures designs. Physiother Res Int [Internet]. 1996 Nov;1(4):221–8. Available from:

http://doi.wiley.com/10.1002/pri.66
71. Sheskin DJ. Handbook of Parametric and Nonparametric Statistical Procedures [Internet]. 5th ed. Boca Raton: Chapman and Hall/CRC; 2020. Available from: https://www.taylorfrancis.com/books/9781439858042
72. Liang K-Y, Zeger SL. Longitudinal data analysis using generalized linear models. Biometrika [Internet]. 1986;73(1):13–22. Available from: https://academic.oup.com/biomet/article-lookup/doi/10.1093/biomet/73.1.13
73. Ballinger GA. Using Generalized Estimating Equations for Longitudinal Data Analysis. Organ Res Methods [Internet]. 2004 Apr 29;7(2):127–50. Available from: http://journals.sagepub.com/doi/10.1177/1094428104263672
74. Group USC. HOW ARE THE LIKELIHOOD RATIO, WALD, AND LAGRANGE MULTIPLIER (SCORE) TESTS DIFFERENT AND/OR SIMILAR? [Internet]. [cited 2021 Jun 4]. Available from: https://stats.idre.ucla.edu/other/mult-pkg/faq/general/faq-how-do-i-cite-web-pages-and-programs-from-the-ucla-statistical-consulting-group/
75. Fox J. Applied regression analysis, linear models, and related methods. Thousand Oaks, CA: Sage Publications; 1997.
76. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol [Internet]. 2006 Jan;3(2):77–101. Available from: http://www.tandfonline.com/doi/abs/10.1191/1478088706qp063oa
77. Gigerenzer G, Mata J, Frank R. Public Knowledge of Benefits of Breast and Prostate Cancer Screening in Europe. JNCI J Natl Cancer Inst [Internet]. 2009 Sep 2;101(17):1216–20. Available from: https://academic.oup.com/jnci/article-lookup/doi/10.1093/jnci/djp237
78. Wegwarth O, Gigerenzer G. "There is nothing to worry about": Gynecologists' counseling on mammography. Patient Educ Couns [Internet]. 2011 Aug;84(2):251–6. Available from: https://linkinghub.elsevier.com/retrieve/pii/S073839911000426X
79. Alexis O, Worsley A. An integrative review exploring black men of African and Caribbean backgrounds, their fears of prostate cancer and their attitudes towards screening. Health Educ Res [Internet]. 2018 Apr 1;33(2):155–66. Available from: https://academic.oup.com/her/article/33/2/155/4850509
80. Chinegwundoh F, Enver M, Lee A, Nargund V, Oliver T, Ben-Shlomo Y. Risk and presenting features of prostate cancer amongst African-Caribbean, South Asian and European men in North-east London. BJU Int [Internet]. 2006 Dec;98(6):1216–20. Available from: http://doi.wiley.com/10.1111/j.1464-410X.2006.06503.x
81. Chan ECY, Vernon SW, O'Donnell FT, Ahn C, Greisinger A, Aga DW. Informed Consent for Cancer Screening With Prostate-Specific Antigen: How Well Are Men Getting the Message? Am J Public Health [Internet]. 2003 May;93(5):779–85. Available from: http://ajph.aphapublications.org/doi/10.2105/AJPH.93.5.779
82. Arroll B, Pandit S, Buetow S. Prostate cancer screening: knowledge, experiences and attitudes of men aged 40-79 years. N Z Med J [Internet]. 2003 Jun 20;116(1176):U477. Available from: http://www.ncbi.nlm.nih.gov/pubmed/12835805
83. Arnold-Reed DE, Hince DA, Bulsara MK, Ngo H, Eaton M, Wright AR, et al. Knowledge and attitudes of men about prostate cancer. Med J Aust [Internet]. 2008 Sep 15;189(6):312–4. Available from: https://onlinelibrary.wiley.com/doi/abs/10.5694/j.1326-5377.2008.tb02047.x

84. Fàbregas Escurriola M, Guix Font L, Aragonès Forès R, Casajuana Brunet J, Ballester Torrens M. ¿Conocen los varones de 50-70 años la efectividad, los beneficios y los riesgos del cribado del cáncer de próstata? Atención Primaria [Internet]. 2008 Jul;40(7):357–61. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0212656708703652
85. Force USPST. Is Prostate Cancer Screening Right For You? [Internet]. 2018. Available from: https://www.uspreventiveservicestaskforce.org/uspstf/recommendation/prostate-cancer-screening

Appendices

Appendix I – Translated version of the survey

1. Introduction

Thank you for participating on this survey about prostate cancer screening.

The present project is carried out as part of a master's dissertation in Health Economics, Policy & Law of Erasmus School of Health Policy and Management (ESHPM) of Erasmus University Rotterdam. The scope of this project serves purely educational purposes, and your help would be greatly appreciated.

This master thesis is led by Francisco Baptista and supervised by Welmoed van Deen and aims to compare three different strategies for obtaining and prioritizing individual preferences for prostate cancer screening through the PSA (prostate specific antigen) test.

Men between 40 and 75 years old and with no previous diagnosis of prostate cancer are eligible to participate. This questionnaire should take you approximately 9 minutes to answer.

Now I would like to confirm that you are man between 40 and 75 years with no previous diagnostic of prostate cancer.

- I confirm that I am a man between 40 and 75 years
- I confirm that I have never been diagnosed with prostate cancer.

2. Informed Consent

You will take part in a project which collects information by means of a questionnaire. The questionnaire should take you approximately 9 minutes to answer. You will be asked to reflect on your preferences for prostate cancer screening in three different tasks. After each task, you will be asked to assess its usefulness for helping you clarify your preferences for prostate cancer screening and ease-of-use. There is also a group of questions on demographic characteristics.

We may ask you questions during this questionnaire which you may feel to be personal due to the sensitive nature of the subject. We only ask these questions in the interest of the project. However, you don't need to answer any questions which you don't want to answer. You are taking part on a voluntary basis and can stop whenever you want to. If you decide to pull out from the survey, this will not have any adverse consequences for you.

We will do everything we can to protect your privacy as much as possible. Data will be collected anonymously, and confidentiality will be safeguarded. Confidential information

or personal data relating to you will not be publicised in any way; no one will be able to trace this information/data back to you.

You will not receive any compensation for taking part in this project.

The data collected will be used for educational purposes in the context of this master thesis. The data will be kept for 1 year and will not be disclosed to any entity.

The responsible parties are identified in the following paragraph and are obliged to maintain professional secrecy both in the collection and data analysis processes.

A. Name: Francisco Baptista, Master student at ESHPM
E-mail address: 573331fp@student.eur.nl
Address: Burgemeester Oudlaan 50, 3062 PA Rotterdam, The Netherlands.

B. Name: Welmoed van Deen, Project supervisor, Assistant Professor at ESHPM.
E-mail address: vandeen@eshpm.eur.nl
Address: Burgemeester Oudlaan 50, 3062 PA Rotterdam, The Netherlands.

If you would like to receive a copy of the results of the study or if you have any questions, please email at 573331fp@student.eur.nl.

I have read and understood the information above and agree to participate in this project.

- Yes
- No

3. Demographic Characteristics

In this next section, we would like to know more who you are.

Age

- 40-44
- 45-49
- 50-54
- 55-59
- 60-64
- 65-69
- 70-75

Race

- White / European origin
- Black / Afro descendent
- Asian
- Other

Employment Status

- Employed
- Unemployed
- Retired

Educational Level

- College graduate
- High school graduate
- Less than high school graduate

Income

Considering an average monthly income of 970 €, your monthly gross income is:

- Below average (less than 970 € / month)
- Above average (more than 970 € / month)
- Prefer not to answer

Family history of prostate cancer

Has any close relative of yours (e.g., father and/or brother) been diagnosed with prostate cancer?

- Yes
- No

Prior PSA testing

When was the last time you did a PSA test for prostate cancer screening?

- Less than 1 year ago
- More than 1 year ago
- Never
- Don't know

4. Information on Prostate Cancer Screening

In this next section, you will be presented some information on prostate cancer screening that will help you complete the questionnaire.

Most prostate cancers end up not being diagnosed, because:

- have a slow progression.

- do not cause any symptoms.
- never progress to a serious health issue.

However, a minority of prostate cancers can:
- evolve quickly.
- spread to other organs.
- cause illness and death.

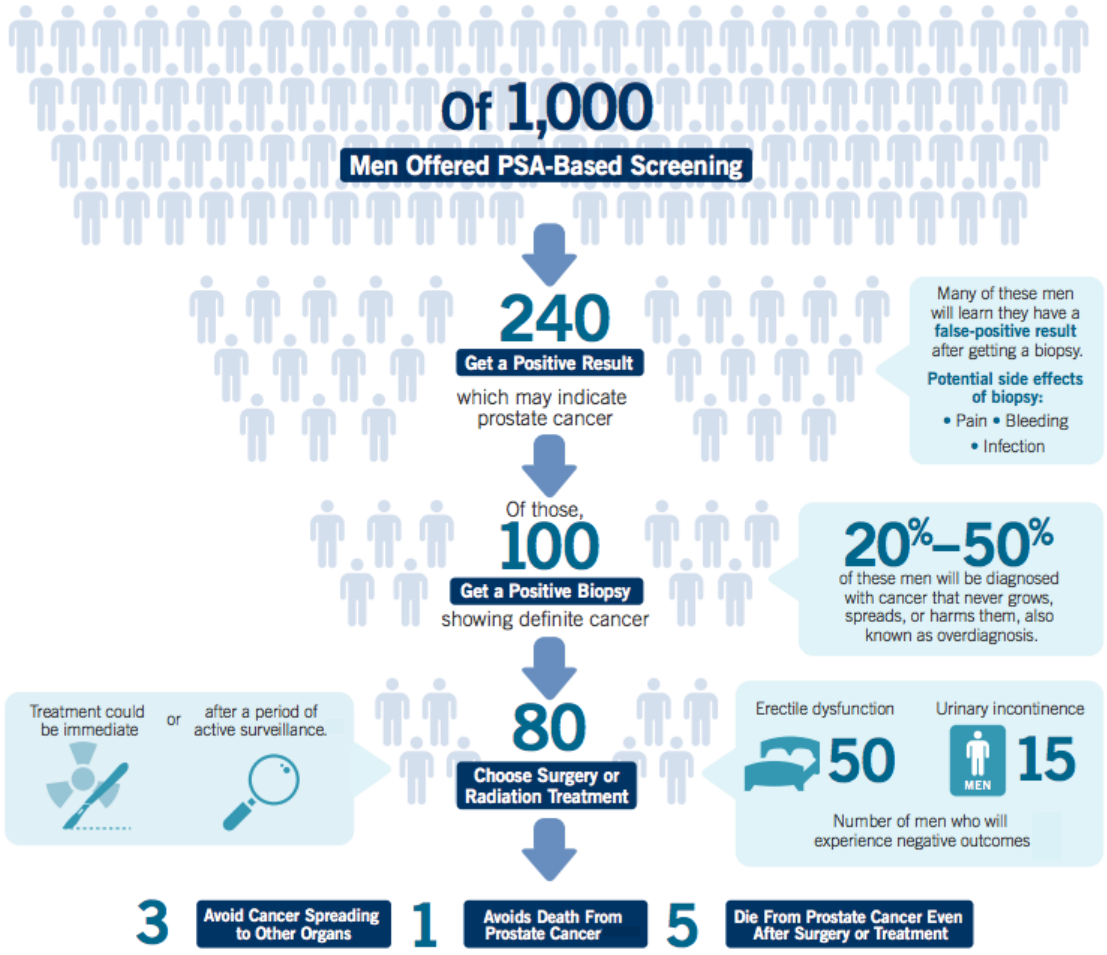
Currently, the extent to which benefits of prostate cancer screening exceed its risks is uncertain.

Benefits of screening
- early detection of more severe types of prostate cancer before any symptoms appear.

Risks of screening
- once cancer is diagnosed, there is no way to predict how the disease will progress.
- some types of cancer may be unnecessarily treated, causing side effects.

Therefore, is up to the patient and his physician to decide whether to perform prostate cancer screening through the PSA test.

Please consider the image below with the potential risks and benefits of screening for prostate cancer.



[English version of the infographic adapted from (85)]

Now imagine that you had to decide if you would like to do a prostate cancer screening.

In the next sections you will assess:
- 4 outcomes of prostate cancer screening.
- the importance of each outcome in the moment of deciding if you want to do a prostate cancer screening.

The outcomes are:

- *Finding prostate cancer if I have it.*
- *Dying from prostate cancer.*
- *Having a prostate biopsy as a result of screening.*
- *Becoming impotent or incontinent as a result of an eventual prostate cancer treatment*

Please proceed to the first task.

5. Ranking Task

For this ranking task, please look at the 4 outcomes of prostate cancer screening.
Please rank the outcomes, based on which ones are the most important for you in the moment of deciding if you want to do a prostate cancer screening.
Mark the most important outcome with a '1', the second most important outcome with a '2' and continue until you mark the least important outcome with a '4'.

- ☐ *Finding prostate cancer if I have it.*
- ☐ *Dying from prostate cancer.*
- ☐ *Having a prostate biopsy as a result of screening.*
- ☐ *Becoming incontinent or impotent as a result of screening.*

6. Acceptability assessment of instrument

Was this task useful and easy to complete?

Please classify the following statements on a scale from 1 to 7.

This task would help me to clarify whether I would like to participate in prostate cancer screening.

Totally Disagree Totally Agree

○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ 6 ○ 7

This task would help me to communicate my preferences for prostate cancer screening to my doctor.

Totally Disagree Totally Agree

○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ 6 ○ 7

This task is easy and understandable.

Totally Disagree Totally Agree

○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ 6 ○ 7

This task is easy to complete.

Totally Disagree Totally Agree

○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ 6 ○ 7

Would you like to leave any comment about the ranking task? (Optional)

_____

7. Rating Task

For this rating task please reflect upon the 4 outcomes of prostate cancer screening individually and rate them according to their importance on a scale of 0 to 5, where 0 is NOT AT ALL IMPORTANT and 5 is VERY IMPORTANT. You may only check one box per question.

	Not at All Important ←					→ Very Important
	0	1	2	3	4	5
Finding prostate cancer if I have it						
Dying from prostate cancer						
Having a prostate biopsy as a result of screening						
Becoming impotent or incontinent as a result of screening						

8. Acceptability assessment of instrument

Was this task useful and easy to complete?

Please classify the following statements on a scale from 1 to 7.

This task would help me to clarify whether I would like to participate in prostate cancer screening.

Totally Disagree Totally Agree

○ ○ ○ ○ ○ ○ ○

1 2 3 4 5 6 7

This task would help me to communicate my preferences for prostate cancer screening to my doctor.

Totally Disagree Totally Agree

○ ○ ○ ○ ○ ○ ○

1 2 3 4 5 6 7

This task is easy and understandable.

Totally Disagree Totally Agree

○ ○ ○ ○ ○ ○ ○

1 2 3 4 5 6 7

This task is easy to complete.

Totally Disagree Totally Agree

○ ○ ○ ○ ○ ○ ○

1 2 3 4 5 6 7

Would you like to leave any comment about the ranking task? (Optional)

_____

9. Point Distribution Task

In this point distribution task, you are asked to distribute 40 points by the 4 prostate cancer screening outcomes, according to how important each one is to you in the moment of deciding if you want to do a prostate cancer screening.
The choice is yours to split the points up any way you like, but you cannot have more than 40 points in total. For example, you could give 10 points to each outcome if they are all equally important, or you might give 40 points to one outcome.

	Points
Finding prostate cancer if I have it	
Dying from prostate cancer	
Having a prostate biopsy as a result of screening	
Becoming impotent or incontinent as a result of screening	
TOTAL	

10. Acceptability assessment of instrument

Was this task useful and easy to complete?

Please classify the following statements on a scale from 1 to 7.

This task would help me to clarify whether I would like to participate in prostate cancer screening.

Totally Disagree Totally Agree

○ ○ ○ ○ ○ ○ ○
1 2 3 4 5 6 7

This task would help me to communicate my preferences for prostate cancer screening to my doctor.

Totally Disagree Totally Agree

○ ○ ○ ○ ○ ○ ○
1 2 3 4 5 6 7

This task is easy and understandable.

Totally Disagree Totally Agree

○ ○ ○ ○ ○ ○ ○
1 2 3 4 5 6 7

Totally Disagree Totally Agree

○ ○ ○ ○ ○ ○ ○
1 2 3 4 5 6 7

Would you like to leave any comment about the ranking task? (Optional)

_____

11. Box to give the respondents an opportunity to leave a comment

12. End of Survey

The survey has ended. Thank you for your participation!

Appendix II – Usefulness and Ease of Use Scores Distribution of Rating Task

Construct	Items	N (%)							
		Strongly Disagree	Disagree	Somewhat Disagree	Neither agree nor disagree	Somewhat Agree	Agree	Strongly Agree	Median
Perceived Usefulness	This task would help me to clarify whether I would like to participate in prostate cancer screening.	5 (4.2)	3 (2.5)	2 (1.7)	17 (14.3)	22 (18.5)	40 (33.6)	30 (25.2)	Agree
	This task would help me to communicate my preferences for prostate cancer screening to my doctor.	5 (4.2)	2 (1.7)	3 (2.0)	12 (8)	27 (22.7)	40 (33.6)	30 (25.2)	Agree
Perceived Ease of Use	This task is clear and understandable	2 (1.7)	3 (2.5)	5 (4.2)	13 (10.9)	28 (23.5)	34 (28.6)	34 (22.7)	Agree
	This task is easy to complete	2 (1.7)	1 (0.8)	4 (3.4)	16 (13.4)	18 (15.1)	39 (32.8)	39 (32.8)	Agree

Notes: 1 (Strongly Disagree) to 7 (Strongly Agree)

Appendix III – Usefulness and Ease of Use Scores Distribution of Ranking Task

Construct	Items	N (%)							
		Strongly Disagree	Disagree	Somewhat Disagree	Neither agree nor disagree	Somewhat Agree	Agree	Strongly Agree	Median
Perceived Usefulness	This task would help me to clarify whether I would like to participate in prostate cancer screening.	6 (5)	4 (3.4)	3 (2.5)	16 (13.4)	29 (24.4)	36 (30.3)	25 (21)	Agree
	This task would help me to communicate my preferences for prostate cancer screening to my doctor.	6 (5)	1 (0.8)	6 (5)	16 (13.4)	28 (23.5)	37 (31.1)	25 (21)	Agree
Perceived Ease of Use	This task is clear and understandable	3 (2)	2 (1.7)	4 (3.4)	17 (14.4)	26 (22)	31 (26.3)	35 (29.7)	Agree
	This task is easy to complete	3 (2.5)	2 (1.7)	5 (4.2)	15 (12.7)	26 (22)	33 (28)	34 (28.8)	Agree

Notes: 1 (Strongly Disagree) to 7 (Strongly Agree)

Appendix IV – Usefulness and Ease of Use Scores Distribution of Point Distribution Task

Construct	Items	N (%)							
		Strongly Disagree	Disagree	Somewhat Disagree	Neither agree nor disagree	Somewhat Agree	Agree	Strongly Agree	Median
Perceived Usefulness	This task would help me to clarify whether I would like to participate in prostate cancer screening.	5 (4.2)	4 (3.4)	6 (5)	24 (20.2)	25 (21)	26 (21.8)	29 (24.4)	Somewhat Agree
	This task would help me to communicate my preferences for prostate cancer screening to my doctor.	5 (4.2)	1 (0.8)	9 (7.6)	16 (13.4)	30 (25.2)	34 (28.6)	24 (20.2)	Somewhat Agree
Perceived Ease of Use	This task is clear and understandable	4 (3.4)	2 (1.7)	14 (11.8)	17 (14.3)	24 (20.2)	27 (22.7)	31 (26.1)	Somewhat Agree
	This task is easy to complete	4 (3.4)	2 (1.7)	10 (8.4)	19 (16)	22 (18.5)	31 (26.1)	31 (26.1)	Agree

Notes: 1 (Strongly Disagree) to 7 (Strongly Agree)