University of Twente / St. Elisabeth Hospital

EXPLICIT TREATMENT ADVICE IN UROLOGY: TRICK OR TREAT?

Determining optimal therapy for localized prostate cancer by using a decision aid

Leonie Hulshof

Master Thesis Health Sciences in the context of Health Services and Management

Faculty: Behavioural Management & Social Sciences

Supervisors of the University of Twente: Dr. J.A. van Til & M.G.M. Weernink MSc

Supervisors of the St. Elisabeth Hospital: Dr. P.J.M. Kil & R.E.D. Lamers MSc

August 17, 2015

Master thesis

"Explicit Treatment Advice in Urology: Trick or Treat?"

Determining optimal therapy for localized prostate cancer by using a decision aid

Leonie Hulshof (Msc)
Kuipersdijk 84
7512 CK Enschede
E: l.hulshof@student.utwente.nl

UT Supervisors:
Mrs. Dr. J.A. van Til
University of Twente
P.O. Box 217 7500 AE Enschede
E: j.a.vantil@utwente.nl
T: +31 53 489 3351
F: +31 53 489 2159

Mrs. M.G.M. Weernink (MSc)
University of Twente
P.O. Box 217 7500 AE Enschede
E: m.g.m.weernink@utwente.nl
T: +31 53 489 3351
F: +31 53 489 2159

St. Elisabeth Hospital Supervisors:
Mr. Dr. P.J.M. Kil & Mrs. R.E.D. Lamers (MSc)
Urology specialists
P.O. Box 90151 5000 LC Tilburg
E: info@elisabeth.nl
T: +31 74 255 4142
F: +31 74 255 4192

UNIVERSITY OF TWENTE.

St. Elisabeth Ziekenhuis

Preface

After acquiring my Master's degree in Psychology, I realized that I wanted to specialize in health care, especially the management area. In 2014, I therefore decided to start a second education: Health Sciences. It has since then been my dream to finish this study with a master thesis incorporating both Psychology and Health Sciences. In addition, it has been my desire to contribute to the existing health care literature by studying a subject that has not gotten the research attention it deserves. I managed to fulfill these wishes by collaborating with the University of Twente and the St. Elisabeth Hospital in Tilburg.

One of the main topics of this study is the patient's search for a treatment for localized prostate cancer, which best fits his characteristics and lifestyle. This exploration is symbolically interwoven throughout the master thesis. The pieces of information on which patients base their decisions are symbolized as twigs at the bottom of each page. The different twigs should be collected in order to find the most appropriate treatment option for the patient. In other words; the patient (with assistance of his specialist) has to find his tree in the wood of treatment possibilities. This wood of treatment possibilities is illustrated on the front page. This thesis ends with the most appropriate treatment symbolically shown as a heart-shaped tree. This heart shape coincides with my beliefs. I believe that the best decision can only be made by using your mind in combination with listening to your heart. Both elements are integrated.

The completion of this master thesis would not have been possible without the participation and assistance of some people whose names cannot all be enumerated. Their contributions are sincerely appreciated and gratefully acknowledged. I would like to express my deep appreciation particularly to the following persons:

Ms. J.A. van Til and Ms. M.G.M. Weernink, my supervisors of the University of Twente, for the sharp comments, the theoretical knowledge and involvement in my research project.

Mr. K.A. Kroeze, Assistant Datalab of the University of Twente, who transformed me into a web developer. Something I never thought I could.

Mr. P.J.M. Kil and Ms. R.E.D. Lamers, specialists of the St. Elisabeth Hospital in Tilburg, for providing me with all features and information that I needed to write this thesis.

All relatives, friends and others who supported me through this venture.

Thank you!

Leonie Hulshof

Enschede, August 17, 2015

Table of Contents

Terms and definitions

A

Active surveillance strategies: In this method (still) no curative treatment is performed. The patient will be closely monitored to avoid unnecessary surgery. If necessary, a treatment is still possible. This concept is illustrated as one of the treatment options.

Adverse effect (complication): An unintended negative outcome which develops as a cause of treatment already present during or after care and either necessitates (adjustment of) treatment or leads to permanent harm.

Attributes: Aspects that characterize alternative possibilities considered in a decision. Attributes may establish preferences about structural aspects of health care (e.g. preferences for information in decision-making), process aspects (e.g. usability of the digital decision aid) and/or outcomes of the health actions (e.g. efficacy or adverse effects of a treatment).

B

Benefit: A positive outcome after or during a treatment.

Balanced spread: When no statement in the experimental designs dominated the other ones.

Best-Worst Scaling experiments (BWS): A method in which persons choose the most and least preferred attributes from a list of three or more presented in a single choice task. Three variants; case 1 (single profile with only attributes), case 2 (single profile with attributes and levels) and case 3 (multi-profile with attributes and levels)

C

Comorbidity of diseases: Pertaining to a disease or other pathological process that occurs simultaneously with another disease.

Computer-adaptive tests: Digital tests designed to adjust their level of difficulty, based on the responses provided, to match the knowledge and ability of the persons. For example: If a person gives a wrong answer, the computer follows up with an easier question. If the person answers correctly, the next question will be more difficult.

Conjoint Analysis: Statistical technique to measure preferences for product features to identify their implicit preferences. For example the Rating method, Ranking method, Discrete Choice Experiments (DCE) and Best-Worst Scaling experiments (BWS).

Control Preferences Scale (CPS): A scale that indicates the degree of control an individual wants to assume when decisions are being made about medical treatment. This is translated in three roles; a passive, collaborative or active role.

Choice-based method: A hypothetical method, which ask persons to make choices based on a hypothetical scenario. An example of a choice-based method is the conjoint analysis.

Curative treatment: A treatment and/or therapies provided to a patient with an intent to improve symptoms and cure the patient's medical problem. Examples are radiotherapy or radical prostatectomy.

D

Decision aid: An information device that besides an informative summary of the main benefits and adverse effects also explicitly considers the personal importance of each benefit or adverse effect.

Decisional Conflict Scale (DCS): Evaluation scale that the personal perceptions and modifiable factors measures, which contribute to uncertainty in choosing options and expressing satisfaction with the choice.

Discrete choice experiments (DCE): A method in which the patient has to choose their preferred scenario from two or more scenarios.

E

External radiotherapy (EBRT): A technique in which the entire prostate will be irradiated from the outside. This concept is illustrated as one of the four treatment options for localized prostate cancer.

F

Full factorial designs (full profiles): An experimental design with all possible combinations of attributes and levels.

G

Gleason score: The Gleason score reflects the rapidly growing tumor. This is determined on the basis of prostate biopsy. The Gleason score ranges from 2 to 10.

H

No definitions

I

Inclusion criteria: Relevant characteristics in the light of the research question. The criteria set out before a study or review. Inclusion criteria are used to determine whether a person can participate in a research study.

In-depth topic interview: An open-ended, discovery-oriented method that is well suited for describing both treatment processes and outcomes from the perspective of the target audience or key stakeholder. The goal is to generate with the use of topics as much opinions, feelings and thoughts as possible.

Internal radiotherapy (Brachy therapy): A technique in which first under general anesthesia or an epidural insertion radioactive 'seeds' are posted into the patient's body and thereafter the entire prostate will be irradiated. This concept is illustrated as one of the four treatment options for localized prostate cancer.

J - K

No definitions

L

Levels: The different possibilities that can be selected on an attribute. The levels could be categorical (e.g. gender), ordinal (e.g. treatment frequency) or ratio/interval scaled (e.g. change of a certain outcome or adverse effect).

Limesurvey: A digital static survey application, which would be used to develop online surveys, collect responses and create statistics that could be exported to other applications as Excel.

M - N

No definitions

O

Orthogonal: Low correlations between the attributes.

P

Patient preferences: The preferences of a patient are defined as an individual appreciation over health outcomes in health care, which results from cognition, experience and reflection.

Partial factorial design (or d-efficient design): A design, which is characterized through the use of a subset of attributes and levels instead of all attributes and levels.

Participants: Patients, who want to participate in an observational study or interview.

Pilotilot test: A small scale preliminary study conducted in order to evaluate feasibility, time, cost, adverse effects, and effect size (statistical variability).

Preference sensitive: There is no single alternative among the available options that is best for every individual, but instead the optimal decision depends on patient preferences

Prostate cancer: A cancer that occurs when cells in the prostate grow unrestrained and form a cancerous (malignant) tumor.
- *Localized prostate cancer:* The tumor is solely located in the prostate.

PSA: Prostate Specific Antigen produced by normal prostate cells or prostate cancer cells. When (localized) prostate cancer is diagnosed, the PSA value increases regularly and is often higher than 4.

Q

Quality of life (QOL): A broad multi-dimensional concept, which usually consists of subjective evaluations of a person about both positive and negative aspects of life. It indicates the general well-being of a person (or society), defined in terms of health and happiness, rather than wealth.

Quantitative method: An objective method directed to digits/numeric data. Examples are experimental research, secondary analysis, survey research and monitoring.

Qualitative method: An interpretative and subjective method. A qualitative method is not about facts and figures, but about the '*why*' and '*how*' question. Examples are interviews, literature studies, observations and case studies.

R

Radical prostatectomy: A medical term for a surgery in which the prostate with the cancer cells will be removed. This concept is illustrated as one of the treatment options.

Ranking scales/method (implicit score): A method in which the patient has to rank different scenarios in order of attractiveness (not attractive - very attractive).

Rating scales/method (real score): A method in which the patient has to rate each scenario on its attractiveness (not attractive - very attractive).

Respondents: Patients, who want to fill in a questionnaire.

Revealed preferences: Preferences derived from observations of the behavior of the patients.

Routing logic: The logic behind the digital process of moving a packet of data from a source to his destination.

S

Satisfaction with Decision Scale (SWD): An evaluation scale which measures satisfaction related to the participation in the decision with a perception of having received the correct information.

Sawtooth software: Computer software for conjoint analysis, also called the survey software of choice.

Selection bias: The selection of individuals, groups or data for analysis such that proper randomization is not achieved, thereby ensuring that the sample obtained is not representative of the population intended to be analyzed.

Self-fullfilling property: The prediction that directly or indirectly causes itself to become true, due to positive feedback between belief and behavior. This results in searching for conformation of the prediction.

Shared decision-making (SDM) (also called clinical decision-making): A mechanism to minimize the informational and power asymmetry between specialists and patients by increasing patients' information, sense of autonomy and/or control over treatment decisions that affect their quality of life.

Stated preferences: Preferences derived from questions focusing on the behavioral intentions or choices of patients on existing and /or hypothetical products or services.

T

Think- aloud technique: A technique in which participants were asked to think out loud and say all the things that come to mind.

U - Z

No definitions

Summary

Background: Choosing the best treatment for patients with localized prostate cancer is a challenge for both treating specialists and patients. Specialists are unaware of the patient's value and preferences and patients have insufficient knowledge about the different treatment options and likely outcomes. There is often no consensus reached about what the optimal treatment is and the consultations are usually too brief to provide and discuss all information that patients need to determine their preferences. A digital preference-based value clarification decision tool could assist patients in gathering information and clarifying their preferences with a personal advice. It also enables them to participate in shared decision-making. Shared decision-making minimizes the informational and power asymmetry between specialists and patients by increasing patients' information, sense of autonomy and/or control over treatment decisions that affect their quality of life.

Research question: *What is the added value of a preference based value clarification decision tool and treatment advice to assist men (50+), who have to choose a treatment for localized prostate cancer, in comparison with existing care?*

Method: First, the data of the existing decision aid of the St. Elisabeth Hospital Tilburg is analyzed to define the additional value of a preference-based value clarification decision tool. Since the added value of the tool was determined at the very beginning of the research, the value clarification decision tool was designed. The design of the tool was based on Best-Worst Scaling experiments (BWS) case 2 and was executed in Limesurvey.
Two pilot tests were performed to evaluate this study. The first pilot test focused on the design of the tool. In the first pilot test seven participants (50+), who have problems with their urinary system, were included. After implementing the recommendations of the first pilot test, a second pilot test was constructed. Ten participants (50+) with a diagnosis of localized prostate cancer (PSA values $\geq$ 20 Gleason score $\geq$ 7) were asked to use and evaluate the value clarification decision tool. The tool was evaluated using (parts of) two valid scientific evaluation scales: the Decisional Conflict Scale (DCS) and the Satisfaction with Decision Scale (SWD).

Results: The overall result of pilot test 1 was that the made choice tasks were comprehensible. Recommendations were more time to fill in the questionnaire, decrease the number of choice tasks, mentioning the effects of some attributes in the statements and rephrasing the question.
Pilot test 2 showed that the participants rated the tool with an average note of 8.5 *(existing decision aid: 7.8)*. Almost all participants recommend the use of the tool to other patients. The willingness to use the

tool was high. The participants felt they were informed about the different treatments, benefits and adverse effects to make a well-informed choice and were very positive about the resulting advice. The participants felt more assisted in making an informed decision with the value clarification decision tool compared to the existing decision aid of the St. Elisabeth Hospital. Overall, an active role was taken and the value clarification decision tool was completed with little interference from their partner (or researcher).

Conclusion: Besides the improved patients' knowledge about localized prostate cancer, the value clarification decision tool also clarifies their preferences and generates a satisfied advice. Therefore, the value clarification decision tool can enhance shared decision-making and could be a valuable addition to existing health care.

Recommendations: Before the value clarification decision tool can be used in practice, the last recommendations of the participants have to be taken into account. Furthermore, a scientific experiment is advisable to confirm the results and increase statistical power. The early triggering of policymakers of existing health care for shared decision-making and the implementation of decision aids, such as the value clarification decision tool, is also recommended. The implementation of the value clarification decision tool can, besides assisting in answering the question *'Trick or treat?'*, also open a new area in existing care by empowering patients and creating new roles in the decision-making process between specialists and patients.

Keywords: value clarification decision tool, shared decision-making (SDM), decision aid, localized prostate cancer, stated preferences, treatment advice, Best-Worst Scaling experiments (BWS)

"It is up to the individual to choose the life he thinks is best."
(Simone de Beauvoir, writer / philosopher, 1908 - 1986)

1. Introduction

It is widely acknowledged that prostate cancer is one of the most common cancers among men in the world. In 2011, 22% of all men with cancer were diagnosed with (localized) prostate cancer (8 out of 1,000 men). Considering the ageing population, the expectation is that the number of men with prostate cancer will increase with 49% between 2011 and 2030 (1, 2).

Fortunately, the survival of patients with (localized) prostate cancer has improved greatly. The 5-year survival rate has increased from 66% in the period 1989-1994 to 87% in the period 2006-2010 (1). The mortality rate has dropped by more than 30% between 1995 and 2012 (3). This drop is due to early diagnosis (PSA testing) and expanding treatment options of cancer (4 - 6). Nowadays, less than 5 out of 1,000 men (0.5%) who are treated for localized prostate cancer will die within 15 years due to this disease (7). The treatment options vary between active surveillance strategies (AS) and curative treatments (radical prostatectomy or internal/external radiotherapy). Figure 1 shows the possible pathways that these patients can choose.

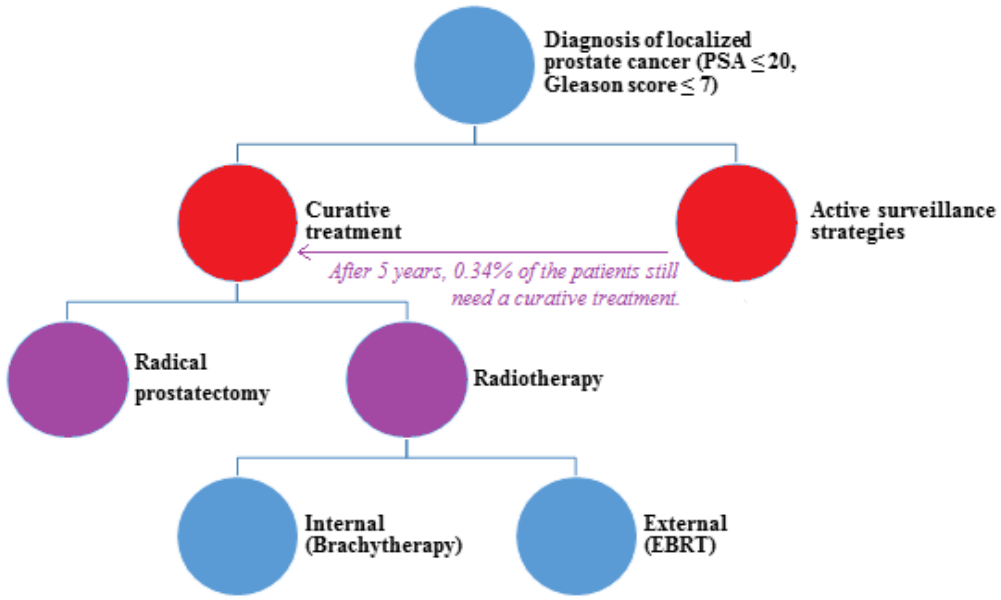

Figure 1: The possible pathways for patients with a diagnosis of localized prostate cancer (1).

The treatment options differ with respect to the benefits and adverse effects (complications). Adverse effects might have a negative impact on the quality of life of patients, and might be important

reasons for patients to opt for an alternative treatment (8, 9). The benefits and adverse effects of the curative treatment and active surveillance strategies (AS) are shown in Table 1 and 2.

Table 1: The benefits of the two main treatment strategies for localized prostate cancer (1).

Curative treatment	**Active surveillance strategies (AS)**
Direct action to cure the identified prostate cancer.	There are no side effects. The physical function remains the same.
There is no risk of being "too late" for a cure.	This strategy will prevent the patient for (possible) unnecessary treatments.
There is no regular schedule of PSA measurements and prostate biopsy, which can cause a lot of tension, stress and uncertainty in patients.	66 of the 100 men (66%), who are actively monitored, do not require curative treatment in the first 5 years.

Table 2: The adverse effects of the two main treatment strategies for localized prostate cancer (1).

Curative treatment	**Active surveillance strategies (AS)**
The patient may undergo unnecessary treatment which does not prolong his life.	The repeated PSA measurements and prostate biopsy can cause a lot of tension, stress and uncertainty in patients.
There is a risk of adverse effects and complications of treatments (e.g. risk of temporary or permanent urinary symptoms, erectile dysfunction after the treatment, intestinal complaints).	34 of the 100 men (34%), who have been actively followed in the first 5 years, still need curative treatment.
The treatment does not necessarily lead to a cure. If the PSA decreases first, but eventually will rise again, this is a sign of tumor growth. Between 20 and 30 out of 100 men, who have been treated, will experience PSA rise again.	In some cases, a curative treatment afterwards is no longer possible.

Due to the fact that every treatment has its own benefits and adverse effects in combination with (almost) no differences in effectiveness and mortality rate (0.5%) between the treatments, no consensus is reached about an optimal treatment for localized prostate cancer by specialists and researchers (7, 10). Therefore, the treatment decision is dependent of the patient's characteristics. The treatment decision differs for older adults (50+) in comparison to younger adults. Older adults are more likely to suffer from comorbidity of diseases and have (often) more sensory impairment, changes in cognition and lower levels of health literacy and technology, all of which has its influence on the treatment decision (11). In addition, older adults may have different priorities than younger adults (e.g. less willing to trade quality of life for survival prolongation) (12). This means that decisions concerning cancer treatments are often *'preference sensitive'*; there is no single alternative among the available options that is best for every individual, but instead the optimal decision depends on patient

preferences (13). The term *'patient preferences'* lacks a consistent definition. In this study, the preferences of a patient are defined as an individual appreciation of health outcomes in the healthcare, which results from cognition, experience and reflection (14, 15). Preferences provide directions for selecting a treatment, just like clinical guidelines do (14). Preferences are relevant because an opposite direction between the preferences of the specialist and patient may hamper the process and outcome of the decision-making process (16).

Historically, specialists were reluctant to inform the patient about the diagnosis localized prostate cancer, due to high mortality rate, fear of losing hope and limited treatment choices. Patient preferences were not taken into account (17, 18). The patient adopted a passive role and the specialist made the decision for a treatment based upon the data from patient's files and his own experiences (19). Adopting a passive role in health related decisions is more observed in men than in women. Women adopt by origin a more cooperative or active role in cancer management related to their own health or health of her family than men do (20 - 23). Figure 2 illustrates the valid and often used Control Preferences Scale (CPS), in which the different roles of the patient are presented (27).

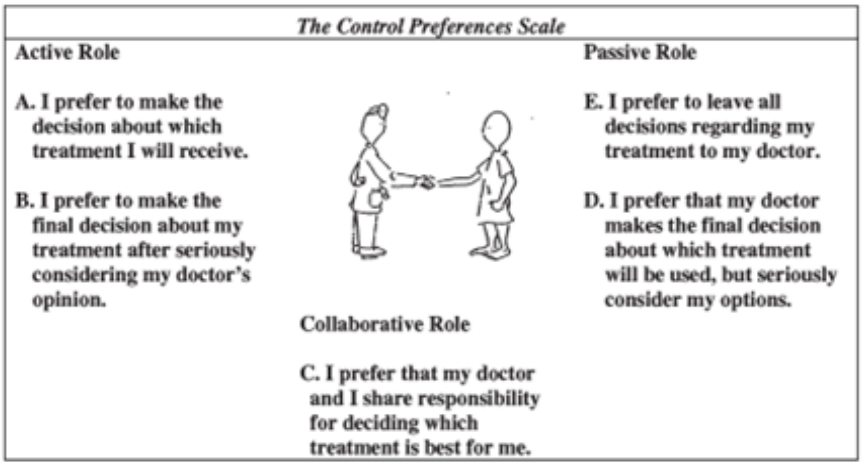

Figure 2: The CPS with the different roles in decision-making. (27)

The combination of the rise of certain social movements (public, patients and feminists) and changes in the organization of the National Health Services (NHS) has led to a growing awareness of the relevance of patient preferences in treatment decision-making (24, 25). Nowadays, more frequently an active role is chosen by the patient (men and women) in the decision-making process (26).

To integrate this active role and patient preferences in the decision process, shared decision-making (SDM) is introduced (28, 29). Shared decision-making (also called clinical decision-making) is a mechanism to minimize the informational and power asymmetry between specialists and patients by increasing patients' information, sense of autonomy and/or control over treatment decisions that affect

their quality of life (30 - 32). This requires a two-way exchange of information. The patient shares his treatment preferences with the specialist and the specialist uses his experience and knowledge to advise the patient. Together with the best scientific evidence, consensus will be built regarding the treatment choice (33). When both parties agree on the decision, the treatment can be started. Shared decision-making results in a higher adherence to the treatment, reduced decisional regret and a higher willingness of the patient to accept the treatment and the possible complications afterwards (34). Furthermore, shared decision-making leads to higher patient satisfaction and more functional status improvements (35 - 37). Notwithstanding all these benefits for shared decision-making, the consultations are usually brief without enough time to discuss all information that the patient needs to clarify his preferences for making his final decision (38, 39). As a result, many patients with localized prostate cancer still have insufficient knowledge over the different treatments and its outcomes and their preferences get little attention (40).

Many information devices are developed to provide information about (localized) prostate cancer as a supplement to the specialist before, during or after the consultation. An information device that, besides an informative summary of the main benefits and adverse effects also explicitly considers the personal importance of each benefit or adverse effect, is called a decision aid (41, 42). A decision aid can reduce decisional conflict and increase patients' participation in shared decision-making (15, 43). Recent studies have shown that decision aids are superior to standard counseling in improving patients' knowledge and realistic expectations about the results of treatments and other procedures. Perceived involvement, agreement between values and choices, and decisional agreement were increased after the use of a decision aid (44).

In the St. Elisabeth Hospital Tilburg (the Netherlands) a digital decision aid for localized prostate cancer is implemented. The purpose of this decision aid is to assist the patient in making his treatment choice. The decision aid provides information about all possible treatments. In addition, value exercises are added to determine the patient's personal opinion about the different treatments to personalize the tool to a patient's situation.

By providing extensive information in the decision aid, patients are able to make a well-informed decision, specialists are able to schedule the consultations in effective way, and more importantly, the communication between the patient and specialist will improve. This will enhance the process and outcome of the shared decision-making, which will lead to a higher satisfaction of the patient and specialist. However, in its current state, this tool does not elicit patient preferences for the different treatments. Furthermore, it does not process the results into a treatment advice. The decision aid only enhances shared decision-making through information provision.

A decision aid that can provide information, elicit patient preferences and give personal advice is the *value clarification decision tool*. With the value clarification decision tool the patient can discover his own preferences by assessing, exploring and determining what his personal preferences are and how this affects personal decision-making (45). It can assist the patient (and the specialist) in making a more informed and preference-based treatment decision, which relates to the quote of Simon de Beauvoir written years ago: *"It is up to the individual to choose the life he thinks is best."*

To study the added value of the value clarification decision tool to health care, the existing decision aid of the St. Elisabeth Hospital will be used as a basis for creating a value clarification decision tool. The value clarification decision tool will not only focus on providing information about localized prostate cancer, but also on providing a personal treatment advice based on the individual preferences. This advice can be discussed in shared decision-making. This leads to the following research question:

What is the added value of a preference based value clarification decision tool and treatment advice to assist men (50+), who have to choose a treatment for localized prostate cancer, in comparison with existing care?

The new value clarification decision tool will be evaluated by Dutch male participants (50+) [1] with localized prostate cancer. Their preferred role and the role of their partner in the decision process will be taken into account.

[1] In this study, a distinction is made between respondents and participants. Respondents are patients, who want to fill in a questionnaire. Participants are patients, who also want to participate in an observational study or interview. (46)

2. Analysis of the existing decision aid of the St. Elisabeth Hospital Tilburg

In this chapter the existing decision aid of the St. Elisabeth Hospital is analyzed. Based on this analysis, the value clarification decision tool will be developed. To avoid confusion in this chapter, *'the existing decision aid of the St. Elisabeth Hospital'* will be referred to as tool A, while the new *'value clarification tool'* will be referred to as tool B. After this chapter, the full names are used again.

2.1. General information about the two experiments carried out in the St. Elisabeth Hospital

In August 2014, the St. Elisabeth Hospital in Tilburg started with the first experiment (randomized trial) regarding tool A. To test this tool it was implemented in nine Dutch Hospitals. Six months later (February 2015), the primary test results from tool A were available. The dataset consisted of 99 respondents diagnosed with localized prostate cancer. One respondent had missing data for variables, which were essential for the analysis, so only the data of 98 respondents were included in this analysis. The dataset provided information about the prior preferences and preferences (choice) after the use of the decision aid (§2.2). This data played an important role in deciding to edit tool A, because it stressed the usefulness of a decision aid and highlighted the patient preferences for the different treatments.

The second experiment, executed in the period of August 2014 till March 2015, also yielded interesting insights. Its dataset consisted of 56 respondents diagnosed with localized prostate cancer. This experiment was executed to measure the need for a decision aid and the satisfaction with tool A, both of which are factors that play a vital role in deciding whether to design tool B or not (§2.3).

2.2. Analysis of prior preferences and preferences after use of the existing decision aid (tool A) *(only data of the first experiment has been useful)*

The dataset consisted of answers provided for 3-14 different statements. The number of statements was dependent of the respondents' preferences and/or the advice of the specialist. A respondent with a PSA of ≤ 20 and a Gleason score of ≤ 7 has a maximum of four options: active surveillance strategies, radical prostatectomy, internal radiotherapy (Brachy therapy) or external radiotherapy (EBRT). Respondents who got a specialists' advice for two treatments or choose a treatment with no choice options (like active surveillance strategies), were provided with fewer statements. These statements consisted of two possible choices. The respondent needed to move the point towards his preference (Figure 3).

Figure 3: Example of a statement in the existing decision aid (tool A) (in Dutch).

A personal preference score was estimated by using the answers for these statements. In some cases, one of the options was excluded due to the advice of the specialist. This led to the following results (N=98); 44 respondents (44%) were able to choose for active surveillance, 94 respondents (95%) were able to choose for radical prostatectomy, 73 respondents (74%) were able to choose for internal radiotherapy (Brachy therapy) and 78 respondents (79%) had the opportunity to choose for external radiotherapy (ERBT).

Before the respondents read the information of tool A, their prior preferences were questioned. Prior preferences are largely shaped by the environment through for example word- to-mouth information, media, internet or books. This information can broaden the knowledge of the patient, but also distort it when patients use unreliable and invalid resources (47). This can have a great impact on their final decision for a treatment. The existing decision aid (tool A) helps to overcome this problem by giving theoretically funded information that assists patients in making their final choice.

The prior preference was questioned to study whether the provided information has led to a change of the preferred treatment option. Changing the patient's preference was not the main purpose of the decision aid, but the information provision may have influenced the final decision of the patient. Table 3 shows the preference of the respondents prior to the use of the decision aid.

Table 3: Question: 'Which treatment options do you prefer the most, before you used the decision aid?'

	Frequency	**Percentage**
Active surveillance strategies (AS)	21	21%
Radical prostatectomy	30	30%
Internally radiotherapy (Brachy therapy)	12	12%
Externally radiotherapy (EBRT)	10	10%
No preference	25	25%
Missings	1	1%
Total	99	100%

Table 3 indicates that 21 respondents (21%) chose for active surveillance strategies (AS) and 52 respondents (52%) opted for a curative treatment (radical prostatectomy, 30 respondents; internal radiotherapy (Brachy therapy), 12 respondents; external radiotherapy (EBRT), 10 respondents). A quarter of the respondents (25%) had no preference beforehand. Tool A should help these respondents making their choice. For the remainder, the tool should help strengthen (or in some cases should change) their choice. Of the 73 respondents (74%), who showed clear prior treatment preferences, 51 of them (70%) chose the same treatment after use of tool A. The tool strengthened their choice. 25 respondents (26%) had no prior preferences beforehand, but after using tool A 14 of them (56%) had a preference. For these respondents, this tool assisted them in making a choice. Only 12 respondents (12%) changed

their choice. The preferences indicated before and after use of tool A are shown schematically in Table 4 and are graphically depicted in the two circle diagrams in Figure 4.

Table 4: The prior preferences and preferences (choice) after the use of the decision aid (tool A).

Prior preferences:	***N***	*Preferences after use:* AS (N)	*Preferences after use:* Radical Prostatectomy (N)	*Preferences after use:* Internal Radiotherapy (N)	*Preferences after use:* External Radiotherapy (N)	*No preferences after use* (N)
Active surveillance strategies (AS)	**21**	16 (76%)	1	0	1	*3*
Radical prostatectomy	**30**	1	25 (83%)	0	0	*4*
Internal radiotherapy (Brachy)	**12**	0	0	6 (50%)	0	*6*
External radiotherapy (EBRT)	**10**	0	0	1	4 (40%)	*5*
No preferences	**25**	4	8	2	0	*11 (44%)*
Total (100%)	***98***	*21*	*34*	*9*	*5*	*29*
Missing	***1***					

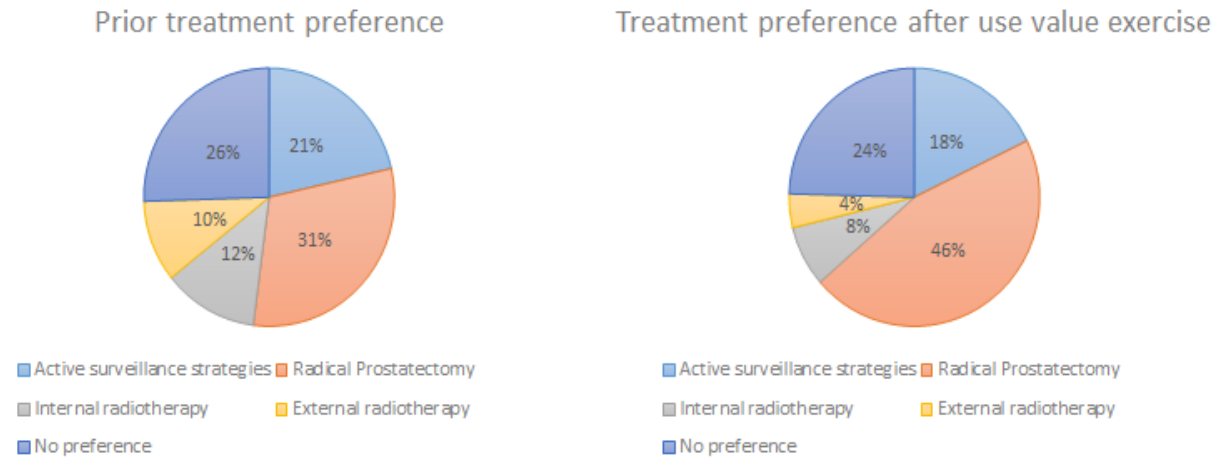

Figure 4: Two circle diagrams with the indicated preferences before and after the use of the decision aid (tool A).

To make Table 4 and Figure 4 more understandable, Figure 5 shows how the different treatments and results are related to each other.

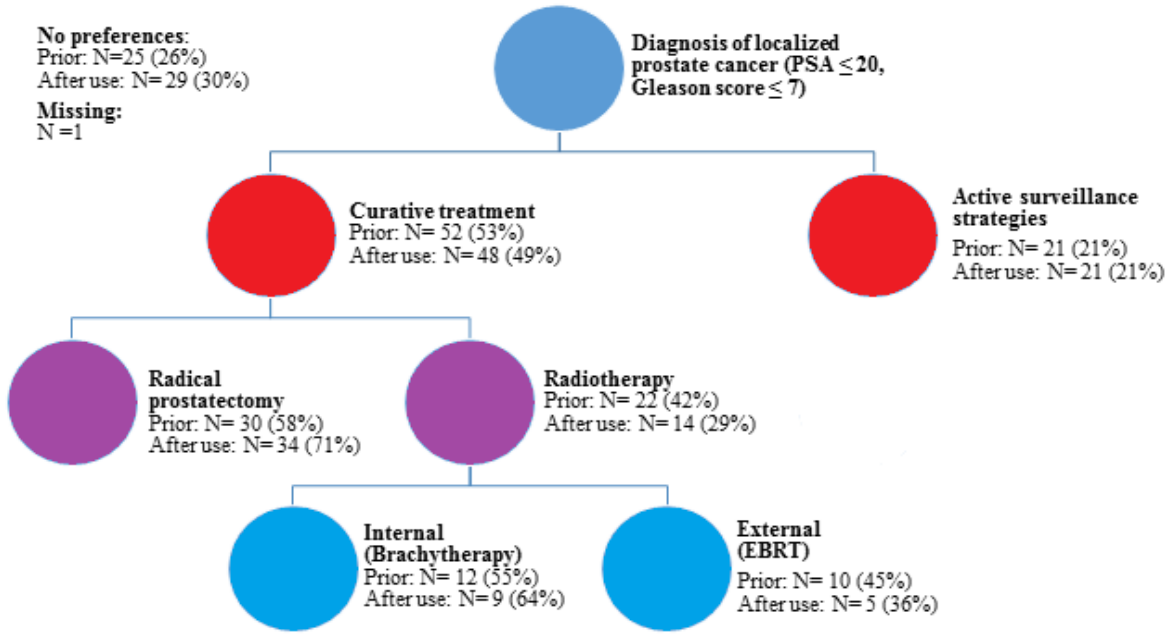

Figure 5: Figurative representation of the preferences (N=99).

Remarkable in these figures is that 29 of the 98 respondents (30%) were indecisive after the use of tool A. For 11 of them (44%) the tool did not show value, these respondents remain indecisive regardless of the use. For these respondents an added advice could be a solution. Since the opinion of the specialist reflects the prior preferences of the respondents in only 51 of the cases (52%), this is especially recommended.

In conclusion, the existing decision aid (tool A) assisted the decision process of the patient by providing information about the different treatment options, the benefits, and the risks. The existing tool strengthened the choice of the patient and enabled the patient to make an informed choice, when he was indecisive beforehand.

The value clarification decision tool (tool B) takes it one step further. Besides providing information, this tool will also advise based on personal patient preferences. The question is whether the patient wants (and will use) a new value clarification decision tool based on their personal preferences. Otherwise, the design of the value clarification decision tool might be a waste of time.

2.3. The need for a new decision aid (tool B) that provides information and produces an advice *(datasets of both experiments were used)*

The results in the previous paragraph have shown that in most situations tool A corresponded the prior treatment choice of the respondent. With this information, the patient is confident and well-informed. However, the patient could also interpret this result as a waste of time due to the corresponding preference. Another reason could be that he wants to take a passive role in the decision process and therefore does not see the value of a decision aid. Thus, it is important to know if the patient is willing to use the value clarification decision tool (tool B) and if he is willing to put effort into doing an exercise to elicit his preferences. To study this, the need for tool B is determined by using the results of both datasets of the existing decision aid (tool A).

In the first experiment (N=98), respondents answered statements about the willingness to use tool A, when it was advised by the specialist. 97 respondents (98%) were willing to use the tool, regardless of the time or effort it takes. Most of them (86%) were very satisfied and would recommend tool A to others. The tool was graded with an average grade of 7.8 on a scale of 0 (very insufficient) till 10 (excellent). Improvements can be made in the items with a lower score. The second experiment (N=56) showed that the improvement should focused on support in decision-making and understanding of the benefits and risks. 31 respondents (59%) indicated tool A assisted them in making their decision and 34 respondents (61%) were stimulated to think about the benefits and adverse effects. Furthermore, 38 respondents (68%) indicated tool A assisted them to clarify which aspect were most important and will influence their choice. Based on these results, measuring preferences is recommended to increase their thinking about the benefits and risks and more importantly their understanding what the most important aspects for them are. The results of the first dataset showed that 97 respondents (98%) appreciated consideration of their own preferences.

In the second experiment (N=56), the respondents were questioned about how preferences could be measured. Receiving a treatment advice is not something the respondents thought about at first. So, this question did not lead to a convincing recommendation for tool B. Only one third of the 56 respondents (36%) prefer an advice after using the decision aids, the other respondents had no opinion (32%) or did not want an advice which clarifies their preferences (32%). The value of a treatment advice in a decision aid, like tool B, needs to be proven.

As indicated earlier the willingness to use a decision aid, the importance of preferences and a preference-based treatment advice are largely dependent of the role the patient wants to take in their decision process (active, collaborative or passive). 74 of the 98 respondents (74.5%) prefer to make the choice for treatment on their own or in consultation with the specialist. These respondents chose a

collaborative or active role and want to engage in shared decision-making. For them, tool B including an advice can be useful and recommendable.

Based on this analysis, the existing decision aid (tool A) is altered into a value clarification decision tool (tool B). The information and the digital accessibility of tool A remained the same. New features were built in to optimize tool A and enhance the assistance in shared decision-making. Shared decision-making has a lot of benefits (mentioned in the introduction) such as increasing patient satisfaction, reducing decisional regret and more functional status improvements (34 - 37). The main requirements are summed up.

- The value clarification decision tool (tool B) should stimulate patients more to think about the treatments, benefits and adverse effects and create a better understanding of the most important outcomes for a patient than the existing decision aid (tool A). For clarifying preferences, understanding the treatment and its outcomes is necessary to engage in shared decision-making.
- The value clarification decision tool (tool B) has to clarify their preferences with a personal preference-based treatment advice. The advice serves as a reward for the efforts the patient takes to complete the process of appreciation of all different preferences. The advice should show the most preferred treatment, which can be used in shared decision-making. The specialist and patient can discuss this and make a more personal preference-based choice.
- The value clarification tool (tool B) also has to measure the strength of the preference, which could be useful for the specialist. If the scores are predominantly mild or neutral, this may indicate either that these patients are truly preferentially neutral about the treatment options or that some patients are uncertain about their attitudes towards the treatment options of interest. These patients need another role of the specialist in shared decision-making than patients with a predominantly strong preference (e.g. more information disclosure/ advice specialist).
  The strength of preference can also indicate if the given advice matches with the final treatment of the patient. This is interesting for the researcher. If the score in the value clarification decision tool indicates a predominantly strong preference for a treatment which is not his final one, this might indicate that the value clarification decision tool is not valid or unreliable (or the patient chooses the wrong treatment). The underlying cause needs to be discovered.

During the development phase, the patient population characteristics (e.g. age, cognitive burden, low technology knowledge) and ease of use were taken into account. In the methods section is written how the value clarification decision tool (tool B) was developed and which choices were made to ensure that the requirements are met.

3. Methods

3.1. Design of the value clarification decision tool

Medical decision-making tools were used for the development of the value classification decision tool based on patient preferences. Medical decision-making tools help to enhance transparency of the decision process and increase patient centeredness (48, 49).

The designed value clarification decision tool for localized prostate cancer exists of two parts. The first part consists of personal medical data (e.*g. PSA values, Gleason score*), general information about prostate cancer and the four different treatment options *(e.g. explanation prostate cancer, risks and benefits of each treatment*). This medical data was obtained from previous visits to the specialist. The general information was similar to the existing decision aid of the St. Elisabeth Hospital. The content remained the same to retain the scientific basis. This part was preserved, because in the second experiment 56 Dutch patients evaluated the existing decision aid of the St. Elisabeth Hospital very positively on the readability and clarity of this information.

The second part is different from the existing decision aid. In this part all mentioned improvements/requirements of chapter 2 were adapted. This design needs further explanation, which is explained below.

3.1.1. Background of preferential technologies

For the development of a scientific value clarification decision tool based on patients preferences, different preferential technologies could be used, like rating scales (real score), ranking scales (implicit score) or choice-based methods (50). These methods have their background in decision theory. Decision theory identifies two sorts of preferences; stated and revealed preferences (51, 52). Revealed preferences are derived from observations of the behavior of the patients. Stated preferences are derived from questions focusing on the behavioral intentions or choices of patients on existing and /or hypothetical products or services (53). This study elicits stated preferences, because the behavioral intentions/choices of patients with localized prostate cancer were the main topic on which the final treatment advice was built.

Figure 6 shows an overview of possible patient preferential technologies. The red pathway represents the choices made in this study. These choices were based on the characteristics needed for the value clarification decision tool (e.g. eliciting preferences, creating an advice, easiness to understand and low cognitive burden), the theoretical relevance and the practical feasibility. The final choice was made in consultation with two medical specialists and two researchers specialized in decision methods and –analysis.

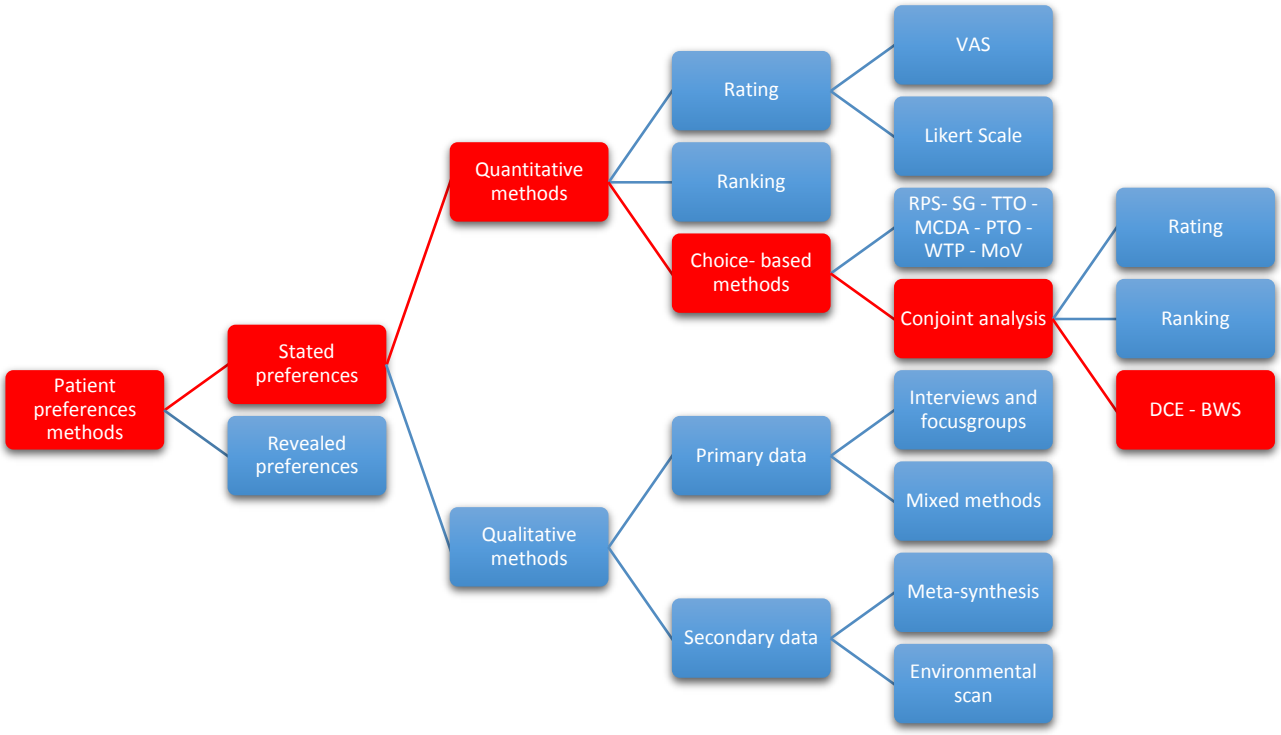

Figure 6: The different methods for stated preferences (52 - 54).

A quantitative method based on measuring process characteristics was chosen. A quantitative method can generate an advice based on the indicated preferences of the patient, without interference of a specialist (55). As a result, the tool could be made digital, one of the criteria for the value clarification decision tool (tool B).

Choice-based methods fit best with the purpose of the tool (eliciting preferences and generating a treatment advice) and the patient population (50+). The advantage of choice-based methods is that these methods allow estimating the marginal value of changing attributes as well as the total value of a good (56) This means that the value of every benefit/ adverse effect and the total value of the treatment could be determined. With the use of choice-based methods a preference-based advice that shows the most preferred treatment, the strength of the preference and the most important characteristics of a treatment, could be generated.

3.1.2. Conjoint analysis

One of the choice-based methods is the conjoint analysis. Conjoint analysis is a statistical technique to measure preferences for product features to identify their implicit preferences. This theory has a strong theoretical background and is based on the theory of Lancaster (57). This theory assumes that

products/services are described by attributes; aspects that characterize alternative possibilities considered in a decision (58). These products/services enter an individual's utility function as a combination of the attributes (59). A conjoint analysis consists of ranking, rating and discrete choice experiments (DCE) (see Figure 6) (50). These scenarios are evaluated by rating or ranking the scenarios or by choosing scenarios. With the rating method, the patient has to rate each scenario on its attractiveness (not attractive - very attractive). With the ranking method, the patient has to rank different scenarios in order of attractiveness. In the DCE, the patient has to choose their preferred scenario from two or more scenarios (60). Within the conjoint methods, the DCE is gaining widespread use in health care and has been applied in a number of areas including; eliciting patient / community preferences in the delivery of health services (57, 61 - 64). A relatively new variant of this, the Best-Worst Scaling experiments (BWS), is on the rise.

3.1.3. Best-Worst Scaling experiments

Recent literature shows that Best-Worst Scaling experiments (BWS) offer a valid method to measure the relative strength of preferences for the designated attributes of the scenarios (65, 66). The patient chooses the most and least preferred attributes from a list of three or more presented in a single choice task. The attributes can be split into levels. For example, the attribute intestinal problems of the treatment option radiotherapy, consists of two levels; level 1 = 0% and level 2 =15%. The difference in the levels of each attribute varies independently over the choice tasks. All attribute levels occur within the choice task at an equal frequency, to maximize the efficacy (67). The use of levels is dependent of the chosen case. There are three cases of BWS; case 1, case 2 and case 3. BWS case 1 only shows the attributes to the patient and does not use the levels. The patient has to choose the one that is best (most attractive) and the one that is worst (least attractive) (68, 69). This case is useful to measure the relative values associated with each list of attributes. The choice sets of BWS case 1 has the similar purpose to those in traditional DCE; with suitable assumptions, it could facilitate inferences about the value of one attribute associated with the other attributes. The overall score of a treatment could not be calculated. In comparison with BWS case 1, BWS case 2 incorporate the attributes and the levels. In each question the levels of the attributes change. This results in a separate measurement of the attribute weight and a separate value for every treatment. With case 2, the values of every treatment could be weighed against each other and a treatment score could be calculated. BWS case 3 presented a specific choice task design (multi-profile) of a treatment option with all attributes (and levels). The patient is not just asked to weigh the attributes (with or without levels), but point out the best profile. This is considered as a heavier method than case 1 or case 2, not only for calculating the scores of the treatments, but also for the patient. The patient has to choose a whole profile instead of defining what he likes best/least about a certain profile (70).

Due to the complexity of the health care decision and personal characteristics of the patients diagnosed with localized prostate cancer, BWS case 2 was chosen. BWS case 2 is not only a straightforward method because the patient is not required to quantify their strength of preference for a total profile, it also resembles real life decisions more than BWS case 1 does and overcomes many of the reliability issues that are inherent with simple rating (71).

3.1.4. Design: Value clarification decision tool with Best-Worst Scaling experiments case 2

As a guideline for the design of the BWS case 2, the checklist by The International Society for Pharmacoeconomics and Outcomes Research (ISPOR) was used (58). The checklist was originally designed for conjoint analysis. For BWS is no checklist available yet. This checklist is used because BWS is a variant of the conjoint analysis, so the data of BWS closely resembles the data of the conjoint analysis.

For editing the existing decision aid of the St. Elisabeth Hospital with eliciting patient preferences and providing an advice, this research focused on the first four phases of the checklist.

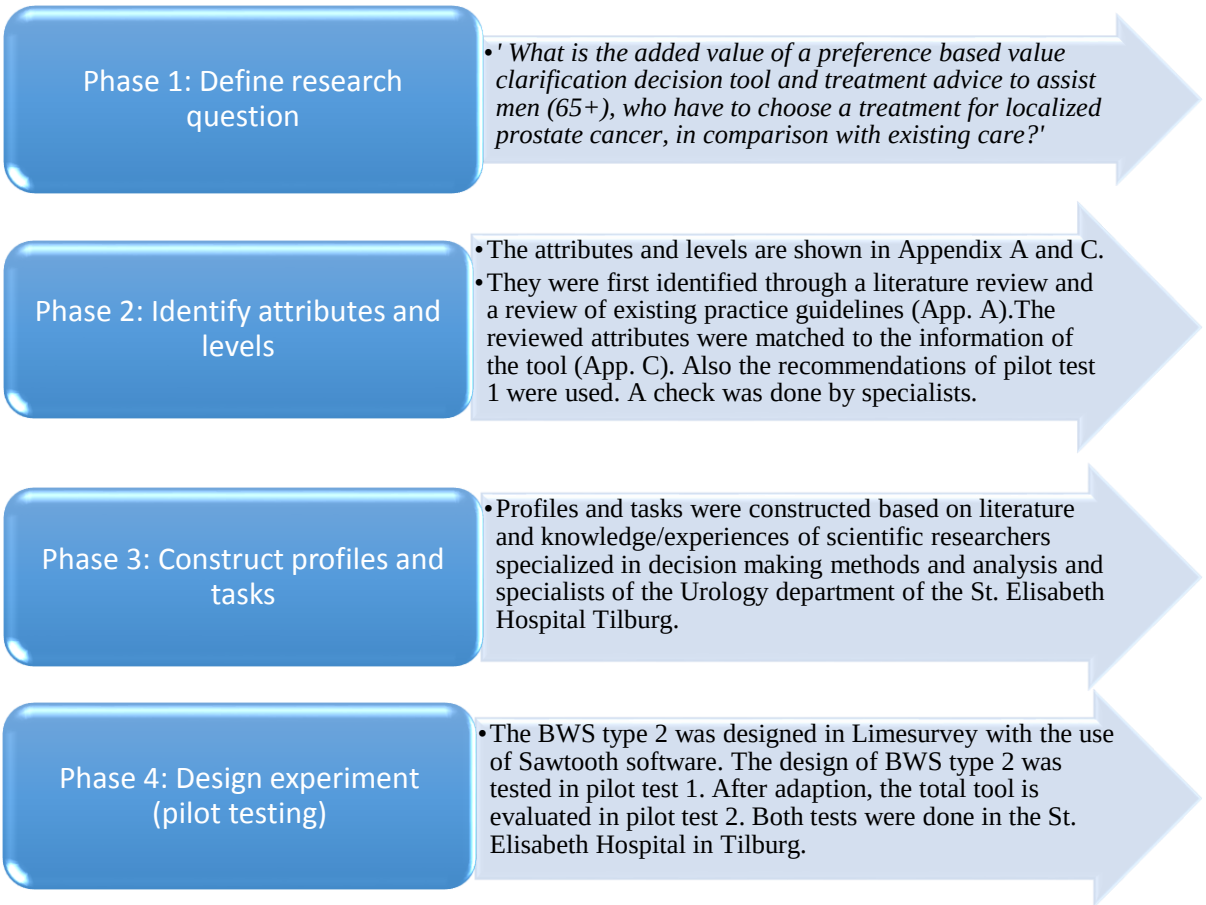

Figure 7: The checklist by The International Society for Pharmacoeconomics and Outcomes Research (ISPOR) (58).

The subsequent other five phases in this checklist focusing on executing an experiment (phase 5), developing an instrumental design (phase 6), collecting data (phase 7), statistical analysis (phase 8) and drawing conclusions (phase 9) were not included in this study. The main purpose of this study was to develop a value clarification decision tool which elicits patient preferences and gives an advice, not generating conclusions about patient preferences (for which phase 5 till 9 are useful).
Below, the first four included phases are described in more detail.

Phase 1: Define the research question

The research question is defined after a literature review and conversations with researchers specialized in decision methods and –analysis and medical specialists of the St. Elisabeth Hospital Tilburg. These conversations took place, because the study should contribute to a larger study funded by the Delectus Foundation and St. Elisabeth Hospital Tilburg, which aimed using high-tech applications in exploring ways to improve the quality and accessibility of care in Urology to meet the growing demand for health care calls for new and creative solutions. This led to the following research question: '*What is the added value of a preference based value clarification decision tool and treatment advice to assist men (50+), who have to choose a treatment for localized prostate cancer, in comparison with existing care?'*

Phase 2: Identify Attributes and levels

In this phase attributes and levels were identified. Attributes may establish preferences about structural aspects of the healthcare (e.g. preferences for information in decision-making), process aspects (e.g. usability of the digital decision aid) and/or outcomes of the health actions (e.g. efficacy or adverse effects of a treatment). All three aspects were represented in this study (14).

The attributes (with the levels) were initially identified through a literature review, analysis of the existing decision aid and a review of existing practice guidelines. (72, 73) The attributes were discussed with specialists of the St. Elisabeth Hospital. All attributes were checked by specialists and researchers on their relevance for the patient and the decision environment (33). An attribute, like *'Risk for being too late for another treatment'*, could be very important for the patient but when it makes no distinction between two treatments (decision environment) it is not relevant. The value clarification decision tool cannot indicate the preference of the patient for a specific treatment.

Once the attributes were identified and checked, levels were designated. The levels could be categorical (e.g. gender), ordinal (e.g. treatment frequency) or ratio/interval scaled (e.g. change of a certain outcome or adverse effect) (74). These three variants were all integrated in this study. The attributes and levels of the different treatment choices are shown in Appendix A: *'The initial attributes and levels.'*

Phase 3: Construct profiles and tasks

Phase 3 focused on designing the value clarification decision tool (tool B) itself. In this phase the choice tasks with statements were constructed, based on the attributes and levels showed in Appendix A. A balanced spread of the attributes (no dominated statements) and low correlations between the attributes (orthogonal) were important for valid and reliable results (75). A high statistical yield and reasonable tasks maximize the statistical value of a design. Unlikely combinations of attributes should be avoided to overcome problems with understanding and choosing between two treatments (76). This resulted in three partial factorial designs. A partial factorial design (or d-efficient design) is characterized through the use of a subset of attributes and levels instead of all attributes and levels, which is the case in a full design (55). With a partial factorial design, all attributes and levels could be used without heightening the cognitive burden of the patient. The designs were made with Sawtooth software to create a full-balanced design. The designs existed all of eight choice tasks with varying number of statements (5, 6 and 7 statements), dependent of the number of attributes and levels.

In the three partial factorial designs, patients were asked to indicate the most preferred (best) and least preferred (worst) characteristic of treatment. An example of a choice task is shown in Figure 8. The number of patients who suffer from adverse effects is indicated in an X-in-100 format, because this correlates with the scientific information of the tool. An example of a question is: *'10 out of 100 patients suffer at (urinary) incontinence due to the treatment.'*

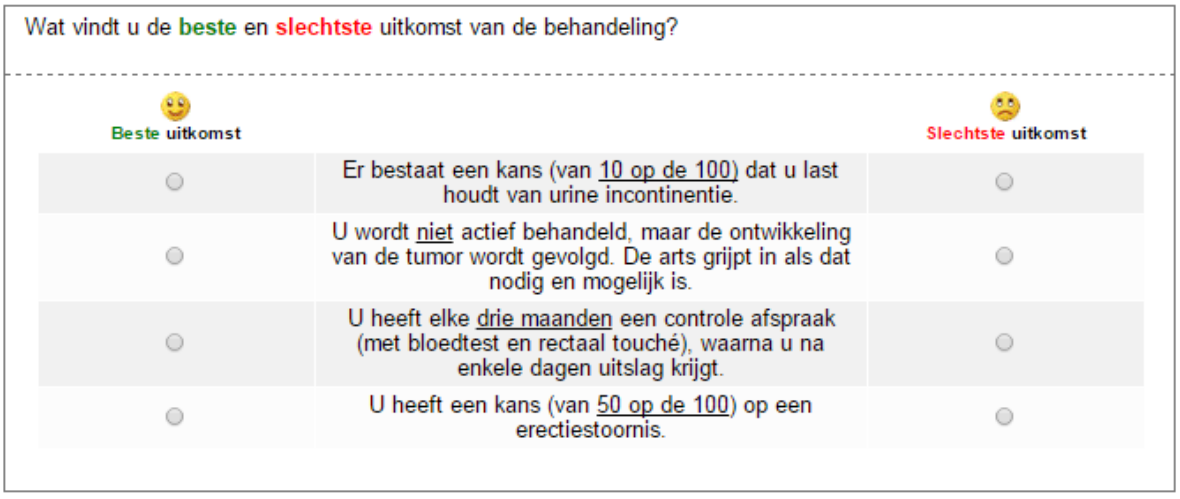

Figure 8: Example of a choice task in the first experimental design of the value clarification decision tool: Active surveillance strategies (AS) - curative treatment (in Dutch).

Besides the information, also the sequence of the existing decision aid is maintained, because 41 of the 56 respondents (73%) indicated that the existing decision aid was very accessible. So, the number of statements was also in the value clarification decision tool dependent of the respondents' preferences and/or the advice of the specialist. Respondents who got a specialists' advice for two treatments or choose a treatment with no choice options (like active surveillance strategies), were

provided with fewer statements. The patients continued until one treatment get the highest score, his preferred treatment/ his advice.

Three summaries were built in at the end of every choice task. These summaries illustrate the strength of preference and most preferred treatment based on the made choice task. After the summary the next choice task (when the preferences are not clear) or the final advice followed (when the preferences are clear). With a minimum of one and a maximum of three choice tasks the value clarification decision tool ended. An example of a summary and final advice is shown in Figure 9.

Figure 9: Example of a summary after completing the first experimental design 'Active surveillance strategies (AS) - Curative Treatment' with a final advice the patient can get. Here: AS or External Radiotherapy (EBRT) (In Dutch).

With the scores between the treatments a separation can be made between mild, moderate and strong preferences; the strength of preferences. In Appendix D: *'Calculation of the strength of preferences'* is explained how this calculation is done.

Phase 4: Design experiment (pilot testing)

For designing an experiment, pilot tests are useful. A pilot test is a small scale preliminary study conducted in order to evaluate feasibility, time, cost, adverse effects, and effect size (statistical variability). A pilot test can indicate the improvement upon a study design before the performance of a full-scale research project can be generated (77). In this study, two pilot tests were executed. In *§3.2 'Pilot testing'* these pilot tests are further explained.

3.2. Pilot testing

As mentioned in the paragraph before, two pilot tests were done. The first pilot test only focused on the design of the value clarification decision tool. This pilot test was included because the literature showed that the cognitive burden, due to task complexity and fatigue associated with the statistical efficiency and plausibility of profiles, is one of the biggest limitations of BWS (78). This cognitive burden is higher when the number of attributes and levels increases and can cause an overall error in the survey response caused by the high number of attributes and levels (79). To minimize this limitation in the value clarification decision tool, the statistical efficiency and the cognitive burden must be balanced. Another factor which is tested in this first pilot test, is the combinations of attributes. Unlikely combinations should be avoided, because it can confuse patients and could increase anxiety, insecurity and loss of control, which led to dysfunctional coping and reduction of processing complex information (80 - 84). As consequence the validity and reliability of the tool will go down. The combinations of attributes should create a plausible profile. So, in the first pilot test the match between the format of the tool and the patient group with localized prostate cancer (50+) was clarified to decrease the limitations of BWS. When the results of the first pilot test were integrated in the value clarification decision tool, a second pilot test was performed to evaluate the customized value clarification decision tool. The evaluation focused on the question whether the given advice based on the preferences of the patients assist them in their decision process. Also an indication of the feasibility of the value clarification decision tool for the patient was measured.

This paragraph explains how and by whom the two pilot tests were conducted. This part is divided into different sections; instrument (§ 3.2.1), participants (§ 3.2.2), procedure (§ 3.2.3) and analysis (§3.2.4). The sections explain these topics for pilot test 1 and 2 in more detail.

3.2.1. Instrument

Pilot test 1

The first pilot test took place in the development stage, so the choice tasks of the value clarification decision tool were printed on paper. The participant got an explanation about the tool and had to make four randomly selected choice tasks in presence of the researcher (and specialist). After these choice tasks of the tool, an in-depth topic interview started to retrieve the participants' experiences about the design. An in- depth topic interview is an open-ended, discovery-oriented method that is well suited for describing both treatment processes and outcomes from the perspective of the target audience (or key-stakeholder). With an in-depth topic interview, as much opinions, feelings and thoughts as possible can be generated (85). Measuring subjective feelings and thoughts is difficult with quantitative and measurable variables, especially a personal topic, such as localized prostate cancer. People will give less information about such topics in questionnaires (86).

The required time for the first pilot test was short, due to the fast consecutive consultations of the participants. A waiting time of maximum 10 minutes after consultation was the directive. The

expected duration of the pilot test was +/- 10 minutes. When the researcher (and specialist) noticed that another participant was waiting for more than 10 minutes, they slowly bring the pilot test with the existing participant to an end. Because of the short amount of time, five open topic questions were prepared. The topics were based on the participant's first impression, the difficulty, the length and the design of the questions and statements. For the questions see Appendix B: *'Interview questions of pilot test 1'.* The researcher made notes of the answers participants gave.

Pilot test 2

The second pilot test was conducted after the recommendations of pilot test 1 were implemented and after the tool was converted in a digital online survey program, called Limesurvey.

For this evaluation are brief and valid questionnaires desirable: these case of questionnaires increase the response rate and limit the cognitive burden on the patient, which is particularly important for cancer patients of older age. (87, 88) The evaluation questions of this study were based on two valid scientific evaluation scales; the Decisional Conflict Scale (DCS) and the Satisfaction with Decision Scale (SWD). These scales measure the content of the value clarification decision tool (information and advice) and patients' satisfaction with the tool. Both scales use a Likert-case response (1= strongly disagree through 5= strongly agree). The use of the same response scales increases clarity, ease of use and willingness to answer (87).

The Decisional Conflict Scale (DCS) measures personal perceptions and modifiable factors contributing to uncertainty in choosing options and expressing satisfaction with the choice (89, 90). Modified factors could be; feeling uninformed, unclear about personal values or unsupported in decision-making. All these factors influence effective decision-making (90, 91). The 16-item DCS is useful for evaluating the effectiveness and efficiency of the value clarification decision tool by making a treatment decision through patients with localized prostate cancer.

An addition to the DCS is the Satisfaction with Decision Scale (SWD). The SWD measures satisfaction related to the participation in the decision with a perception of having received the correct information (92). By using this scale, the satisfaction of the treatment information and the created personal advice could be more explicitly evaluated. Another reason for using the SWD is that this scale also gives information about the likelihood that the patient sticks to the given treatment advice. Several versions of the SWD are available. The version used in this study was developed for postmenopausal hormone-replacement therapy decisions. This six- item scale has excellent reliability (Cronbach's alpha= 0.86) and established a high correlation between satisfaction and '*decisional confidence*' (0.64) (93).

The most important questions of these two scales are used. These questions are selected by the researcher, based on overlap between the two scales and the main topics of the evaluation (clarifying understanding of the adverse effects/benefits and treatment, the opinion about a preference-based

treatment advice and an indication about the feasibility of the tool). An example of the chosen evaluation questions processed in Limesurvey is shown in Figure 10.

Two questions were discussed after finishing the Limesurvey;

1. Do you want to use the value clarification decision tool, regardless of the time and effort it takes?

2. Do you have the feeling that the value clarification decision tool could assist you in making your decision?

These questions were used to make a comparison with the existing decision aid and were added after finishing Limesurvey and before pilot test 2 started.

In hoeverre bent u het eens/oneens met onderstaande stellingen?

	Helemaal mee eens	Mee eens	Mee eens/mee oneens	Mee oneens	Helemaal mee oneens
Ik ben op de hoogte van de verschillende behandelmogelijkheden voor prostaatkanker.	○	○	○	○	○
De informatie die ik heb gekregen over de behandelmogelijkheden is volledig.	○	○	○	○	○
Ik ken alle voordelen van de verschillende behandelingen.	○	○	○	○	○
Ik ken alle nadelen van de verschillende behandelingen.	○	○	○	○	○
Het advies dat ik heb gekregen komt overeen met mijn voorkeur.	○	○	○	○	○
Ik verwacht dat ik bij het gekregen advies blijf.	○	○	○	○	○
Ik vind het eenvoudig om een behandelkeuze te maken met deze keuzehulp.	○	○	○	○	○
Ik vind de tijd die ik kwijt ben aan de keuzehulp acceptabel.	○	○	○	○	○
Ik heb de vragen in de keuzehulp gemaakt zonder hulp van anderen.	○	○	○	○	○
De keuzehulp is van toegevoegde waarde ten opzichte van de overige informatie.	○	○	○	○	○
Ik zal het gebruik van de keuzehulp aanbevelen.	○	○	○	○	○

Figure 10: An example of the integrated evaluation in Limesurvey (in Dutch).

A final supplement of this evaluation is the generation of test results (and advices) of a number of participants. The participants had to fill in the value clarification decision tool first, before they could evaluate the tool. These results (and advices) were analyzed to see whether the advice corresponds with their indicated preferences and final treatment.

The indicated test time for the second pilot test was +/- 40 minutes. In this test were no waiting times, because the researcher planned enough space between the visits. No time limit was used in order to resemble the home situation as closely as possible.

In both pilot tests, the participant had to think out loud, so the researcher could actively follow his thoughts. The think-aloud technique may be the most widely used method in usability testing (94). This technique ask participants to think out loud and say all the things that come to mind. The thinking-aloud technique is used in order to obtain additional information and points of interest to improve the value clarification decision tool.

3.2.2. Participants

The participants in both pilot tests were recruited through the use of purposive sampling. Purposive sampling is a research technique in which research units are selected on pre-defined criteria. The purpose of this technique is twofold. First, purposive sampling must ensure that the targeted selection exists of patients, who have as many as possible relevant characteristics in the light of the research question (inclusion criteria). In addition, it has to provide enough diversity to take the impact of the different patient characteristics into consideration (95).

In this first pilot test, the used inclusion criteria were men with problems with their urinary system, an age of 50 or older and Dutch speaking. The diversity was checked on age to ensure that the tool is usable for all older age categories. In the second pilot test, the inclusion criteria were men with a diagnosis of localized prostate cancer (PSA values ≥ 20 Gleason score ≥ 7), aged 50 years or older and Dutch speaking. Diversity was ensured among the sample. The diversity was checked on the advised treatment options of the specialist, to avoid that every participant ends in the same set of the value clarification decision tool.

Next to these participant characteristics, the accessibility of the target group and the size of the amount of participants were taken into account (96). The accessibility of the target group in pilot test 1 was determined through the urologist after consultation and checked through the secretariat before the researcher (and other involved specialist) were called. The accessibility in pilot test 2 had already been established at the end of testing the existing decision aid of the St. Elisabeth Hospital a couple months ago. At the end of these two tests, respondents were asked to sign informed consent to participate in follow-up research to further improve the existing decision aid. The list of these respondents was retrieved and a random number of patients was selected. Before the selected patients were approached, a check was done in the medical database to view the progress of their treatment process and their personal situation. When his treatment process and his personal situation seemed normal, the patient was contacted by telephone whether he was willing to participate in this study. The size of the amount of participants in a pilot study, often causes confusion. No minimal number is indicated (95). A minimum of five participants was chosen in the first pilot test, because the purpose of this study was only to evaluate the design. When more patients could or would participate, they were accepted. In the second pilot test, a minimum of ten participants was set, fewer participants would jeopardize the validity and reliability of the results.

3.2.3. Procedure

Both pilot tests were executed in the St. Elisabeth Hospital in Tilburg, because this tool will be first implemented in this hospital. Each pilot test was conducted by the same researcher and specialist.
In the first pilot test, a specialist asked if the patient wanted to join in the pilot test after their consultation. When the patient agreed, the researcher and another specialist were cabled through the secretariat to pick up the patient. In the second pilot test, the participants were asked by telephone to visit the St. Elisabeth Hospital and report their presence to the secretariat of the Department of Urology. The secretariat contacted the specialist and researcher, who called the participant in.

Both tests were performed in one of the treatment rooms of the urology department of the St. Elisabeth Hospital to control interfering factors and enhance the reliability. The purpose of the tool, the pilot test and the ethical rules were explained. Also, the think- aloud technique and the observer role of the researcher (and specialist) were clarified. In the second pilot test, some extra explanation was given. To the participant was explained that the questionnaire in Limesurvey was the leading tool in this process. The information was provided on paper due to the limitations of Limesurvey. If the participant encountered difficulties during this pilot test, he could contact the researcher and/or specialist. The participant used the tool individually or in the presence of his partner. Participants were given a beverage to make them feel more comfortable. After pilot test 2, the participants received a box of chocolates to show the appreciation for their participation and effort to visit the St. Elisabeth Hospital.

3.2.4. Analysis

The analysis of both pilot tests is almost the same, because both used the think- aloud technique. All observation notes and quotes were analyzed. The most important quotes got a label; open encoding of codes and labels. (95) These codes were encrypted and organized axially around some categories (e.g. pilot test 1: *first reaction, difficulty questions, number of statements etcetera*) (97, 98). Organizing axially around categories makes it easier to draw conclusions from the data (99). The main points were selected by the researcher and used for evaluating the design of the value classification decision tool (pilot test 1) or as support and clarification of the answers of the digital evaluation questions (pilot test 2).[2] The data of the digital evaluation questions were analyzed in Excel. One of the biggest limitations of a coding process in qualitative research is the influence of the subjectivities of the researcher (100). To overcome this limitation, another neutral researcher, not directly involved in this study, checked the codes and the coding process.

The method sections are summarized and shown schematically in Table 5. A separation is made between pilot test 1 and pilot test 2 to show the differences.

[2] On request, inspection in the observation notes is possible.

Table 5: Summary of the method sections.

SECTION		PILOT TEST 1	PILOT TEST 2
INSTRUMENT	**Phase in process**	Development stage	End stage
	Presentation value clarification decision tool	Paper	Digital online survey program, called Limesurvey
	Instrument	In-depth interview with topic questions (5 questions).	Parts of DCS/SWD integrated in value clarification decision tool. Two evaluation questions orally added.
	Duration	+/- 10 minutes (waiting time also +/- 10 minutes)	+/- 40 minutes (no waiting time)
PARTICIPANTS	**Inclusion criteria**	Men with problems with their urinary system, an age of 50 or older and Dutch speaking	Men with a diagnosis of localized prostate cancer (PSA values ≥ 20 Gleason score ≥ 7), aged 50 years or older and Dutch speaking
	Recruitment	Purpose sampling	Purpose sampling
	Minimum number required	5 participants	10 participants
PROCEDURE	**Performer**	Researcher (and specialist)	Researcher and specialist
	Task participant	Make four randomly selected choice tasks and think out loud.	Make all questions and evaluation questions after getting their advice. Think out loud again.
	Used method during pilot test	In- depth topic interview and the think- aloud technique	Think- aloud technique and observing
	Pilot Location	Treatment rooms of the urology department of the St. Elisabeth Hospital	Treatment rooms of the urology department of the St. Elisabeth Hospital
ANALYSIS	**Data Analyzed**	Observation notes and quotes were analyzed and coded	Observation notes and quotes were analyzed en coded. The data of the digital evaluation questions were analyzed in Excel

4. Results

4.1. Background of the pilot tests

For the first pilot test focusing on the design of the tool, seven Dutch men were selected. These men had an average age of 70.3 and had varying problems with their urinary system. This data is shown in Table 6.

Table 6: The included participants with their diagnosis (pilot test 1).

Participant	**Age**	**Disease**
♂ *Participant 1*	62	Puddle problems (LUTS)
♂ *Participant 2*	80	Radical Prostatectomy and bladder cancer.
♂ *Participant 3*	72	Kidney tumor in urinary system.
♂ *Participant 4*	71	Radical prostatectomy, recidive biochemical.
♂ *Participant 5*	71	Radical prostatectomy.
♂ *Participant 6*	59	Bladder cancer.
♂ *Participant 7*	77	Check for Prostate cancer.

In five of the seven cases, the participant was accompanied by his partner. The role of the partner was not addressed in this first pilot test, because the participants were emphatically asked to take an active role to generate as much recommendations as possible. In the second pilot test, focusing on the evaluation of the value clarification decision tool, the role of the participant as well as the role of his partner (when they came together to make the test) were analyzed in greater detail.

The second pilot test exists of ten Dutch men with an average age of 62.1. Two participants were waiting for a treatment and eight participants just started with a treatment. The treatment options were dependent of their personal characteristics and the size of the tumor indicated by their specialist. The characteristics and suitable treatment options of the selected participants are presented in Table 7.

Table 7: The included participants with their diagnosis (pilot test 2).

Participant	**Age**	**PSA value**	**Gleason score**	**Treatment option***				**Treatment status**
				1	***2***	***3***	***4***	
♂ *Participant 1*	62	Lower than 10	Lower than 7		X	X		Waiting
♂ *Participant 2*	58	Lower than 10	Lower than 7	X	X	X	X	Started
♂ *Participant 3*	60	Lower than 10	Lower than 7	X	X	X	X	Started
♂ *Participant 4*	67	Lower than 10	Lower than 7	X		X		Started
♂ *Participant 5*	75	Lower than 10	7	X	X	X	X	Started
♂ *Participant 6*	64	Lower than 10	Lower than 7	X	X	X	X	Waiting
♂ *Participant 7*	51	Lower than 10	7		X	X	X	Started
♂ *Participant 8*	66	Between 10 and 20	7		X	X		Started
♂ *Particpant 9*	58	Lower than 10	Lower than 7		X	X	X	Started
♂ *Paticipant 10*	60	Lower than 10	Lower than 7		X	X	X	Started

**1: Active Surveillance strategies*
2: Radical Prostatectomy
3: Internal radiotherapy (Brachy therapy)
4: External Radiotherapy (EBRT)

This result section is split in three paragraphs. The first paragraph (§5.2) focuses on the design of the tool, the second paragraph (§5.3) on the preference for a treatment, while the last paragraph will pay attention to the feasibility of the tool (§5.4).

4.2. Design of the value clarification decision tool

The generated experimental designs were tested in pilot test 1. This pilot test had a duration of 10 minutes and therefore the designs were divided over the participants (Table 8).

Table 8: The three experimental designs divided over the participants.

Design	Participant number *
First experimental design: *'AS or curative treatment'*	1, 4, 5, 7
Second experimental design: '*Radical prostatectomy or radiotherapy'*	2, 6
Third Experimental design: *'Internal (Brachy therapy) or external radiotherapy (EBRT).'*	3

** Most participants got the first experimental design due to the short test time.*

The first questions focused on the of BWS case 2. All participants did not experience difficulties with the task, but mainly with the content of the tool.

Participant 1: *"I could fill in the questions with statements directly. The content with all the disadvantages was confrontational for me."*
Participant 4: *"The questions were not confusing. I know that I have to choose between two worst scenarios."*

To specify the first reaction of the participants, attention was paid to the influence of the attributes and the levels in the statements on the difficulty of the content. This gave a remarkable result. All seven participants only paid attention to the attributes *(e.g. urinary incontinence)* and not to the levels presented in the statements *(e.g. 10 out of 100 patients)*.

Participant 1: *"When I read the statement, my attention goes to (urinary) incontinence and I actually do not read the number."*
Participant 6: *"I looked at (urinary) incontinence and the number faded."*
Participant 7: "*The number of patients has no addition. Even if you write down 'half of the patients', it does not get my attention."*

The participants were highly impressed by the negative attributes in the choice tasks which influenced the questionnaire. Two participants (20%) indicated that the formulation of the question was too neutral and did not match with their impression. There should be more emphasis on the unpleasantness of localized prostate cancer. Taking this into account, the question '*What is most and least annoying of a treatment?*' should be more suitable than *'What is the best and worst outcome of a treatment.'* Besides the formulation of the question, the participants also indicated that their tested experimental design existed of too many questions. These participants appreciate fewer questions.

Participant 5: *"When I have to fill in all questions, it will be too much. The questions look similar. They are all negative and that makes it confusing and tedious for me."*
Participant 3: *"I think that the questionnaire exists of too many heavy questions. Six questions would be much better."*

About the number of statements per profile, no participant had a comment. All reactions were positive.

Participant 2: *"Enough statements for every question. Not too much."*
Participant 6: *"I experience no problems with the number of statements."*
Participant 7: *"The statements are clear for me. You already noticed it because of my quick finish."*

Two other recommendations given at the end of the pilot test were: enough time to fill in the questionnaire and mentioning the effects of some attributes. These recommendations were integrated before pilot test 2 started.

First, the question was adapted to *'What is most and least annoying of a treatment?'* This is done for a better fit with the content of the statements. Second, new attributes and levels were created (*Appendix C: 'The attributes and levels redesigned'*) to diminish the number of questions and reformulate the statements. The statements were adapted to the given information in the tool. The consequences were built in and statements with levels *(e.g. 10 out of 100 patients)* were equalized with the information. The levels were not removed, because these levels made a distinction between the different treatments. This is important for the decision environment (*see paragraph 3.4.1; Phase 2: Identify attributes and levels).* An example of a customized choice task is shown in Figure 11.

Figure 11: Example of a redesigned choice task in the first experimental design of the value clarification decision tool; Active surveillance strategies (AS) - Curative Treatment (in Dutch).

The recommendation about the time to fill in the questionnaire will be automatically solved in practice, because patients could use the tool at home. For pilot test 2, there is no time limit implemented. The customized value clarification decision tool in Limesurvey is used in pilot test 2 to assess the preferences for a treatment and get an indication of the feasibility of the tool.

4.3. Preferences for a treatment

The preferences for a treatment and the corresponding preference-based advice were dependent of the possible treatments given by their specialist. Table 9 showed the preference-based advice with the strength of preference and the final treatment (almost) chosen through the participants of pilot test 2. No separation is made in the table between the prior preference and final treatment, because all participants indicated these as the same.

Table 9. An enumeration of their final treatment, the treatment advice, treatment score and strength of preference.

Participant	**Prior preference / Final Treatment**	**Advice *Value Clarification Decision Tool***	**Treatment Score**	**Strength of Preference**
Participant 1	Internal radiotherapy	Radical prostatectomy	Internal radiotherapy: -1 Radical prostatectomy: 3	Mild
Participant 2	AS	AS	AS: 4 Curative treatment: - 4	Strong
Participant 3	AS	AS	AS: 3 Curative treatment: -2	Moderate
Participant 4	AS	AS	Internal radiotherapy: 1 External radiotherapy: -2	Mild
Participant 5	Internal radiotherapy	Internal radiotherapy	AS: 1 Curative treatment: -1	Mild
Participant 6	Internal radiotherapy	AS	AS: 1 Curative treatment: -1	Mild
Participant 7	Radical prostatectomy	Radical prostatectomy	Radical prostatectomy: 7 Internal radiotherapy: -5 External radiotherapy: -4	Strong
Participant 8	Radical prostatectomy	Radical prostatectomy	Radical prostatectomy: 2 Internal radiotherapy: -5	Moderate
Participant 9	Radical prostatectomy	Radical prostatectomy	Radical prostatectomy: 5 Internal radiotherapy: -5 External radiotherapy: -3	Moderate
Participant 10	AS	AS	AS: 3 Curative treatment: -4	Moderate

Table 9 shows that five participants (50%) got an advice for active surveillance strategies, four (40%) an advice for radical prostatectomy and one participant (10%) got an advice for internal radiotherapy. None was advised to choose for external radiotherapy. These advices correspond in eight out of the ten cases (80%) with the participant's final chosen treatment. The treatment advice only differs for the two participants who were waiting to get their final treatment. These participants were asked for a possible explanation.

Participant 1: *"I think the cause lies in the degree of incontinence and impotence. This is not clear for me. Therefore, I find it difficult to compare the statements with each other. Another explanation could be the moment of the day. If I did heavy work with lifting and that kind of stuff, I would attach great importance to the (urinary) incontinence for example. If I just came out of bed, I would maybe attach more value to lower risk of being impotent. It is just what is on my mind in that moment."*

Participant 6: *"I understand very well why this is my advice. If I had to make a choice now, maybe I had chosen for active surveillance strategies. However, the shape of my prostate was strange and next to the preference for adverse effects as little as possible, I also want to treat the cancer. That is the reason that the specialist and I decided to choose for the radiotherapy. "*

Based on these data, nothing could be concluded about adherence to the generated treatment advice of the value clarification decision tool. The participants, who already chose a treatment, were engaged in the treatment process and therefore will adhere. Participant 1 did not dare to say anything about adherence and chose a neutral opinion. Participant 6 reported that he would not adhere to the generated advice of the value clarification decision tool, because later medical research has proven that this treatment is too risky for him.

The generalized treatment advice is based on the preferences of the participant. These preferences are distracted of the treatment scores and could be categorized in mild, moderate and strong. Table 8 shows that four participants have a mild, four a moderate and two a strong preference for their treatment. The explanation of participant 6 about his problem with choosing between active surveillance strategies (‘*few adverse effects as possible’*) and curative treatment (*‘treating the cancer’*), which is the cause of the wrong advice, seems logical. He has a very mild (neutral) advice for active surveillance strategies, according to the value clarification decision tool. So, the difference between mild and strong preferences depends of the score on the chosen statements related to the treatment options.

Most participants switched between three or more statements which were indicated as least or most annoying of the treatment (best/worst outcome). Seven participants (70%) considered the statements related to the lowest chance of adverse effects and treat the cancer as least annoying (best outcome). Half of the participants (50%) found it most annoying to undergo an unnecessary treatment which does not prolong their life (worst outcome). The citations confirm the result that the statements related to the attributes *‘the probability of adverse effects due to the treatment’, ‘the probability to undergo an unnecessary treatment’* and *‘the degree of anxiety experienced by the patient for a life with a tumor’* were the utmost important attributes for making their decision.

Participant 7: *“My main points about the preferred cancer is that the cancer must leave my body. I prefer that it leave my body with less as possible adverse effects.”*

Participant 5: *“Why would you choose an unnecessary treatment and get all kinds of adverse effects, while this is not needed and you also can choose for the option where the adverse effects can be avoided.”*

On the question whether the participants want to receive an advice based on their preference depends on earlier mentioned roles in the decision process (passive, cooperative or active). The statements of the Control Preferences Scale (CPS) were asked and the results are shown in Figure 12.

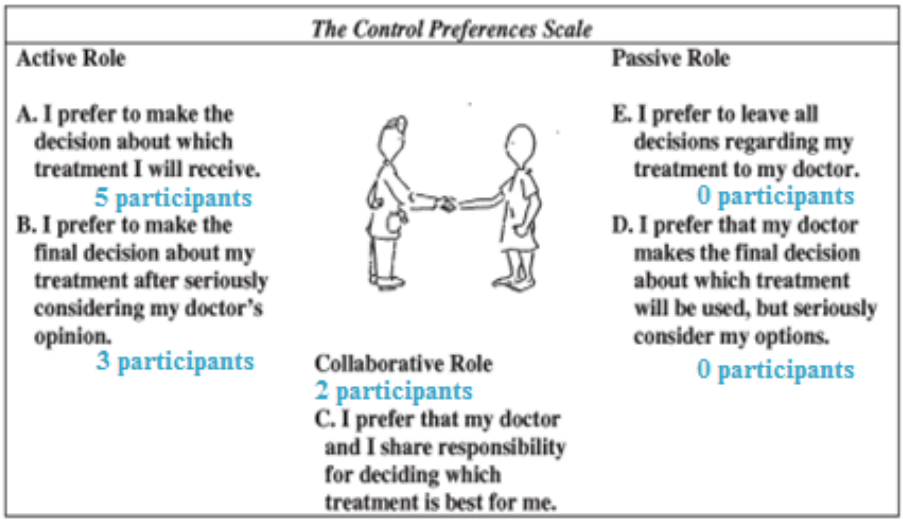

Figure 12. The CPS and the numbers of the participants, who indicated their most preferred role in their decision process.

Figure 12 shows that nobody had a preference for a passive role in the decision process. Most of them (80%) wanted an active role in which they make their own decision after information of the specialist/doctor. An active role is important for shared decision-making and therefore also for the use of the value clarification decision tool (27, 36). Besides the role of the participant in the decision process, the role of his partner by making the value clarification decision tool is also observed. The partner can influence the decision process consciously or unconsciously by steering her husband towards her own preferences. In four of the ten cases (40%) the partner was present. Three out of the four partners took a cooperative role; read the information and statements carefully, but only interrupted when she totally disagreed with her husband. For the main part she did not interfere. One of these four partners chose for an active role. This participant wanted to do everything in cooperation with his partner. Participant 10 came by himself, however he indicated that he included his family in the choice for a treatment. Probably his partner had an influencing role as well. This, however, could not be observed.

Participant 10: *"Every decision I make about my health is discussed with my family; my wife and children. The specialist has less influence on my decision."*

4.4. The feasibility of the value clarification decision tool

Before the feasibility of the value clarification decision tool is questioned, the need of the tool as support is analyzed as well in this study. Only three out of the ten participants (30%) were not able to make an informed choice between the treatments before they get the value clarification decision tool. Other participants were able to make a choice because they already had enough time to collect information on the internet, brochures or other peers.

Participant 1: *"Before I went on consultation for the diagnosis, I read all I could find about prostate cancer. However, in comparison with the information I found, the information in this tool is very clear to me."*

Participant 4: *"Due to the long preliminary phase, I could deepen my knowledge on the Internet. I also spoke my neighbor, who also had prostate cancer. The information given in the value clarification decision tool helped with the finishing touches. "*

The opinion of participant 1 (and his partner) was confirmed by other participants. The evaluation questions showed that the information about treatment options, advantages and disadvantages was complete and clear for most participants. Eight participants (80%) indicated that the received information of the value clarification decision tool was complete. Nine participants (90%) were aware of the different treatment options and knew all benefits of the different treatments after use of the value clarification decision tool. Eight participants (80%) knew all adverse effects. The participants, who were neutral on one of these aspects, missed some information, like the degree of the adverse effects, the function of the prostate in his body or the 'new' robot technique. This was the reason for participant 1 that he would not recommend the tool. He called the missing information an error. All other participants (90%) would recommend the tool to other patients.

Participant 1: *"I do not know if I would recommend the tool, because it is difficult and there are still errors in it. These things should be adjusted before I want to recommend it. I don't know if I recommend the value clarification decision tool afterwards. This depends on how it looks after the changes."*

The difficulty with the profiles tested in pilot test 1 popped up in this study. Three out of the ten participants (30%) indicated that it was difficult to start with making the profiles. This difficulty had two reasons. The first reason was identified by participants, the second was identified by the observers. The participants experience problems with comparing the different characteristics of the treatments.

Participant 3: *"It is difficult enough to choose between all those statements. What do I find the least and most annoying?"*

Participant 5: *"Life is not easy at all. These statements reflect this."*

The observers saw that most participants (80%) did not read the explanation about how the choice tasks works and started directly. After a few minutes, confusion was visible due to a lack of understanding. When the participants got some short explanation of the researcher or specialist, they made the choice tasks quickly.

Finally, the evaluation questions gathered information about the time the participant had spent with using the value clarification decision tool. All participants saw at least one experimental design dependent of the indicated treatment options given by their specialist. The average time spent on the total questionnaire was 30.52 minutes (median: 25.12). The average time per section and per question is shown in Table 10.

Table 10. Average time per section and question.

Section	**Average time per section (mins).**	**Average time per question (mins).**
Demographic background	6:26	1.39
Active surveillance strategies (AS) or Curative Treatment	13:28	2.21
Radical prostatectomy, Internal (Brachy therapy) or External (EBRT) Radiotherapy.	16:27	2:20
Operation (Radical prostatectomy) or External Radiotherapy (EBRT)	Not chosen	Not chosen
Evaluation questions	6:31	0:48

Table 10 indicates that there is almost no difference (1 sec) between the times spending per question in the two used designs. The time is evaluated as very acceptable through all participants.

On the basis of all these characteristics, the participants rated the tool. The average rating of the value clarification decision tool was 8.5. Some reactions are shown below.

Participant 1: *"A tool like this is convenient, because the treatment choice is hard to make. Every situation is different."*

Participant 4: *"I experienced the use of the value clarification decision tool as very positive. I liked to use it. I found it informative and I got a good advice."*

Participant 9: "*I'm very enthusiastic about this tool. What an improvement, I'm impressed. Only loving words, that's the reason why I gave it a note of a 10."*

All participants indicated that the use of the value clarification decision tool (with a personal advice) has an added value to the other received information. The participants were willing to use the value clarification decision tool regardless of the time and effort it takes and saw the value clarification decision tool as an improvement in comparison with the existing decision aid.

Participant 6: *"If I have to be very honest, the statements of the existing decision aid don't say anything for me. I find the questions too simple and too monochrome. When I said that I do not want anesthesia, the option operation immediately disappeared. This is surely not only dependent of a single characteristic, right? Hence, I'm happy with this new designed tool."*

Participant 7: *"I agree with the editing of the existing decision aid. The existing one was very limited. The questions were not detailed enough. I finished the existing decision aid in two minutes, while it was one of the most important decisions in my life at that moment."*

Besides these first positive reactions, all ten participants (100%) indicated the value clarification decision tool could assist them in making a decision after imagine that they did not make one yet. In the existing decision aid, 31 respondents (59%) indicated this. One reason might be the choice tasks assisted them with understanding what the most important aspects for making their decision were. In the existing decision aid only 38 of the 56 respondents (68%) understand the most important influencing aspects in their decision.

As with every pilot test, recommendations were made by the participants to improve the value clarification decision tool. The most recommendations were related to the formulation of the statements. Participant 6 preferred shorter and more concise statements. Participant 2 and 4 wanted percentages instead of numbers of patients, who suffered an adverse event. This made it more understandable for them. Participant 3 and 6 disagreed and prefer the numbers instead of percentages. According to participant 2 and 5, the word *'treatment'* was not correct for active surveillance strategies. In their perception active surveillance strategies is not a treatment, but a coping mechanism with regular checks. Participant 7 should change the words *'radical prostatectomy'* and *'radiotherapy'* in one word (for example *'treatment'),* to make the statements more neutral. Other recommendations were more attention for the new robot operation technique in the information and the opportunity to read all information, even when it not covers the patients' advice or advised treatment options through the specialist. This opportunity exists in the existing decision aid, but was difficult to build in Lime Survey. The features of Limesurvey were limited.

5. Conclusion

In this study is a new decision aid for (localized) prostate cancer is developed, the so-called value clarification decision tool. This tool assists patients in their decision-making process. The (central) research objective was to examine the added value of this preferences based value clarification decision tool and treatment advice in comparison with existing care, for men (50+) who have to choose a treatment for localized prostate cancer.

The results of the second pilot test (N=10) of this study pointed out that the value clarification decision tool is a valuable addition to existing care. It was shown that the tool enhanced shared decision-making by giving necessary information and ultimately providing a treatment advice. It improved patients' knowledge, showed patient preferences and stimulated patients to think about benefits and adverse effects of the different treatments. In comparison with the existing decision aid the value clarification decision tool has some extra features, like clarifying their preferences, preference calculations (mild, moderate, strong) and a personal preference-based advice, which were all positively evaluated. By using the value clarification decision tool patients got a better understanding of the important aspects of the decision-making process in comparison with the existing decision aid. The satisfaction rate of the value clarification decision tool is therefore higher than that of the existing decision aid (8.5 vs. 7.8). However, it should be noted that the results of the existing decision aid are based on experiments (N= 56; N=99) and the results of the value clarification decision tool have only been based on pilot tests (N=7; N=10).

Both tools showed that patients often want to play an active/collaborative role in their decision process, which is not strange given the current participation society these patients live in. So, the patient is ready for the new value clarification decision tool. Is health care ready too?

6. Discussion

'The patient is ready for the new value clarification decision tool. Is health care ready too?'

The value clarification decision tool was well received by patients. The average satisfaction rate of the value clarification decision tool was high (8.5). This high satisfaction rate can be explained by the rise of the collaborative/active role in decision-making and degree to which the tool takes patient preferences into account. Earlier research of Castell (1994) and Van der Rijen & Van der Kraan (2004) showed that patients want their treatment preferences to be honored (8, 9). Another conclusion of both sources was that patients frequently abandon a proposed treatment by the specialist, when the patient gets more time and space to think about the different treatment options on his own before making a decision. Through the use of the digital value clarification decision tool at home, both factors; honoring of their treatment preferences and more time and space to think, will be taken into account.

Another positive result was that participants were willing to use the tool no matter how much effort or time it takes. The time it takes to fully use the new value clarification decision tool was reduced by implementing the recommendations of pilot test 1. After this pilot test the value clarification tool was customized to a design in which the patient just gets one experimental design after the introduction instead of two or three. One exception is possible. This is when the patient is qualified for active surveillance strategies (AS) and two or three other treatment possibilities and his first advice is curative treatment. In this case he gets two experimental designs. First to decide if he prefers a curative treatment or active surveillance strategies (AS) and when he prefers a curative treatment, the second experimental design has to clarify which one exactly. In the second pilot test of this study, no participant ended up in this situation. The duration and the satisfaction of this exception have to be measured with another test.

The positive evaluation with a high satisfaction rate and high willingness to use the value clarification decision tool indicate that the patient is ready for implementing the value clarification decision tool. However, before the implementation of the value clarification decision tool could take place in practice, the last phases of the checklist by The International Society for Pharmacoeconomics and Outcomes Research (ISPOR) (Figure 7) should be carried out to generate more valid and veracious conclusions. Due to the few participants (seven and ten patients) in the pilot tests, a valid comparison with the data of the existing decision aid is not possible. Another problem with this few participants is the limited generalizability of the findings and opinions of the participants to the entire patient population with localized prostate cancer (low external validity) (101). The next phase (Phase 5: Execute an experiment) could show whether the value clarification decision tool is more effective and

enhance the statistical power. After positive results of the scientific experiment(s), the use of the value clarification decision tool could be extended to other diseases.

For designing a scientific experiment, special attention should be paid to the formulation of the statements, explanation of BWS choice tasks at the beginning, the manner of selection participants and the advice of the value clarification decision that not corresponding with their final treatment. The last recommendation focuses on changing the used software. These aspects will be discussed more extensively below.

The formulation of the statements.

The results showed different opinions about the formulation of the statements, especially about the length and the use of percentages. Some participants appreciate short and powerful statements with percentages, while others appreciate longer statements that also explain the effects and indicate the number of participants. No agreement is reached about the most valued one. Therefore, the formulation of the statements is participant dependent. The used statements in this value clarification decision tool were based on the first pilot test and the written information of the existing decision aid. Newer research indicated that not an X-in-100 format, but a 1-in-X format is the best format. This format yields higher perceived likelihoods and it appears to be the easiest format to interpret. (102) An experiment in which both formats were presented to the same participants, could clarify if the statements (and information) needs to be adapted.

Explanation of BWS choice tasks at the beginning.

Another point for improvement is the explanation of the use of the BWS choice tasks. Although the anxiety caused by the adverse effects in the choice tasks was decreased after reformulating the question and the statements, the difficulty with using the BWS choice tasks was increased. When the participants only could read how the technique works, no efforts were made. After receiving a short explanation of the researcher, this problem was solved. Therefore, a short oral explanation or an example question at the release is advisable.

Different manner of selecting participants.

The selection process caused two main problems: selection bias and self-fulfilling property. Selection bias is seen in the low difference between the treatment advices. The most participants (90%) got a treatment advice for active surveillance strategies (AS) or radical prostatectomy. This is probably due to selection bias. The patients were selected based on the diversity of treatments, to avoid this selection bias. However, the patients who have chosen the treatment options active surveillance strategies (AS) or radical prostatectomy were more willing to participate. Due to the fact that the specialist selected the patients on basis of the predefined criteria and the researcher, who was blinded

for their medical background, contacted them, this could happen. Another reason could be that these two treatments were the most recommended ones by specialists of the St. Elisabeth Hospital. Through the selection bias proper randomization is not achieved, thereby ensuring that the sample obtained is not representative of the population intended to be analyzed (103). Testing the value clarification decision tool with more patients, who chose different treatments, should reduce this problem.

A related problem with occurs through the use of these participants, who already chose a treatment, is the psychological mechanism self-fulfilling property. Self-fulfilling property is the prediction that directly or indirectly causes itself to become true, due to positive feedback between belief and behavior (104). This results in searching for conformation of the prediction (105). In this study, self-fulfilling property is established. The participants, who signed informed consent for continued research on the decision aid, have mostly already chosen a treatment. This led to the result that the chosen treatment always was indicated as their preferred treatment. This can, in turn, influence the test results and the validity and reliability of the advice, especially when the participant recognized the statements associated with his treatment. A solution could be to test the value clarification decision tool at the same way as the second pilot test with more patients, who are not already involved in a treatment but are waiting. This could also clarify the wrong treatment advices given to the waiting patients.

<u>No corresponding advices for patients, who were waiting for a treatment.</u>

A remarkable result of this study was that the only two patients, who were waiting for a treatment got another advice of the value clarification decision tool than the treatment they waiting for. All the patients who were starting with the final treatment got the same advice. The patients' explanations for the mismatch were the difficulty with judging the degree of the adverse effect/benefit and medical checks which indicate that during the moment of time that they had to make their decision, the advised treatment was not available anymore. A more methodological reason is that the implicit assumption used in this study may not hold. The implicit assumption was that the addition of deleted attributes to the choice tasks will not alter the results regarding attributes, which were important for the decision-making process. It is possible that there may be omitted attributes that, if presented, would let to different results. Also here is an experiment similar to the second pilot test recommended. More data of patients provides more insight and could clarify if this result really needs attention or just is caused by accident.

<u>Changing software.</u>

Finally, Limesurvey is not the most appropriate software to make a value clarification decision tool. Limesurvey is basically a static survey that is not meant for making choice-tasks and creating an advice. For calculating preferences and generating an advice the scores depend on all previous answers, which

is almost impossible in static surveys. The *'routing logic'* could be built in, but this generally go not beyond the cosmetic changes and asking whether or not a choice task based on the specialists advice has to be answered. Software to administer tests, e.g. computer- adaptive tests are more suitable. More research needs to be done to the most suitable one.

After implementation of these last recommendations, followed by experiments and the last phases of the checklist by The International Society for Pharmacoeconomics and Outcomes Research (ISPOR), the question about the readiness of existing Dutch health care for the value clarification decision tool remains. Dutch health care is diverse and complex (106). This makes it difficult to implement new innovations. However, health care is slowly changing to a system in which patient involvement in medical decision-making is encouraged (107). This are slight signs for changes directing to shared decision-making. The value of decision aids in this process needs to be proven, before the implementation will be considered. This requires not only more experiments that confirm the positive results, also enthusiastic teams, like the specialists of the St. Elisabeth Hospital, should communicate the changes and success stories about the implementation of decision aids. Decision aids can open a new area in existing care by empowering patients and creating new roles in the decision-making process between specialist and patients, which brings us a step closer to a more personalized and therefore a more satisfied answer to the question: '*Trick or treat?'*, one of the most difficult questions for patients with a diagnosis of localized prostate cancer.

Don't base your decisions on advice from people who don't have to deal with the results.

- Unknown

List of Tables

List of Figures

References

1. Integraal Kankercentrum (IKNL). Nederlandse Kankerregistratie. n.d. (cited 2013 January 1). Available from: http://www.iknl.nl/prostaatkanker&wordsMode=0
2. Oosterhof GON. Waarom wordt er geen bevolkingsonderzoek naar prostaatkanker georganiseerd? Vademecum permanente nascholing huisartsen. 2006: 3441-3442. ISBN: 978-90-313-8807-3. DOI: 10.1007/978-90-313-8808-0_1805.
3. National Institute for Public Health and the Environment (RIVM). Hoe vaak komt prostaatkanker voor, hoeveel mensen sterven eraan en neemt dit toe of af? Nationaal Kompas Volksgezondheid. 2014.
4. Bosetti C, Bertuccio P, Chatenoud L. Trends in mortality from urologic cancers in Europe. Eur Urol J. 2011; 60(1): 1-15.
5. Etzioni R, Legler JM, Feuer EJ, Merrill RM, Cronin KA, Hankey, BF. Cancer Surveillance Series: Interpreting Trends in Prostate Cancer—Part III: Quantifying the Link Between Population Prostate-Specific Antigen Testing and Recent Declines in Prostate Cancer Mortality. J Natl Cancer Inst. (JNCI). 1999; 91(12): 1033-1039. DOI: 10.1093/jnci/91.12.1033.
6. Neppl-Huber C, Zappa M, Coebergh JW, Rapiti E, Rachtan J, Holleczek B. Changes in incidence, survival and mortality of prostate cancer in Europe and the United States in the PSA era: additional diagnoses and avoided deaths. Ann Oncol. 2012; 23(5): 1325-1334.
7. Cutler DM. Where Are the Health Care Entrepreneurs? The Failure of Organizational Innovation in Health Care. Innov Policy Econ. 2011; 11(1): 1-28. ISSN: 0-226-47338-4.
8. Cassell EJ. The nature of suffering and the goals of medicine. Oxford University Press. 1994.
9. Van der Rijen AJG. & Van Der Kraan, WGM. Arts, patiënt en gebruik van de zorg. In: Met oog op gepaste zorg. RVZ. 2004: 217-252.
10. Du Benske G, Gustafon DH, Shaw BR, Cleary JF. Web-based cancer communication and decision-making systems: connecting patients, caregivers and clinicians for improved health outcomes. Med Decis Making. 2010; 30(6): 732-744. DOI: 10.1177/0272989X10386382.
11. Puts, MTE, Tapscott B, Fitch M, Howell D, Monette J, Wan-Chow-Wah D, et al. A systematic review of factors influencing older adults' decision to accept or decline cancer treatment. Elsevier. 2015; 41 (2): 197-215. DOI: http://dx.doi.org/10.1016/j.ctrv.2014.12.010
12. Yellen, SB, Cella DF, Leslie, WT. Age and clinical decision-making in oncology patients. J Natl Cancer Inst. 1994; 86(1): 1766–1770.
13. Blinman P, King M, Norman R, Viney R, Stockler MR. Preferences for cancer treatments: an overview of methods and applications in oncology. Ann Oncol. 2012; 1 (1). DOI: 10.1093/annonc/mdr559.

14. Brennan PF, Strombom I. Improving Health Care by Understanding Patient Preferences: The Role of Computer Technology. JAIMA. 1998; 5(3): 257-262.
15. Street RL, Elwyn G, Epstein RM. Patient preferences and healthcare outcomes: An ecological perspective. Expert Rev Pharmaecon Outcomes Res. 2012; 12(2): 167-180.
16. De Bekker-Grobb WE. Discrete Choice Experiments in Health Care. Theory and Applications. Erasmus University Rotterdam; Department of Public Health Erasmus MC. 2009. ISSN: 978-90-9024498-3.
17. Annunziata MA, Foladore S, Magri MD, Crivellari D, Feltrin A, Bidoli E, et al. Does the information level of cancer patients correlate with quality? A prospective study. Tumori. 1998; 84(6): 619-623.
18. Finlayson CS, Chen YT, Fu MR. The Impact of Patients' Awareness of Disease Status on Treatment Preferences and Quality of Life among Patients with Metastatic Cancer: A Systematic Review from 1997-2014. J Palliat Med. 2015; 18(2): 176-179. DOI: 10.1089/jpm.2014.0222.
19. Sommers BD, Beard CJ, Amico AV, Kaplan I, Richie JP, Zeckhauser RJ. Predictors of Patient Preferences and Treatment Choices for Localized Prostate Cancer. Cancer: 2008; 113 (8): 2058-2067.
20. Van Boven I, Zielman M, Bijsterbosch K. Gendermarketing. Customer insights. Christelijke Hogeschool Windesheim. School of Business & Economics. 2011: 3-5.
21. Davey HM., Barratt AL, Davey E, Butow PN, Redman S, Houssami N, et al. Medical tests: women's reported and preferred decision-making roles and preferences for information on benefits, side-effects and false results. Health Expect. 2002; 5(4): 330-340. DOI: 10.1046/j.1369-6513.2002.00194.
22. Wong F, Stewart DE, Dancey J, Meana M, McAndrews MP, Bunston T, et al. Men with prostate cancer: influence of psychological factors on informational needs and decision-making. J Psychosom Res. 2000; 49(1): 13-19. DOI: 10.1016/S0022-3999(99)00109-9.
23. Zhang AY, Siminoff LA. The role of the family in treatment decision-making by patients with cancer. Oncol Nurs Forum. 2003; 30(3): 1022-1028. DOI: 10.1188/03.ONF.1022-1028
24. Joy SM, Little E, Maruthur NM, Purnell TS, Bridges JFP. Patient Preferences for the Treatment of Case 2 Diabetes: A Scoping Review. Pharmacoecon. 2013: 31(10): 877-892.
25. National Health Services (NHS). Management Executive. Local voices; the views of local people in purchasing for health. London: NHS Management Executive. 1992.
26. Peep FM, Stalmeier PFM, Van Tol-Geerdink, JJ, Van Lin ENJ, Schimmel E, Huizenga H, et al. De patiënt kiest: werkbaar en effectief. Neds Tijdschr Geneeskunde. 2009; 153 (74): 1-7.
27. Atherton PJ, Ten Broeck-Smith MA, Singh JA, Huntington J, Diekmann BB, Huchka M. et al. The relation between cancer patient treatment decision-making roles and quality of life. Cancer. 2013; 119 (12): 2342-2347.

28. Bridges JF, Jones C. Patient-based health technology assessment: a vision of the future. Int J Technol. Assess Health Care. 2007: 23(1); 30-35.
29. Gagnon MP, Desmartis M, Lepage-Savary D, Gagnon J, St-Pierre M, Rhainds M. et al. Introducing patients and the public's perspectives to health technology assessment: A systematic review of international experiences. Int J Technol Assess Health Care. 2011; 27(1): 31-42. DOI: 10.1017/S0266462310001315.
30. Eddy DM. Anatomy of a decision. JAMA. 1990: 263(441).
31. Emanuel EJ, Emanuel LL. Four models of the physician-patient relationship. JAMA, 1992; 267(1): 2221-2222.
32. Ryan M. The Economic Theory of Agency in Health Care: Lessons from Non-Economists for Economists. Health Economics Research Unit Discussion Paper. 1992; 3(92).
33. Charles C, Gafni A, Whelan T. Decision-making in the physician-patient revisiting the shared treatment decision-making model. Soc Sci Med. 1999; 49(1): 651-661.
34. Agency for Healthcare Research and Quality (AHRQ). Achieving Patient-Centered Care with Shared Decision-making: A Brief for Administrators and Practice Leaders (Workshop Curriculum: Tool 9), Rockville. 2014: 4-5. Pubnr. 14-0034-9-EF.
35. Lerman CE, Brody DS, Caputo GC, Smith DG, Lazaro CG, Wolfson HG. Patients' perceived involvement in care scale: Relationship to attitudes about illness and medical care. J of Gen Intern Med. 1990; 5(1): 29-30.
36. Mahler HI, Kulik J. Preferences for health care involvement, perceived control and surgical recovery: A prospective study. J Health Econ. 1990; 12(1): 125.
37. Pierce PF. Deciding on breast cancer treatment: A description of decision behavior. Nursing Res. 1993; 42(22).
38. Legare F, Ratte S, Gravel K, Graham ID. Barriers and facilitators to implementing shared decision-making in clinical practice: update of a systematic review of health professionals perceptions. Patient Educ Couns, Elsevier. 2008; 73(3): 526-535. DOI: 10.1016/j.pec.2008.07.018
39. Wennberg JE. Unwarranted variations in healthcare delivery: implications for academic medical centers. BMJ. 2002; 325(7370): 961.
40. Krahn M, Naglie G. The Next Step in Guideline Development. Incorporating Patient Preferences. JAMA. 2008: 300(4); 436-438. DOI: 10.1001/jama.300.4.436.
41. Barnato, AE, Lleyellyn-Thomas HA, Peters EM, Siminoff L, Collins ED, Barry MJ. Communication and decision-making in cancer care: setting research priorities for decision support/patient' decision aids. Med Decis Making. 2007: 27(5); 626-634.
42. O'Connor AM, Bennett CL, Stacey D, Barry M, Col NF, Eden KB, et al. Decision aids for people facing health treatment or screening decisions. CDSR. 2003; 1(1). DOI: 10.1002/14651858.CD001431.pub2.

43. Stacey D, Samant R, Bennet C. Decision-making in oncology: a review of patient decision aids to support patient participation. CA. 2011; 58(5): 293-304.
44. Elwyn G, O'Connor AM, Stacey D, Volk R, Edwards A, Coulter A, et al. Developing a quality criteria framework for patient decision aids: online international Delphi consensus process. BMJ. 2006: 333(1). DOI: 10.1136/38926.629329.AE
45. Mosby JS. Medical Dictionary, 8th edition. Elsevier. 2009.
46. ZIGT. Van respondent naar participant. Hoofddorp; ZIGT Mediabureau. 2014.
47. Tversky A, Kahneman D. The framing of decision and the psychology of choice. Science. 1981; 221(4481): 453-458. DOI: 10.1126/science.7455683.
48. Diaby V, Goeree R. How to use multi-criteria decision analysis methods for reimbursement decision-making in healthcare: a step-by-step guide. Exp Rev Pharmacoecon Outcomes Res. 2014; 14(1); 81-99.
49. European Medicines Agency (EMA). Benefit-risk methodology project. n.d. (cited 2010 March 30). Available from: http://www.ema.europa.eu/ema/index.jsp?curl=pages/special_topics/document_listing/document_listing_000314.jsp.
50. Adamsen JM, Rundle-Thiele S, Whitty JA. Best-Worst scaling. Reflections on presentation, analysis, and lessons learnt from case 3 BWS experiments. J Market Soc Res. 2013; 21(1): 9-11.
51. Bridges JF. Stated preference methods in health care evaluation: an emerging methodological paradigm in health economics. App Health Econ Health Policy. 2003; 2(4): 213-224.
52. Vriens M. Het meten van preferentie structuren. Jaarboek MOA. 1997.
53. Edwards W, Fasolo B. Decision Technology. Annu Rev Psychol. 2001; 52(1): 518-606. DOI: 0066-4308/01/0201-0581$14.00.
54. Weernink MGM., Janus SIM., Van Til JA, Raisch DR., Van Manen, JG, IJzerman MJ. A Systematic Review to Identify the Use of Preference Elicitation Methods in Healthcare Decision-making. Pharmaceutical med. 2014; 28(3). ISSN: 1178-2595 DOI: 10.1007/s40290-014-0059-1.
55. Ungar WJ, Hadioonzadeh A, Najafzadeh M, Tsao NW, Dell S, Lynd LD. Quantifying preferences for asthma control in parents and adolescents using best-worst scaling. Respiratory med, 2014; 108: 842-851.
56. Carson RT, Louviere JJ. A common nomenclature for stated preference elicitation approaches. Environ Res Econ. 2010. DOI 10.1007/s10640-010-9450-x
57. Ryan M, Scott DA, Reeves C, Bate A, Van Teijlingen ER, Russel EM. Eliciting public preferences for healthcare: a systematic review of techniques. Int J Technol Assess Health Care. 2001; 5(5): 1-186.

58. Bridges JF, Hauber AB, Marshall D, Lloyd A, Prosser LA, Regier DA. Conjoint analysis application in health. A checklist: a report of the ISPOR Good Research Practices for Conjoint Analysis Task Force. Value in Health. 2011; 14(4): 403-413.
59. Bingham MF, Johnson FR, Miller D. From efficacy (to effectiveness) to market share: Modeling choice behavior for new pharmaceutical products. Value in Health. 2001; 4(1): 32-44.
60. Crawford MJ, Rutter D, Manley C, Weaver T, Bhui K, Fulop N. Systematic review of involving patients in planning development of health care. BMJ. 2002; 325(1): 1263-1265.
61. Bowling A. Research methods in health. Investigating health and health services. Buckingham: Open University Press: 1997.
62. Bryan S, Buxton M, Sheldon R, Grant A. Magnetic resonance imaging for the investigation of knee injuries: an investigation of preferences. Health Econ, 1998: 7(1); 595–603.
63. Ferguson RP, Wetle T, Dubitzky D, Winsemius D. Relative importance to elderly patients of effectiveness, adverse effects, convenience and cost of antihyperintensive medications. Drugs Aging, 1994; 4: 56–62.
64. Ratcliffe J, Buxton M. Patients' preferences regarding the process and outcomes of life. Int J Technol Assess Health Care. 1999; 15(1): 340–51.
65. Flynn TN, Louviere JJ, Peters TJ, Coast J. Best-worst scaling: What it can do for health care research and how to do it. J Health Econ. 2007; 26: 171-189.
66. Marley AAJ, Louviere JJ. Some probabilistic models of best, worst and best-worst choices. J Math Psychol. 2005; 49(1): 464-480.
67. Johnson FR, Lancsar E, Marshall D, Kilambi V, Muhlbacher A, Regier DA. Constructing experimental designs for discrete-choice experiments: report of the ISPOR Conjoint Analysis Experimental Design Good Research Practices Task Force. Value of Health. 2013; 16: 3-13.
68. Hensher DA, Rose JM, Greene WH. Applied choice analysis: A primer. Cambridge University Press, 2005.
69. Louviere JJ, Street D. Stated preference Methods. Handbook in Transport I: Transport Modelling. Pergamon (Elsevier Science); Amsterdam. 2000.
70. Marley AAJ, Flynn TN. Best Worst Scaling: Theory and Practice. International Encyclopedia of the Social and Behavioral Sciences. 2nd Edition. 2012; 2-15.
71. Bryan S, Dolan P. Discrete choice experiments in health economics. HEPAC. 2004; 5(3): 199-202.
72. Becker A, Berube D, Chad A, Dolovich M, Ducharme FM, D'Urzo T. Canadian pediatric asthma consensus guidelines. CMAJ. 2005: 173(1); 12-14.
73. Cope SF, Shade DM, Irvin CG, Mastronarde JG, Hanania NA, Castro M. International differences in asthma guidelines for children. Int Arch Allergy Immunol. 2009; 148(1): 275-278.

74. Erdmann, T.D. Statistical evaluation of binary measurement systems. The institutional repository of the University of Amsterdam (UVA). n.d. (cited 2011 Novembre 21). Available from: http://www.narcis.nl/publication/RecordID/oai%3Auva.nl%3A376607
75. Huber J, Zwerina K. The importance of utility balance in efficient choice designs. J Marketing Res. 1996; 28: 307-317.
76. Ryan M, Gerard K, Amaya-Amaya M. Using Discrete Choice Experiments to Value Health and Health Care. Springer. 2008; 11: 13-112. ISBN: 978-1-4020-4082-5.
77. Hulley SB. Designing Clinical Research. Lippincott: Williams & Wilkins. 2007: 168-169.
78. DeShazo JR, Fermo G. An exploration of the interval properties of three commonly used marketing research studies: a magnitude estimation approach. Journal of the Marketing Research Society. 2002: 29(3); 317- 399.
79. Lanscar E, Donaldson C. Discrete choice experiments in health economics. European Journal of Economics. 2005; 6(4): 314-316.
80. Boehmke MM, Dickerson SS. Symptom, symptom experiences, and symptom distress encountered by women with breast cancer undergoing existing treatment modalities. Cancer Nursing. 2005: 28(5); 382-389. DOI: 10.1097/00002820-200509000-0008.
81. Knobf MT, Sun Y. A longitudinal study of symptoms and self-care activities in women treated with primary radiotherapy for breast cancer. Cancer Nursing. 2005: 28(3); 210-218.
82. Larson M, Hedelin B, Athlin E. Needing a hand to hold: Lived experiences during the trajectory of care for patients with head and neck cancer treated with radiotherapy. Cancer Nursing. 2007; 30(4): 324-334. DOI: 10.1097/01.NCC.0000281722.56996.07.
83. Larson M, Hedelin B, Athlin E. Lived experiences of eating problems for patients with head and neck cancer during radiotherapy. JCN. 2003: 12(4); 562-570. DOI: 10.1046/j.13656-2702.2003.00751.x.
84. Svenson O, Maule AJ. Time pressure and stress in human judgment and decision-making. NY; Plenum Press. 1993; 293-308.
85. Center for Marketing Intelligence & Research (CMIR). Diepte-interviews. Right Marktonderzoek. 2015.
86. Ritchie J, Lewis J. Qualitative Research Practice. A Guide for Social Science Students and Researchers. In- depth Interviews. London; SACE Publications Ltd. 2003;138-145. ISBN: 076197116.
87. Giesen D, Meertens V, Vis-Visschers R, Beukenhorst D. Vragenlijst ontwikkeling. CBS. 2010: 5-54. ISBN: 18760333.
88. Stalmeier PFM, Roosmalen MS. Concise Evaluation of decision aids. Patient Educ Couns, Elsevier. 2009; 74(1): 104-109.

89. Chiavari L, Gandini S, Feroce I, Guerrieri-Gonzaga A, Russell-Edu W, Bonanni B, et al. Difficult choices for young patients with cancer: the supportive role of decisional counseling. Support Cancer Care. 2015. DOI: 10.1007/s00520-015-2726-5.
90. O'Connor AM. User Manual- Stage of Decision-making. Ottawa: Ottawa Hospital Research Institute. 1993.
91. O'Connor AM. Validation of a decisional conflict scale. Med Decis Making. 1995; 15(1): 25-30.
92. Redman BK. Measurement Tools in Patient Education. Springer Publishing Company. 2003. ISSN: 2002036570.
93. Holmes-Rovner M., Kroll J, Schmitt N, Rovner DR, Breer L, Rothert ML, et al. Patient satisfaction with health care decisions. The Satisfaction with Decision scale. Med Decis Making. 1996; 16(1): 58-64.
94. Boren MT, Ramey J. Thinking Aloud: Reconciling Theory and Practice. IEEE Transactions on professional communication. 2000; 43(3).
95. Plochg T, Juttmann, RE, Klazinga NS, Mackenbach, JP. Handboek gezondheidszorgonderzoek. Kwalitatief onderzoek. Bohn Stafleu van Loghum; Houten. 2007: 78-92. ISBN: 9789031348961.
96. Baker SE, Rosalind R. How many qualitative interviews is enough. NCRM. 2012.
97. Boeije, H. Analyseren in kwalitatief onderzoek. Denken en doen. De onderzoekspraktijk. Boom uitgevers Amsterdam. 2008: 10- 220. ISBN: 9085060788.
98. DiCicco-Bloom B, Crabtree BF. The qualitative research interview. Med Educ. 2006; 40(4): 314–321.
99. Gill P, Stewart K, Treasure E, Chadwich B. Methods of data collection in qualitative research: interviews and focus groups. BDJ. 2008: 204(6), 291–295.
100 Harris L, Brown G. Mixing interview and questionnaire methods: Practical problems in aligning data. Pract Assess, Res Eval. 2010; (15)1: 1-19.
101 Scholten RJPM, Tuut MK, Kremer LCM, Assendelft WJJ. Beoordelen van de Kwaliteit van Medisch wetenschappelijk Onderzoek, in Evidence-based richtlijnontwikkeling, Houten; Bohn Stafleu van Loghum. 2003: 158-171. DOI: 10.1007/978-90-313-9397-8_13, ISBN: 978-90-313-9397-8
102 Oudhof JP, Timmermans DR. The effect of different graphical and numerical likelihood formats on perception of the likelihood and choice. Med Decis Making. 2015: 35(4). DOI: 10.1177/0272989X15576487.
103 Heckman J, Hidehiko I, Smith J, Todd P. Characterizing selection bias using experimental data. Econometrica. 1999; 64(5): 1017-1098.
104 Merton, RK. The Self Fulfilling Prophecy, Antioch Review. 1948; 8(2): 195. DOI: 10.2307/4609267, ISSN 0003-5769.

105 Darley JM, Gross PH. A Hypothesis-Confirming Bias in Labelling Effects, in Stangor, Charles, Stereo cases and prejudice: essential readings, Psychol Press. 2000: 212-213, ISBN 978-0-86377-589-5.

106 Van Gerven M, Huisman J, Nijnens H. Duurzame gezondheidszorg. Visies en praktijkervaringen. Zeist; Christofoor. 2011: 199-200. ISBN 9789060386699.

107 Smith M, Higgs J, Ellis E. Factors influencing clinical decision-making in clinical reasoning in the health professions. Oxford: Butterworth-Heinemann. 2008: 89-100.

Appendix A: The initial attributes and levels

In this appendix, the different attributes and levels of the first three experimental designs are displayed. These designs were tested in pilot test 1. At the bottom of every experimental design, numbers are visible, which were made by Sawtooth software. These numbers corresponding with the oblique numbers in the tables and were used to make the BWS case 2 questionnaire.

Curative treatment or active surveillance strategies (AS).

Attributes	Nr. Levels	Active surveillance strategies (AS)	Curative treatment
The probability of (urinary) incontinence due to the treatment. ***(attr. 1)***	***2***	(Urinary) incontinence due to the treatment does not occur. ***(1)***	There is a probability (10 of the 100) of (urinary) incontinence. ***(2)***
The degree of anxiety experienced by the patient for a life with a tumor. ***(attr. 2)***	***2***	The tumor is not actively treated, but the development of the tumor is followed. The specialists intervenes when it's necessary and possible. ***(1)***	The tumor is actively treated. ***(2)***
The frequency content of the follow up. ***(attr. 3)***	***2***	Every three months, a checkup (with blood test, re-biopsies and rectal touché) takes place. After a couple days, you receive the results. ***(1)***	Every 6 to 12 months, a checkup (with blood test) takes place. After a couple days, you receive the results. ***(2)***
The probability of erectile dysfunction due to the treatment. ***(attr. 4)***	***2***	Erectile dysfunction due to the treatment does not occur. ***(1)***	There is a probability (50 of the 100) of erectile dysfunction. ***(2)***

Set	Attr.1	Attr.2	Attr.3	Attr.4
1	1	1	1	1
2	2	2	1	1
3	2	1	2	1
4	1	2	2	1
5	2	1	1	2
6	1	2	1	2
7	1	1	2	2
8	2	2	2	2

Radical prostatectomy or radiotherapy.

Attribute	Nr. Levels	Radical prostatectomy	Radiotherapy
The probability of (urinary) incontinence due to the treatment. ***(attr. 1)***	2	There is a probability (10 of the 100) of (urinary) incontinence. ***(1)***	There is a probability (1 of the 100) of (urinary) incontinence. ***(2)***
The probability of intestinal problems due to the treatment. ***(attr. 2)***	2	Intestinal problems due to the treatment does only occur in the first weeks after the prostatectomy. ***(1)***	There is a probability (15 of the 100) of intestinal problems. ***(2)***
The extent in which the patient finds it important that his prostate (body parts) are saved. ***(attr. 3)***	2	The tumor and prostate are both removed. ***(1)***	The tumor cells are treated, but the prostate is saved. ***(2)***
The duration of the rehabilitation after undergoing a treatment. ***(attr. 4)***	2	The treatment lasts several hours and the rehabilitation period is up to a month. ***(1)***	The treatment lasts several weeks and the rehabilitation period lasts weeks to months. ***(2)***
The probability of wearing a catheter after the treatment for (at least) two weeks. ***(attr. 5)***	2	A catheter is required for two weeks. ***(1)***	There is a probability (10 of the 100), that you have to wear a catheter for two weeks. ***(2)***
The degree of anxiety experienced by the patient for a surgery with general anesthesia. ***(attr. 6)***	2	The treatment is performed under total anesthesia. ***(1)***	The treatment can take place under total or local anesthesia. ***(2)***
The probability that after the first treatment another treatment is possible. ***(attr. 7)***	2	After the treatment, another treatment is possible, when the tumor returns. ***(1)***	After the treatment, another treatment has too high risks. ***(2)***

Set	Attr.1	Attr.2	Attr.3	Attr.4	Attr.5	Attr.6	Attr.7
1	1	1	1	1	1	1	1
2	2	1	2	1	2	1	2
3	2	2	1	1	2	2	1
4	1	2	2	1	1	2	2
5	2	2	2	2	1	1	1
6	1	2	1	2	2	1	2
7	1	1	2	2	2	2	1
8	2	1	1	2	1	2	2

Internal radiotherapy (Brachy therapy) vs. external radiotherapy (EBRT).

Description	Nr. Levels	Internal Radiotherapy	External Radiotherapy
The probability of intestinal problems due to the treatment. ***(attr. 1)***	*2*	There is a probability (7 of the 100) of intestinal problems. ***(1)***	Intestinal problems due to the treatment does only occur in the first weeks after the prostatectomy. ***(2)***
The duration of persistent fatigue caused by radiotherapy. ***(attr. 2)***	*2*	The duration of fatigue after the treatment goes up to 1 week. ***(1)***	The duration of fatigue after the treatment goes up to 7 weeks. ***(2)***
The probability that the patient temporarily has to avoid children and pregnant women due to the nuclear load. ***(attr. 3)***	*2*	Avoid temporarily contact with children and pregnant women is needed because of the nuclear load. ***(1)***	Avoid contact with others is not necessary. ***(2)***
The degree of perceived anxiety for the insertion of radioactive seeds and the consequences. ***(attr. 4)***	*2*	Internally radioactive 'seeds' are posted into your body. ***(1)***	No internally radioactive 'seeds' are posted into your body. ***(2)***
The probability of wearing a catheter after the treatment for (at least) two weeks. ***(attr. 5)***	*2*	The probability (10 of the 100) that you must wear a catheter for two weeks. ***(1)***	The probability (1 of the 100), that you must wear a catheter for two weeks. ***(2)***
The duration of the treatment in which the patient cannot do his daily activities. ***(attr. 6)***	*2*	The treatment is once and has a duration of 90 minutes with a rehabilitation period of one week. ***(1)***	The treatment has a duration of 10 minutes per day for 7 weeks. ***(2)***

Set	Attr.1	Attr.2	Attr.3	Attr.4	Attr.5	Attr.6
1	2	2	2	2	2	2
2	1	2	1	2	1	2
3	1	1	2	1	2	2
4	2	1	1	1	1	2
5	1	1	1	2	2	1
6	2	1	2	2	1	1
7	2	2	1	1	2	1
8	1	2	2	1	1	1

Appendix B: Interview questions of pilot test 1

In this appendix, the five open questions of pilot test 1 are shown.

The questions:

1. You made the first questions and statements of the value clarification decision tool. What is your first reaction?

2. What is your opinion about the questions?

3. What is your opinion about the statements?

4. What do you think of the number of statements per question?

5. Do you have general improvements and / or comments to improve this tool?

6. (Extra). If you look carefully to certain statements you see that some statement exist of two components: a problem (e.g. intestinal problems) and a number (e.g. 15 of the 100 patients).

a. What is your opinion about the combination of numbers (percentages) and the problem in the statement?

b. Which of these two components catch your attention in the first place?

These five (six) open questions were used in all six interviews.
The interviewer asked in more detail if this was possible.
The answers of the participants are available on request.

Same questions in Dutch:

1. U heeft de eerste vragen en stellingen gemaakt van de nieuwe keuzehulp. Wat was uw eerste reactive?

2. Wat is uw mening over de vragen?

3. Wat is uw mening over de stellingen?

4. Wat vindt u van het aantal statements per vraag?

5. Heeft u algemene verbeterpunten of opmerkingen voor de verbetering van de keuzehulp?

6. (Extra). Als u goed heeft gekeken naar sommige stellingen dan ziet u dat sommige stellingen uit twee componenten bestaan: een probleem (bijvoorbeeld darmproblemen) en een aantal (bijvoorbeeld 15 van de 100 patiënten).

a. Wat is uw mening over de combinatie van het aantal (percentages) en het probleem gecombineerd in één stelling?

b. Welk van deze twee componenten trekt als eerste uw aandacht?

Appendix C: The attributes and levels redesigned

In this appendix, the edited attributes and levels of the three experimental designs are displayed. At the bottom of every experimental design, numbers are visible, which were made by Sawtooth software. These numbers corresponding with the oblique numbers in the tables and were used to make the BWS case 2 questionnaire. The decision for these three experimental designs was made in collaboration with Urology specialists, who indicated that it is almost always possible to choose for Radical prostatectomy or External radiotherapy (EBRT).

Active surveillance strategies (AS) or curative treatment

Attributes	Nr. Levels	Active surveillance strategies (AS)	Curative Treatment
The degree of anxiety experienced by the patient for a life with a tumor.	2	The treatment of your prostate is deferred till the time it is absolutely necessary. ***(1)***	Your prostate will be directly treated. ***(2)***
The probability of adverse effects due to the treatment.	2	There are no adverse effects. Your physical function remains the same. ***(3)***	You are at risk of adverse effects and complications of the treatments, such as a risk of urinary problems, bowel problems and / or erectile dysfunction. ***(4)***
The probability of needing another treatment.	2	34 of the 100 men, who have been actively monitored in the first 5 years, require another treatment because of the rise in PSA score. ***(5)***	25 of the 100 men, who have in the first 5 years a treatment, still require a different treatment because of the rise in PSA score. ***(6)***
The probability to undergo an unnecessary treatment.	2	You are prevented for unnecessary treatments which do not extend your life. ***(7)***	You may undergo unnecessary treatment which does not prolong your life. ***(8)***
The degree of anxiety experienced by repeated PSA measurements and prostate biopsies.	2	The repeated PSA measurements and prostate biopsies can cause a lot of tension and insecurity. ***(9)***	There is no fixed schedule of tests that can produce tension and stress. ***(10)***

Version	Set	Item1	Item2	Item3	Item4	Item5
1	1	6	10	4	1	7
1	2	7	6	2	9	3
1	3	9	1	6	3	8
1	4	1	8	10	4	5
1	5	5	7	3	2	10
1	6	4	5	9	8	2

Radical prostatectomy, internal (Brachy therapy) or external radiotherapy (EBRT).

Attribute	Nr. Levels	Operation (Radical Prostatectomy)	Internal Radiotherapy (Brachy therapy)	External Radiotherapy (EBRT)
The degree of anxiety experienced by PSA score.	2	The prostate is removed in its entirety. You PSA score is 0. ***(1)***	After treatment, you keep a measurable PSA score. ***(2)***	After treatment, you keep a measureable PSA score. ***(1)***
The probability that after the first treatment another treatment is possible.	2	After an operation, you can still be irradiated upon return of the tumor. ***(3)***	After radiotherapy, you cannot be operated upon return of the tumor. ***(4)***	After radiotherapy, you cannot be operated upon return of the tumor. ***(4)***
The probability of (urinary) incontinence due to the treatment.	2	5 to 10 of the 100 men, who undergo surgery, have persistent incontinence. ***(5)***	1 to 2 in 100 men undergoing radiotherapy have persistent incontinence. ***(6)***	1 to 2 in 100 men undergoing radiotherapy have persistent incontinence. ***(6)***
The probability of intestinal problems due to the treatment.	2	Prolonged intestinal problems are rare. ***(7)***	0 to 15 of the 100 men will get permanent intestinal problems. **(8)**	Less than 1 of the 1000 mannen get after a long period permanent intestinal problems. ***(9)***
The size of the surgery.	3	You will need to undergo a major surgery with the risk of complications, such as bleeding and infections. ***(10)***	You must undergo a small surgery with a small risk of complications, such as bleeding and infections. ***(11)***	You do not need a surgery. ***(12)***
The probability of erectile dysfunction due to the treatment.	3	50 to 60 of the 100 men, who had surgery, get erectile dysfunction. ***(13)***	40 to 85 of the 100 men, who have erectile dysfunction, get irradiated. ***(14)***	40 to 52 of the 100 men, who have been irradiated, receive erectile dysfunction. ***(15)***

Version	Set	Item1	Item2	Item3	Item4	Item5
1	1	2	15	11	3	9
1	2	9	3	6	14	12
1	3	3	2	10	8	13
1	4	15	10	8	6	4
1	5	4	13	7	12	6
1	6	10	14	2	7	5
1	7	14	11	5	1	8
1	8	11	4	1	15	7
1	9	5	9	13	12	1

Radical prostatectomy or external radiotherapy (EBRT).

Attribute	Nr. Levels	Operation (Radical Prostatectomy)	External Radiotherapy (EBRT)
The degree of anxiety experienced by PSA score.	2	The prostate is removed in its entirety. You PSA score is 0. ***(1)***	After treatment, you keep a measureable PSA score. ***(2)***
The probability that after the first treatment another treatment is possible.	2	After an operation, you can still be irradiated upon return of the tumor. ***(3)***	After radiotherapy, you cannot be operated upon return of the tumor. ***(4)***
The probability of (urinary) incontinence due to the treatment.	2	5 to 10 of the 100 men, who undergo surgery, have persistent incontinence. ***(5)***	1 to 2 in 100 men undergoing radiotherapy have persistent incontinence. ***(6)***
The probability of intestinal problems due to the treatment.	2	Prolonged intestinal problems are rare. ***(7)***	Prolonged intestinal problems are rare. ***(8)***
The size of the surgery.	2	You will need to undergo a major surgery with the risk of complications, such as bleeding and infections. ***(9)***	You do not need a surgery. ***(10)***
The probability of erectile dysfunction due to the treatment.	2	50 to 60 of the 100 men, who had surgery, get erectile dysfunction. ***(11)***	40 to 52 of the 100 men, who have been irradiated, receive erectile dysfunction. ***(12)***

Version	Set	Item1	Item2	Item3	Item4	Item5	Item6
1	1	8	2	11	6	4	10
1	2	10	12	2	7	5	3
1	3	5	10	8	3	11	1
1	4	11	1	7	4	6	9
1	5	6	8	9	2	12	4
1	6	9	3	12	5	1	7

Appendix D: Calculation of the strength of preferences

The separation between mild, moderate and strong preference is calculated with the number of questions and the number of given responses. This is done with the attributes and levels of appendix C. These attributes and levels were used in pilot test 2. In this pilot test, scores of patients were generated.

For the version *'Active surveillance strategies (AS) or curative treatment'* and *'Radical prostatectomy or external radiotherapy (EBRT)'* the score is based on 6 questions with 12 responses (6 choice tasks with one as least and one as most annoying). When the difference between the scores of the treatments is 4 or lower, this will be indicated as a mild preference. A difference of the scores between 4 and 8 is called a moderate preferences and a score of 8 or higher is called a strong preference.

In the version '*Radical prostatectomy, internal (Brachy therapy) and external radiotherapy (EBRT)'* 9 questions were used with a total of 18 responses (9 choice tasks with one as least and one as most annoying). A mild preference is indicated when the difference between the scores is 6 or less, a moderate preference is given when the difference between the scores is between the 6 and 12 and a strong preference when the difference between the scores is 12 or higher.