FP7-287811

MobiGuide

Guiding Patients Anytime Everywhere

Collaborative projects - Large-scale integrating project (IP)

Start date: 1-Nov-2011 Duration: 48 months

Deliverable 4.2: Decision Models

Delivery due date: 31-10-2014

Actual submission date: 30-10-2014

Coordinator	HU
Deliverable Leading Partner	UNIPV
Contributing Partners	Silvana Quaglini (UNIPV), Silvia Panzarasa (UNIPV), Lucia Sacchi (UNIPV), Stefania Rubrichi (UNIPV), Carla Rognoni (UNIPV), Carlo Napolitano (FSM)
Revision	Version 0.1

Project co-funded by the European Commission within the Seventh Framework Programme (2007-2013)		
Dissemination Level		
PU	Public	**X**
PP	Restricted to other programme participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the consortium (including the Commission Services)	

Table of Contents

List of Figures

1. Executive Summary

This Document presents the implementation of the Decision Model framework described in Deliverable D2.1, which will be available within the MobiGuide prototype. In particular, Section 3 presents the Decision Trees (DTs) that will be delivered with the prototype. One of these models is described in more detail, as it has been developed during the third year of the project. The other DT has already been introduced in D2.1, but it underwent some modifications due to the release of an updated version of the guideline and the introduction of new oral anticoagulant drugs on the market. Such modifications are detailed in the present document. Moreover, we have implemented a visualization framework that allows inspecting the results of running a DT, to make the results' interpretation easier for patients and the results' explanation easier for the physician.

Section 4 is related to the description of the prototype version of the Utility Coefficients Elicitation (UCE) interface. We present the new release of this component, which has been implemented during the third year of the project. The new interface features are shown using a set of screenshots that follow the schema of a typical elicitation encounter. In Section 5 we present some concluding remarks.

2. Introduction

An important topic in the MobiGuide project is the one related to introducing shared-decision making (SDM) into a Clinical Practice Guideline (CPG)-based decision support system. The definition of such a framework has been one of the objectives of a number of the project's work packages, namely WP2, WP4 and WP5.

In this deliverable, we present the complete framework, highlighting also the most recent implementation advancements that have been proposed to make the SDM task effectively exploitable into the clinical practice together with the MobiGuide system. We have devoted specific attention to increasing the number of decision models available and to presenting the results in a more intuitive way, both to physicians and to patients.

3. Decision Tree Models

3.1. Decision trees for medical decision making

As introduced in D2.1, the decision models that were selected are Decision Trees (DTs). Decision models are built to address specific guidelines recommendations that suggest involving the patient in the choice of a specific treatment alternative. After a careful analysis of the guidelines related to the MobiGuide applications, two decision points of this type were found in the Atrial Fibrillation (AF) guideline:

- Selection of ablation therapy for treatment of AF
- Selection of treatment for preventing thromboembolism.

In both of these decision points, more than one legitimate alternative exists, and the patient can share the decision by specifying their preferences.

In the Gestational Diabetes Mellitus (GDM) guideline, no decision points with more than one possible alternative were identified by the medical experts. Thus, no shared decision making was modeled in that domain.

The model related to the selection of ablation therapy for treatment of AF has been integrated in the DT component. In the following, we briefly introduce the structure of the DT, which relies on a model presented in the literature [1] and which was further refined on the basis of domain knowledge, with the help of the FSM cardiologists.

Figures 1–3 summarize the structure, health states, and possible transitions between health states used in the model. A DT was combined with two Markov models, one for patients completing an ablation, and another for patients undergoing an Antiarrhythmic Drug (AAD) therapy. The DT starts with a decision node, which distinguishes the strategies in comparison (i.e., Ablation and AAD). After the initial decision, continuing on the ablation branch, patients may die as a result of the procedure, have non-fatal

complications, or have a normal course without complications. All patients who survive ablation enter the Markov process for the ablation (Figure 2).

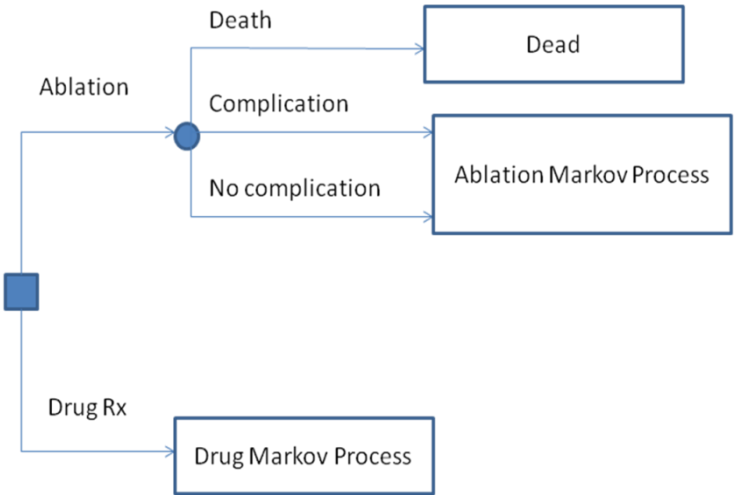


Figure 1- Simplified structure of the decision model. The square node indicates the decision point, while the round nodes represent the probabilistic events that may occur as a result of the choice.

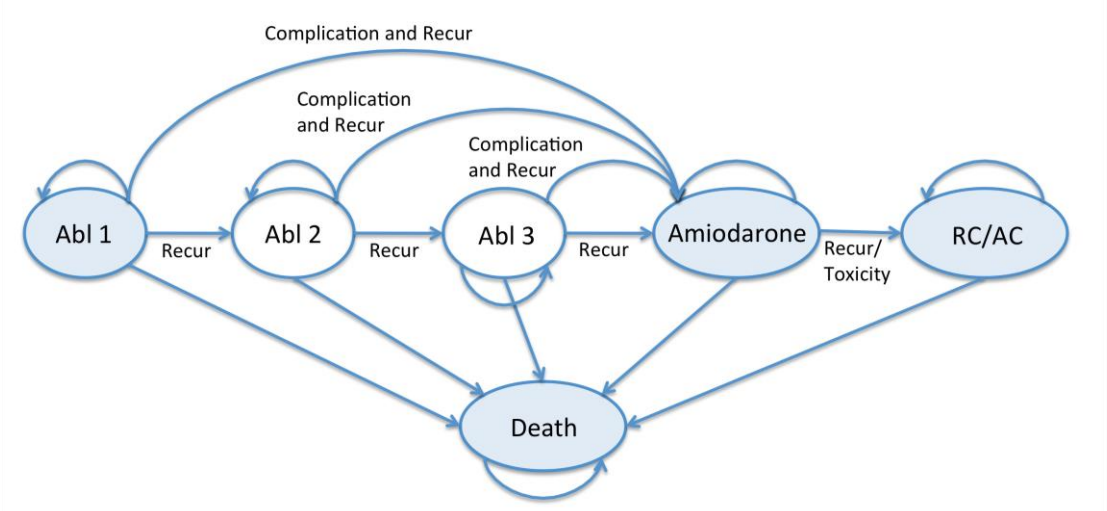


Figure 2 - Markov process for ablation. The diagram shows the main health states (boxes) and the possible transitions between them (arrows) in the ablation branch of the DT. In case a complication occurs after one of the ablations, the process follows the health states shaded in light blue. In each state the patient has a risk of death associated with its age.

Such a process assumes that patients will progress stepwise from one therapy to the next, based on whether or not they experience symptomatic AF recurrences on their current treatment. Since there is consensus that recurrences of AF after ablation procedures can be best controlled with re-ablation procedures, those patients showing a recurrence after a first ablation may repeat ablation up to two times. Following guideline recommendations, the process includes the treatment with the previously ineffective guideline-recommended first-line AAD drug, either sotalol or flecainide, for the first two months after ablation. Patients with recurrent AF despite the third ablation proceed to treatment by the Amiodarone AAD drug. Reynolds et al. assume that patients undergoing ablation will not be subject to drug treatment with Amiodarone. After consultation with our clinical partners, and taking into account what is indicated in the guidelines, we decided instead to consider Amiodarone at this stage. Moreover, we considered the possibility of encountering death due to the ablation procedure and its complications also after both the second and the third ablation. It was assumed that only patients who have undergone ablation without any complications could repeat the procedure. On the other hand, patients who have experienced non-fatal procedural complications incur costs and disutilities in the short term, and in the case of AF recurrence, proceed to treatment with Amiodarone.

Patients failing second-line drug treatment cease further efforts at rhythm control and are treated with pharmacologic rate control.

The AAD Markov process is shown in Figure 3.

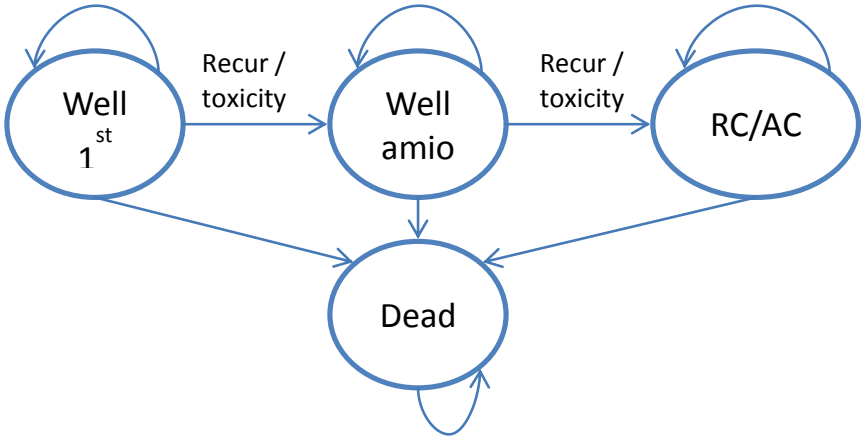


Figure 3 - Markov process for AAD. The three main treatment options in the group of AAD are: 1) sotalol / flecainide ('first drug'); 2) Amiodarone ('amio'); 3) medications for rate control and anticoagulation ('RC / AC').

Patients initially receive a first-line drug (sotalol or flecainide), entering the “Well 1st drug” state. In the event of toxicity or therapeutic failure, they start treatment with amiodarone (“well amio” state), and in the event of amiodarone failure, are treated with rate control (“RC/AC”). Amiodarone was chosen as the second-line agent for all patients in the drug “arm” based on its superiority over other drugs at maintaining sinus rhythm, however it is associated with more severe side effects.

For all patients should be taken into account the mortality rate related to age and sex, so each state can lead to death. Except for the very small risk of death associated to ablation and the fatal toxicity of the drug, the Reynolds model assumes that the risk of death is the same for all health states, except for stroke after ablation. We, however, do not consider stroke different from other complications as in the study by Reynolds et al., assuming that the incidence of stroke is the same for both therapies and we do not consider explicitly the stroke as an outcome. As reported in previous studies, after the success of ablation, we consider a three months therapy with anticoagulants and antiarrhythmic drugs.

The risk of toxicity related to antiarrhythmic drugs are obtained from the literature and are applied both to patients who undergo ablation and to those who follow pharmacological therapy only.

The DT related to the selection of treatment for preventing thromboembolism in low-risk AF patients has been described in D2.1. However, the model has been modified according to new evidence related to the use of dabigatran, a new anticoagulant (NOAC) drug recently introduced on the market. In particular, the model has been enhanced to add a new decision option, according to the following recommendation [2]:

Dabigatran is useful as an alternative to warfarin for the prevention of stroke and systemic thromboembolism in patients with paroxysmal to permanent AF and risk factors for stroke or systemic embolization who do not have a prosthetic heart valve or hemodynamically significant valve disease, severe renal failure (creatinine clearance 15 mL/ min) or advanced liver disease (impaired baseline clotting function) (Class I, Level of Evidence B[1]).

3.2. Visualization of DT Results

In deliverables D2.1 and D5.1, we have described the design and technical details of the web interface we have developed to present DT models to physicians during face to face encounters with their patients. This interface relies on TreeAge Pro, the most known commercial package used to implement DTs and TreeAge Pro Interactive, which allows to create web interfaces to make the models available to users. With this tool it is possible to visualize and edit model parameters, run the DT and show the results on terms of the selected outcomes to the patients. In the first version of the tool, such results were presented only in table format.

[1] Class I: Conditions for which there is evidence and/or general agreement that a given procedure/therapy is beneficial, useful, and effective. Level of Evidence B: Data derived from a single randomized trial, or nonrandomized studies (20).

Since graphical representations of data are usually easier to understand than written text or numeric tables, we added a method to graphically summarize the results of a DT. To implement the method, we exploited some advanced functionalities of TreeAge Pro. TreeAge results are further elaborated and combined to generate more comprehensive graphical outputs. To do that, we faced a dual challenge:

(i) to design the most appropriate format that would enable a patient to understand the health benefits and harms that are foreseen to influence his future not only as average/cumulative values, but also as time function

(ii) to provide information regarding results and inter-patients variability.

Our method is aimed at representing the expected evolution of a disease in a comprehensive but easily understandable format. Specifically, our idea is to display an average lifetime health path for each possible treatment choice. A lifetime health path consists of a representation of the expected transitions through different health states (e.g. from absence of disability to post-stroke disability), but it also includes punctual health events (e.g. a surgical intervention) over time within a lifelong temporal horizon. In the standard result presentation, a decision tree shows the expected values of the payoffs for each decision option; when a simulation analysis is performed, also a probability distribution of the outcomes may be shown, but in any case the results do not easily explain which are the health paths those values come from. To this end, we exploited the Markov modeling facilities provided by the package TreeAge Pro.

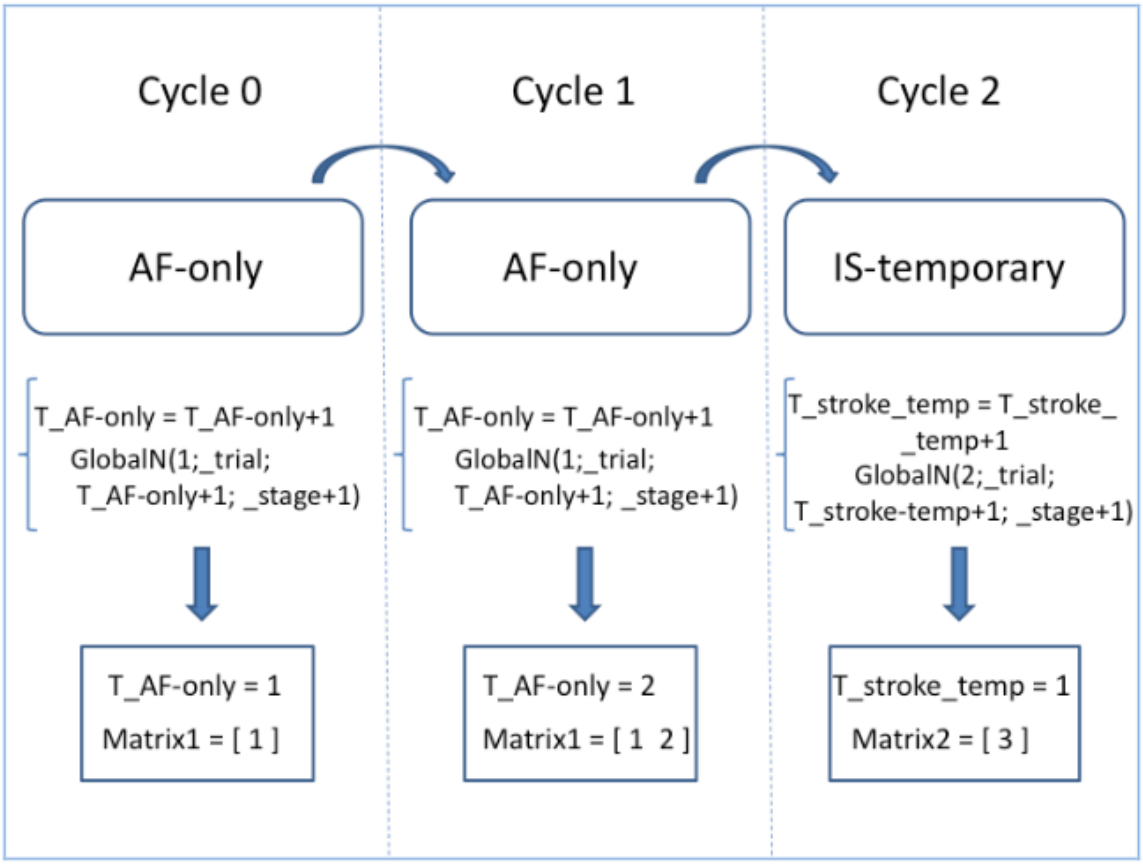
In the setting of our work, Monte Carlo microsimulation (MS) is a valuable way to capture the inherent variability of real world contexts as well as to efficiently track prior history. Unlike traditional cohort analysis, MS retains memory of previous events from one cycle to the next one, recording information about individuals' history through the Markov model [3,4]. Monte Carlo MS determines the pathway of a large number of individuals with identical initial characteristics. Individuals traverse the model one by one, and at each transition phase use random number sequences to select a single path through the model. Therefore, every individual will experience his own path. Running a number (n) of simulations (or trials) will result in a list of n randomly-chosen

outcome measures, one for each of the trials (e.g. n life years values, n lifelong costs, etc). Random sampling is performed according to the probability distributions in the model, and after a large number of trials, the central tendency can be observed. However, as already mentioned, we are interested in knowing more about the *distribution* of the final outcome. In TreeAge Pro, time-dependent models may be handled using some built-in variables (such as *_stage*, *_tunnel*, *tracker*), which allow including time reference in any expression of the model. Among these variables, the *trackers* represent a very useful and flexible means for keeping track of individuals time through the model. More precisely, by conveniently defining and updating a tracker variable at an event node and running a MS, it is possible to count the number of times an individual experiences that particular event as well as its duration. Then, while simulating individuals' paths, we can record the value of such variables and subsequently determine the life course/pathway of hypothetical individuals. These simulated life histories can be aggregated to estimate population-level effects on disease progression, but can also be represented graphically, in order to provide the patient with a more understandable result.

Tracking a patient's history within a Markov process requires to know the health events he experiences, their duration as well as the Markov cycle in which they occur. Hence, in addition to each tracker variable, we used a global matrix to store all the information. Global matrices are an advanced TreeAge Pro function that allows saving values globally and then using them for calculation or reporting purpose. The expression *GlobalN (n; row; col; value)* sets the cell *(row, col)* in the matrix *n* equal to the specific *value*. All the information we need can be simply retrieved by defining the tracker variable at the event node as follows.

The variables *_trial* and *_stage* are some built-in TreeAge Pro variables which count the MS iterations and the Markov process cycle, respectively. When running MS, if a simulation trial encounters a node with the tracker modification above, the current value of the tracker for that patient's trial is incremented by 1 and the cell (_trial, tracker, i.e. the event number) of the corresponding Global matrix is set to a value equal to the corresponding Markov cycle.

To illustrate how our framework works, we show an example on the DT related to



thromboembolism prevention (Figure 4).

Figure 4 - Using trackers to keep trace of every event in a patient's life path.

Our hypothetical patient begins the process in the "AF-only" state. We assume that the patient suffers a temporary ischemic stroke (IS) at stage two and so he moves to the state "temporary IS". At the end of cycle 2, the tracker variables that count the events "AF-only" and "temporary IS" will be then increased by 2 and 1 units, respectively and the associated global matrices will be set. At the start of each individual trial, all trackers are reset to 0. After the simulation is complete, the entire contents of the matrix can be dynamically saved to a text file or excel sheet for further elaborations.

The results of the MS analysis have been further elaborated to obtain a format suitable for supporting the patient's understanding of the benefits and risks associated to the different decision options. We propose several graphical formats, iteratively designed through active collaboration with clinicians. In the following, we will refer to the DT related to thromboembolism prevention in AF for presenting some examples of the proposed visualizations. It has to be pointed out that this visualization functionalities are available for all the implemented models. Figures 5 and 6 report results related to two of the decision options available for the considered DT, namely Dabigatran and No_Treatment.

Figure 5 depicts patients' lifetime paths using a stacked bar chart. Each one of the 100 bars displays the expected-life of a single patient. Each bar of the graph is divided into sections representing the different health states a patient goes through. Their position along the bar corresponds to the onset of the health state they represent. The length of each rectangle proportionally depicts the part of the patient's life spent in the corresponding state. Small triangles represent temporary events, such as temporary IS, intracranial hemorrage (ICH) and extra-cranial bleedings (ECB). The color of both triangles and bars relates to the severity of the condition (darker color indicating more severe condition). This graphical form provides a valuable tool for integrating several information. In particular, for each decision option, it allows conveying information about the survival trends as well as the course of disease in terms of succession of different health states.

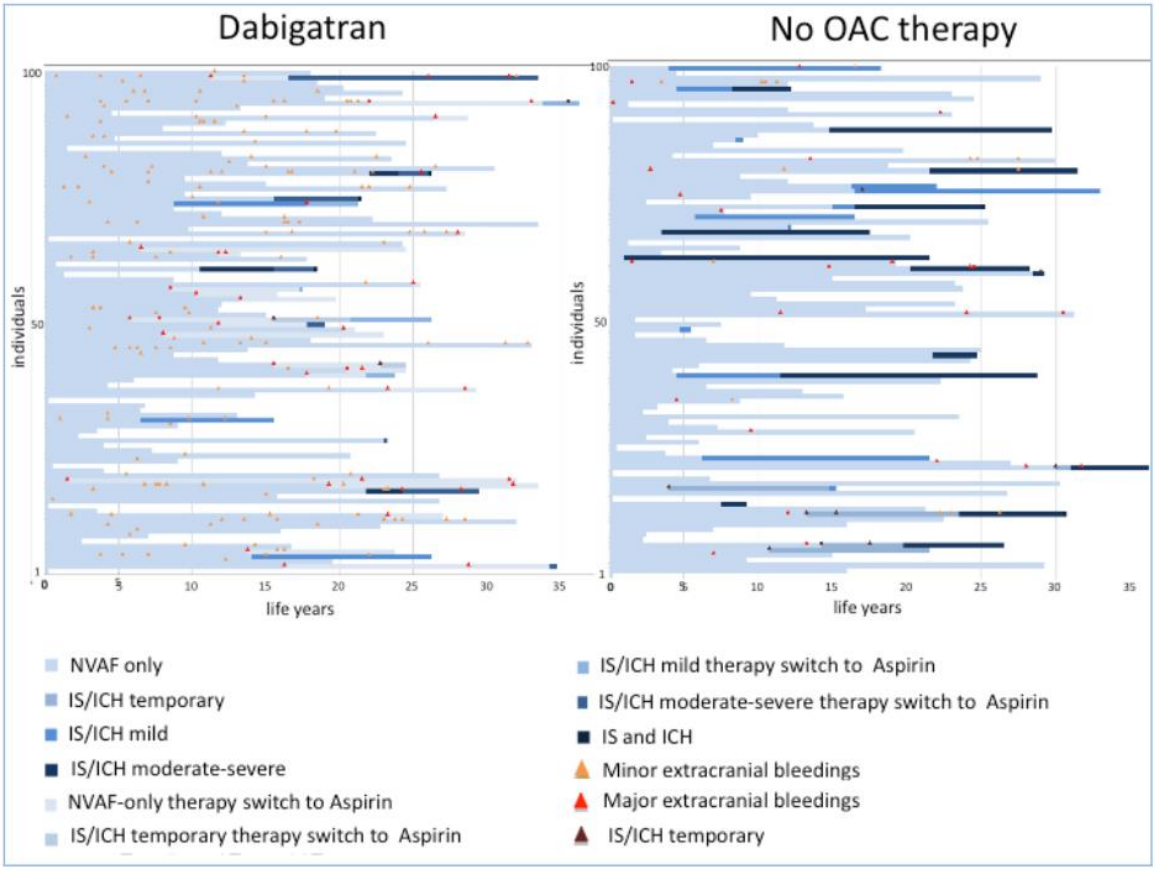


Figure 5 - One of the proposed graphical outputs. (left) Expected health-paths of a hundred non valvular atrial fibrillation (NVAF) patients treated with Dabigatran. (right) Expected health-path of a hundred patients not treated with OAC. Key domain-specific events in each patient's disease course are displayed along that patient's personal timeline. (IS: Ischemic Stroke, ICH: intracranial hemorrage, NVAF: non valvular atrial fibrillation)

Bars can be sorted according to different criteria. Sorting according to the expected QALYs makes the survival trend more explicit, with some hints on the quality of life. Sorting according to the utility coefficient of the worst health state that the patient has experienced helps to focus on the incidence of the worst health states, that is those states that result in different levels of impairment.

Figure 6 proposes an additional presentation format, which is based on a more compact display and allows summarization of data. We use a bubble chart to represent the

survival curve through xy location of the bubbles. The probability of having a temporary event (i.e. minor/major extracranial bleeding, IS/ICH temporary) is proportional to the bubbles size. The colored bar on the bottom represents the average QALYs over time (also in this case darker colors indicate worsen conditions).

Figure 6, besides showing higher 10-years survival for Dabigatran with respect to No OAC therapy (0.87 vs. 0.76), shows that the quality of life is higher for patients taking Dabigatran (lighter colors in the orange bar). This is due to a lower incidence of permanent disability events (see Figure 5). The drawback of Dabigatran therapy is the higher incidence of temporary events, which, given their short duration, don't impact on the overall quality of life. The higher incidence of temporary events is shown by the bigger bubbles size in the Dabigatran graph.

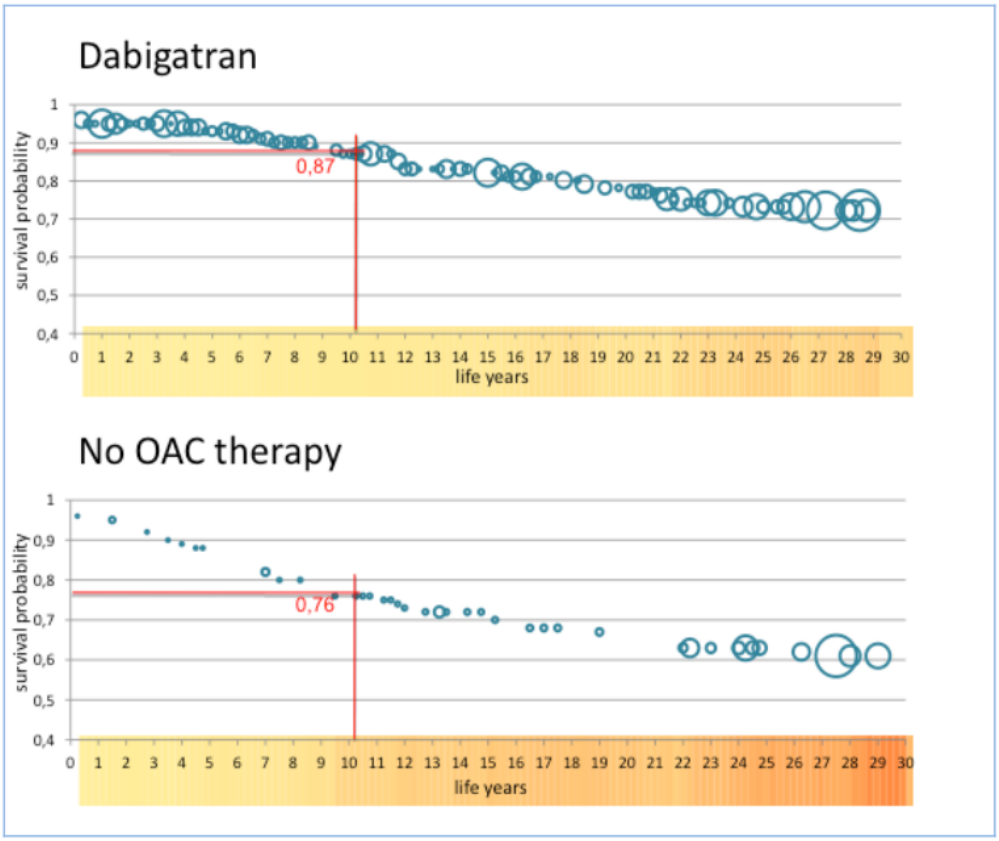


Figure 6 - Survival curves "weighted" for event probability, according to the treatment group (Dabigatran at the top, No OAC therapy at the bottom). Please note that we considered causes of death related to AF

4. Utility Coefficients Elicitation (UCE)

As described in deliverable D5.1, the UCE Interface is a web-based interface responsible for elicitation of utility coefficients (UCs) for DTs. It implements three different methods: Standard Gamble, Time Trade off and Rating Scale. It can be called from the caregiver interface when a recommendation suggests the possibility of taking a decision considering also patient preferences. Moreover, it can be accessed through the caregiver IF whenever the physician has the perception that a new elicitation is necessary for the patient (e.g. a new event has occurred). It stores utility coefficients that are then used by DT tool to run the decision trees.

With respect to the first prototype presented in D2.1 and D5.1, the interface has been completely re-designed to take into account the results of the preliminary evaluation (described in deliverable D5.2.2 and summarized in Appendix A of this document) and an analysis of previous research about utility coefficients elicitation. As the current version of the interface has been developed to be accessed from the caregiver interface also without being triggered by a specific recommendation, the starting page gives the possibility of selecting also the decision tree (besides health state and elicitation method), as shown in Figure 7.

Figure 7 – Home page of the UCE interface.

The available elicitation methods are the ones described in D2.1, namely: Rating Scale (RS), Time Trade-Off (TTO) and Standard Gamble (SG). According to the feedbacks collected during the pilot evaluation phase both by patients and physicians, we have modified and improved the GUI and the usability for the three methods. In all cases, we keep the media contents page separate from the elicitation page. In this way, the physician is able to easily skip the media page if needed (Figure 8).

Figure 8 - Multimedia resources for health states description.

The RS method is not used for QALY evaluation in the decision models but is very convenient to define a patient-based ranking of the available states. This is possible due to the simplicity of the method that allows patients to understand it quickly. We have improved the elicitation page by allowing the user to see a visual, instead of a numerical representation of values, by dynamically switching to a set of smiles. This further facilitates patient's understanding and makes administration of this method even easier for the physician. To make the elicitation process easier to understand, we have added legends to the numerical scale to point to the best and worst health state values. Moreover, we have improved the question description on top of the page, to support the physician in the explanation of the method during the encounter. The new page for the RS method is shown in Figure 9.

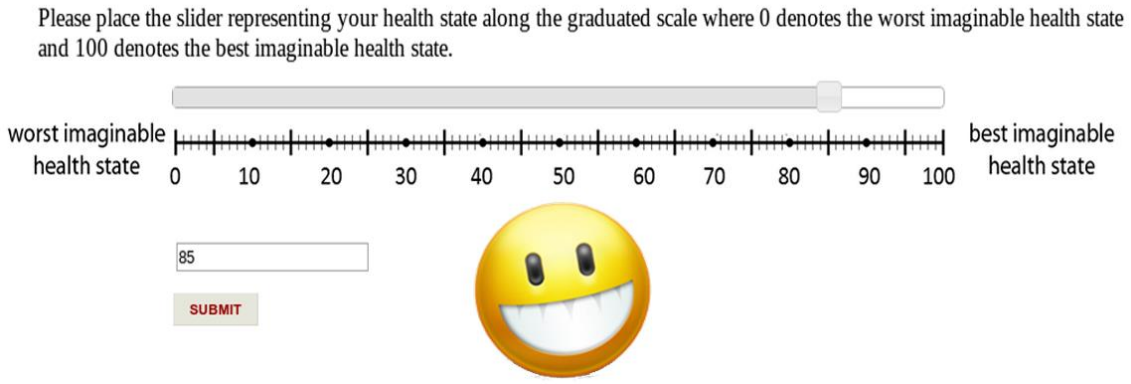


Figure 9 - The Rating Scale elicitation page.

As regards TTO and SG methods, some implementation improvements have been included in the present version of the interface. These have been described in details in D5.2.3. Summarizing, the major modifications have involved the following aspects:

- use of the values elicited with the RS method to initialize TTO and SG parameters
- implementation of a bisection algorithm that automatically chooses the next question, to optimize the tradeoff between number of questions and the accuracy of the elicited UC

The interface related to TTO has been restructured and is shown in Figure 10.

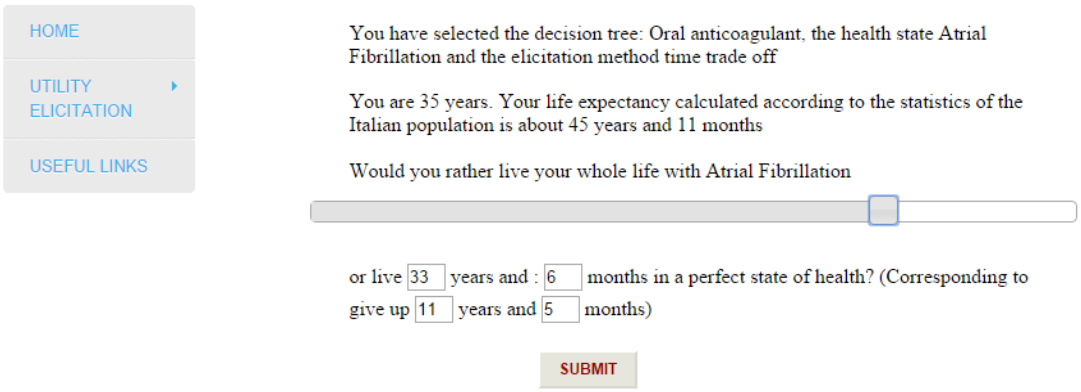


Figure 10 - The Time Trade-Off elicitation page

The SG elicitation phase has been divided into two steps. In the first, the patient is asked if he would be available to accept a risk of death (any) during a possible surgical intervention that might completely restore his health condition (Figure 11).

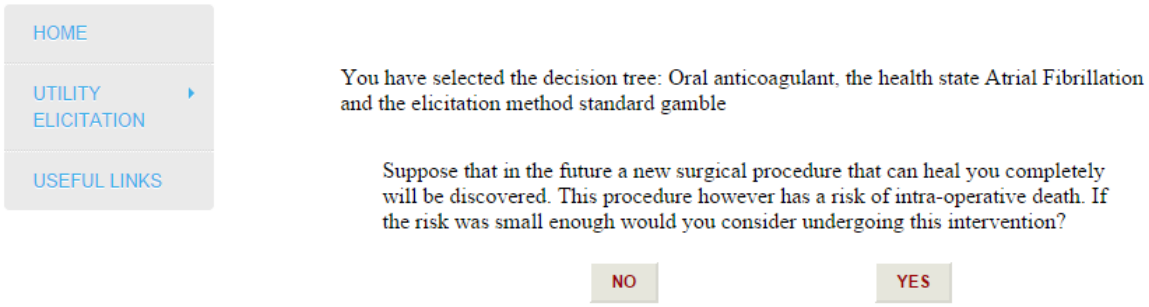


Figure 11 - The first page for eliciting UCs using the Standard Gamble method.

In case the patient declares he is not willing to consider such risk, the elicitation phase is stopped for this method. In case the answer is yes, the patient is directed to the SG page for elicitation. Here, according to the first patients' feedbacks, we have included the possibility deciding whether to use a graphical aid for risk assessment or not. In this representation, red icons reflect the portion of patients that would die according to the risk of death being presented to the patient. The graphical representation of the risk has two different visualizations: the random visualization (Figure 12), where red icons are randomly positioned on the panel, and the standard visualization (Figure 13), where red icons are positioned in a row.

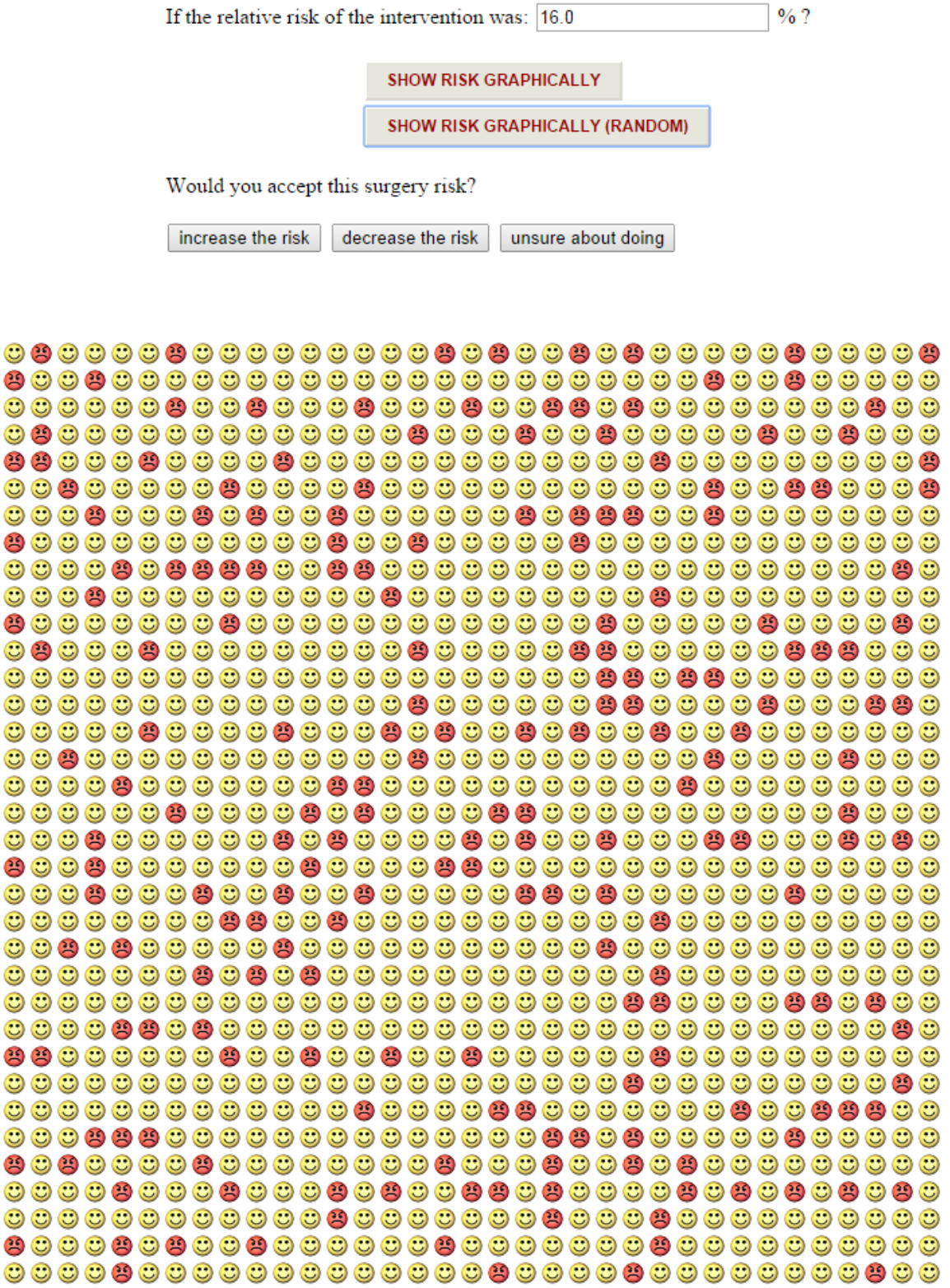


Figure 12 - Standard Gamble elicitation with the "Random" risk visualization option.

If the relative risk of the intervention was: 16.0 % ?

SHOW RISK GRAPHICALLY

SHOW RISK GRAPHICALLY (RANDOM)

Would you accept this surgery risk?

increase the risk | decrease the risk | unsure about doing

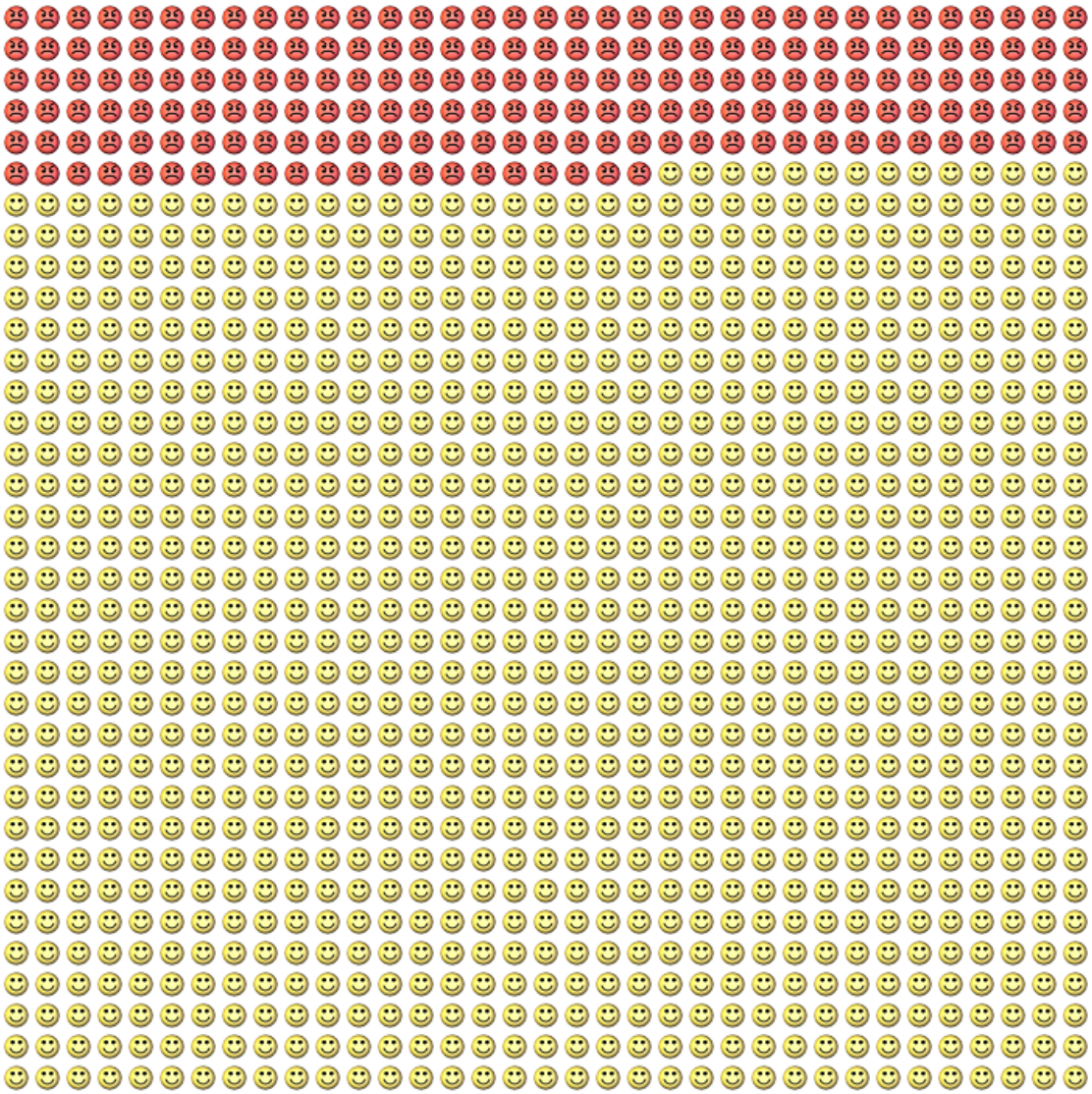


Figure 13 - Standard Gamble elicitation with the "Standard" risk visualization option.

Once UC_{TTO} and UC_{SG} have been elicited, the UCE interface also provides a functionality to calculate the average value. This value is able to potentially better summarize the QoL experienced by the patient. Moreover, a weighted mean, giving more importance to SG or to TTO, can also be calculated, according to the doctor's feeling about the patient's understanding of the two methods (motivating that feeling in a mandatory text field). This functionality is in line with findings reporting that some patients perceive the two methods with different difficulty and may misunderstand the questions or get confused with the required tasks [5]. Figure 14 shows the average computation for a patient on the Atrial Fibrillation health state. The average has been modified by the physician, who provided a motivation for this change.

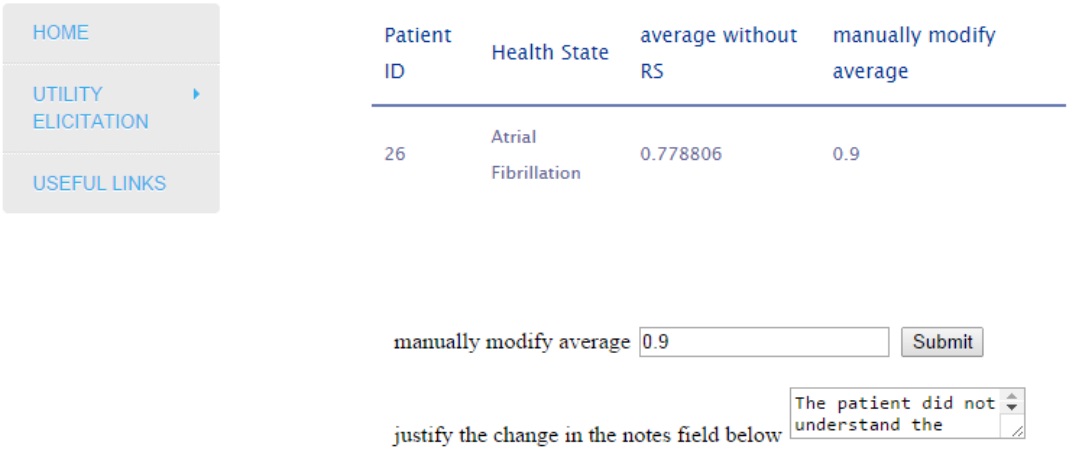


Figure 14 - Summary page showing the average elicited coefficient for the health state "Atrial Fibrillation".

5. Discussion and Conclusion

In this document, we have presented a framework that encapsulates decision models and instruments to elicit patients' preferences into a single tool, which in turn is integrated within the PHR GUI, thus enabling physicians to exploit electronic data management, evidence-based medicine and shared decision making in the same encounter. The availability of a single system that performs decision analysis and preferences elicitation enables the doctor to have all the needed instruments at hand. The use of these instruments offers several advantages with respect to simply carrying out an open dialogue with the patient. First, knowledge-based decision models depict in a systematic way all the possible consequences of decisions, thus leading to less subjective conclusions. Second, using the same model for all the patients allows adopting a homogeneous approach, helping the care professional to present the problem in the most complete way and giving similar information to all the patients. The system implements an automated direct triggering of the shared decision-making framework by CPG recommendations, thus customizing the guideline to the preferences of each patient. Having a unified framework to elicit patients' preferences and consequently run personalized decision trees is also a way to limit the amount of time needed to perform these kind of analyses during an encounter.

In the MobiGuide project, we have focused on the cases where a guideline explicitly states that patients' preferences might be considered in the decision process. Of course, taking into account the patients' opinion is important in any decision. The proposed framework is well suited to be adapted also to those situations where the guideline offers an unambiguous recommendation. As a matter of fact, any decision could in principle be represented as a decision tree that can be used to show to the patients not only the outcome related to the best choice, but also the one related to the alternative options, including the "do nothing" case. When strong evidence exists for a specific option, the decision tree can help the patient to better realise the appropriateness of the decision proposed by the guideline. Moreover, guidelines usually

consider only the health care related outcomes, while the use of decision trees makes it possible to take into account also other options of interest to the patient, such as out of pocket costs. This will be a future direction for our work.

6. References

[1] Reynolds MR, Zimetbaum P, Josephson ME, Ellis E, Danilov T, Cohen DJ. Cost-effectiveness of radiofrequency catheter ablation compared with antiarrhythmic drug therapy for paroxysmal atrial fibrillation. Circ Arrhythm Electrophysiol. 2009 Aug;2(4):362-9. doi: 10.1161/CIRCEP.108.837294

[2] Anderson JL, Halperin JL, Albert NM, Bozkurt B, Brindis RG, Curtis LH, DeMets D, Guyton RA, Hochman JS, Kovacs RJ, Ohman EM, Pressler SJ, Sellke FW, Shen WK, Wann LS, Curtis AB, Ellenbogen KA, Estes NA 3rd, Ezekowitz MD, Jackman WM, January CT, Lowe JE, Page RL, Slotwiner DJ, Stevenson WG, Tracy CM, Fuster V, Rydén LE, Cannom DS, Crijns HJ, Curtis AB, Ellenbogen KA, Le Heuzey JY, Kay GN, Olsson SB, Prystowsky EN, Tamargo JL, Wann S. Management of patients with atrial fibrillation (compilation of 2006 ACCF/AHA/ESC and 2011 ACCF/AHA/HRS recommendations): a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines. J Am Coll Cardiol. 2013 May 7;61(18):1935-44. doi: 10.1016/j.jacc.2013.02.001.

[3] Gray AM, Clarke PM, Wolstenholme JL, Wordsworth S. Applied methods of cost-effectiveness analysis in health care. Oxford University Press. 2011.

[4] Hiligsmann M, Ethgen O, Bruyère O, Richy F, Gathon HJ, Reginster JY. Development and validation of a Markov microsimulation model for the economic evaluation of treatments in osteoporosis. Value Health. 2009;12(5):687-96.

[5] M. Ryan, V. Watson, M. Amaya-Amaya. Methodological issues in the monetary valuation of benefits in healthcare. Expert Reviews in Pharmacoeconomic Outcomes Research (2003); 3(6), 717-727

7. Appendix A

Summary statistics (mean ± standard deviation) of the scores for the two questionnaires (AFEQT and EuroQol) and for the three direct utility elicitation methods (SG, TTO and RS) are reported in Table A.1. In addition to the overall score, the four AFEQT subscores are reported. Concerning the direct methods, our results are in agreement with the literature, reporting that RS usually provides lower and more variable values [Morimoto T, Fukui T. Utilities measured by rating scale, time trade-off, and standard gamble: review and reference for health care professionals. J Epidemiol (2002) Mar;12(2):160-78].

Table A.1- Mean and standard deviation (SD) of the scores obtained with the questionnaires and with the three direct elicitation methods

	AFEQT					EuroQol	U_{TTO}	U_{SG}	U_{RS}
	Overall score	Symptoms	Daily activities	Treatment concern	Treatment satisfaction				
Mean ± SD	67.71 ±19.02	76.09 ±20.31	66.67 ±25.17	63.58 ±22.44	67.13 ±20.10	0.586 ±0.369	0.979 ±0.058	0.977 ±0.219	0.669 ±0.196

Since the questionnaires we selected are validated tools to assess QoL, we expected to find correlations among their scores and the UCs elicited with the three implemented direct methods. To investigate this issue, we computed all the correlation coefficients (using the Pearson's method) between the scores obtained with the questionnaires and the scores obtained using TTO, SG and RS. The full correlation matrix summarizing this analysis is provided Table A.2.

Table A.2 - Correlation coefficients (p-values) between the quality of life values elicited with different methods

	AFEQT	EuroQOL	SG	TTO	RS
EuroQOL	0.32 (ns)				
SG	Overall:0.32(ns) Symptoms:0.49 (0.04)	0.58(0.02)			

	Daily activity: 0.43 (0.07)				
TTO	0.56(0.015)	-0.18(ns)	0.26(ns)		
RS	0.24 (ns)	0.35(ns)	-0.04(ns)	-0.07(ns)	
Avg(TTO,SG)	Overall: 0.56 (0.02) Symptoms: 0.67 (0.002) Daily activity: 0.63 (0.005)	-0.1(ns)	-	-	-0.14 (ns)

We herein report a summary of the most relevant results. We obtained significant correlations between the overall AFEQT score and the TTO method and between the AFEQT symptoms subsection score and the SG method. From this observation it emerges that the values elicited with the different methods are in general not well-correlated. This can be due to different reasons. First of all, since, as mentioned, the direct methods may be difficult to understand, some of the values provided by patients could be not reliable. Second, the two questionnaires are quite different, being the EuroQol very general, and the AFEQT specific for AF. An interesting result regards the AFEQT symptom and daily activities subscores. In this case, the most significant correlation coefficient was not found with one of the elicitation methods, but with the average value obtained from TTO and SG. In particular, for the symptoms subscore, we found a correlation coefficient of 0.67 ($p<0.002$) (see supplementary material, Suppl. Figure 2), whereas for the daily activities we found a correlation coefficient of 0.63 ($p=0.005$). These results suggest that SG and TTO probably capture different aspects of the QoL, and must be jointly considered to have a global picture of the patient's perspective. This is an example of how the collected data can be used to gain further insight into theoretical models of patients' preferences.

Another point we have evaluated is the time needed to administer the different methods. This data is relevant because visit duration is always a concern for physicians. As a consequence, it is important to inform them about how longer visits would be with the introduction of a new task. Average minutes necessary to administer the tools were: 11.7 for AFEQT (range 5-20), 3.2 for EuroQol (range 2-5), 1.5 for RS (range 1-2), 3.2 for TTO (range 2-5) and 5.9 for SG (range 2-15).