TreatmentExplorer: an Interactive Decision Aid for Medical Risk Communication and Treatment Exploration

Lyndsey Franklin
HCIL, University of Maryland
College Park, MD
lyndsey@cs.umd.edu

Kazi Minhaszur Rahman
University of Maryland
College Park, MD
minhasz@umd.edu

Catherine Plaisant
HCIL, University of Maryland
College Park, MD
plaisant@cs.umd.edu

Ben Shneiderman
HCIL, University of Maryland
College Park, MD
ben@cs.umd.edu

ABSTRACT
Medical treatments carry unique benefits and risks which patients must understand in order to decide which option is best for them. Prior research has demonstrated that patients are ill-equipped to understand the statistical information presented to them through standard decision aids. We describe a prototype decision aid, TreatmentExplorer, which supports patients’ needs by presenting treatment success rate, onset of symptoms, and treatment side effects using a novel graphic representation introduced with staged animation and text-based narration. We report on a pilot study (n=24) and a main user study (n=42) which characterize the benefits of TreatmentExplorer over a text only decision aid as well as a version without staged animation. We conclude with guidelines for designers.

Author Keywords
Risk Communication; Animations; Guided Narration; Evaluation; Interactive Interfaces; Information Visualization

ACM Classification Keywords
H.5.1. Information interfaces and presentation (e.g., HCI): Multimedia Information Systems. H.5.2. Information interfaces and presentation (e.g., HCI): User Interfaces.

INTRODUCTION
Consider the following scenario: Donna is 60 years old and has a family history of breast cancer. She also recently tested positive for the BRCA gene mutations so she is at high risk of developing cancer herself. She now has to decide whether or not she wants to opt for preventative treatments like surgery or medication to reduce her risk.

Evidence-based medicine [1] promotes the ideal that patients like Donna should be given personalized data about their treatment options so that they can come to a logical decision based on treatments that have the greatest efficacy – usually with their physicians. They must become knowledgeable - sometimes in a short period of time - about their condition and treatment options, and for each treatment option they need to understand the expected benefits (e.g. a reduction of the probability of getting cancer) and risks (e.g. side effects or possible secondary illnesses brought by the treatment). All those benefit and risk measures also have a temporal element: when is this likely to happen? How long will this side effect last?

Unfortunately, patients are often unable to use the information presented to them about treatments and make poor choices [2] [3]. Many patients are ill-equipped to understand both medical terminology and the statistical data presented to them in decision aids [2]. Even educated patients have difficulties reading and understanding textual presentations of treatment information [2]. Misinterpretation of both numeric and textual information is a frequent hazard [4]. Patients also have their own opinions on their medical care in terms of which lifestyle-impacting side effects they are willing to cope with as part of their treatment [5]. It has been reported that even with the data to reach a logical decision, patients often base their treatment decisions on their emotional reactions rather than rational decision-making [6] [7]. Thus, the design of a useful decision aid faces challenges of risk communication, health literacy, and placement within a medical workflow.

These medical challenges are a new domain for HCI researchers who are in the position to apply existing usability guidelines and heuristics to the problem of supporting patients in their decision making. In this work, we describe a prototype decision aid, TreatmentExplorer (Figure 1), which provides patients with a novel representation of temporal treatment information, presented...
in an interactive interface using staged animation and text-based narration. We report on the evaluation of the prototype with four domain experts and the results of a pilot study (n=24) and a main user study (n=42). Finally, we present a set of guidelines based on the results of our design and evaluation experiences.

RELATED WORK
TreatmentExplorer draws its inspiration from the work of risk communication, health literacy, and medical decision aid research.

Risk Communication: Research investigating the communication of risk in healthcare has consistently demonstrated one great challenge: patients are generally poor with statistics, even when educated. In their commentary, Fagerlin, Zikmund-Fisher, and Ubel provide 10 specific recommendations for improving risk communication which include recommendations on use of language, text, graphics, order of information, use of comparison, and presentation of time [8]. They note that frequencies (i.e. 1 of 4) are preferred for providing information about absolute risks and/or highlighting changes between levels of risk. They also recommend repeatedly drawing attention to the time interval over which a risk occurs and the inclusion of graphs and summary tables.

Han et al. have examined the topic of uncertainty in risk communication and formed a taxonomy of types of uncertainty within health care in the attempt to help clarify the problem of its expression [9]. Follow up work endeavored to produce novel visualizations capable of representing randomness and its effect on uncertainty [10].

Health Literacy: Health literacy research focuses on producing health-related messages that non-experts can understand and use in making informed treatment choices. The goal is to be able to produce health materials that patients with deficient skills are able to access as well as patients without such deficiencies. For comparing high-level information such as the quality of treatment-supporting evidence or a summarized rating of a treatment’s success, work has shown that icons such as star ratings or symbols are preferred over other representations [11] [12]. Other research suggests that the presence of graphics increases the believability of information [13]. Verbal expressions such as “low”, “medium”, and “high” have been shown to lead to misunderstandings between physicians and patients and interfere with patient understanding [4].

Medical Decision Aids: Decision aids are often evaluated as part of larger clinic systems where the focus is on the feasibility of their deployment and their role in patient-physician communication. Decision aids used for physician-patient communication often produce anxiety, which interferes with patients’ ability to reason about their healthcare and that this impacts their decisions [14]. However, patients have also credited decision aids with more productive and efficient physician consultations [15]. Visual representations provide memory prompts which reduce the cognitive load of patients during decision-making [14]. Study results suggest that decision-aided patients are able to express more emotions, use more cognitive terms, and verbalize decision-relevant feelings [14].

A framework [16] was recently proposed for describing the content of medical decision aids, which includes six dimensions: decision type, time scale, measurement types, data source, personalization level, and information format. It identified an understudied design space: interactive and animated decision aids. This is the gap we are trying to address with TreatmentExplorer.

Storytelling is regarded as the next step for visualization [17] and researchers have focused on telling stories about the analysis process or the data itself. TreatmentExplorer uses aspects of storytelling by using narration to help users learn how to read an unfamiliar visual representation.

We use animation which has been shown to be helpful in some cases e.g. to reveal trends over time or transformations of graphical representations during exploration [18], [19] but new examples and further user studies are needed to sharpen our understanding of which situations benefit from animation and which ones do not [20]. The design of TreatmentExplorer was inspired by the Lifeflow visualization which summarizes temporal patterns of electronic health records [21].

TREATMENTEXPLORER
While this paper focuses on the information presentation aspects of the decision aids it is useful to understand the source of the data used by the decision. We first introduce our data-driven approach, and then describe the main characteristics of TreatmentExplorer using the framework dimensions [16].

Data-driven decision aids and personalization
Unlike the vast majority of decision aids (which present results gleaned from the medical literature) or the more recent model-based decision aids (which compute somewhat personalized information based on answers to short questionnaires, e.g. [22]) our prototype summarizes data from Electronic Health Records (EHRs) of patients with similar profile and history, to produce personalized decision aids. This data-driven approach [24] allows TreatmentExplorer to reflect the most up-to-date data relevant to a condition and patient. It also enables TreatmentExplorer to be used for a great variety of conditions. Deployed in a clinic environment, a fully-operational system would be useful to many patients if access to EHR data for enough similar patients was available. Coming back to our original imaginary scenario, Donna’s physician would first select relevant elements from
her patient record to guide a search for similar patients [24], then review the usefulness of the data aggregated from those similar patients, then ask Donna to use TreatmentExplorer - either on her own or together with the physician [25] - to compare the effects of the different treatments for patients like her. While this vision may still be futuristic, TreatmentExplorer’s interface could already be useful today using data generated by models.

We now review the main characteristics of TreatmentExplorer and how it further differs from existing decision aids.

**Multi-Option Decision Support**

TreatmentExplorer supports the visualization and comparison of multiple treatment options. By supporting comparison, patients using TreatmentExplorer are relieved of some of the cognitive burden of assimilating data from multiple decision aids using different formats. Patients are able to explore the differences between treatments as well as learn the details of individual treatments. Animated transitions facilitate comparison between treatments, first slowly guided by text-based narration, then more rapid for back and forth comparison.

**Measurement Types**

The majority of decision aids focus on comparing a single measure of effectiveness for each treatment (for example the proportion of patients who get cancer) [16]. Only text decision aids have attempted to provide the rich picture of the likely future of patients selecting a treatment over another one. Decision aids using graphics become quickly crowded with the many charts and graphs needed for each measure so limit themselves to one or two measures. In contrast TreatmentExplorer attempts to combine multiple measures in a single interactive visual display: i.e. the proportions of patients 1) remaining healthy after treatment, 2) experiencing a specific outcome (e.g. breast cancer), and 3) afflicted by side effects and other secondary outcomes linked with the treatment (e.g. another type of cancer). Vertical sizing of each visual element is controlled by the proportion of patients represented. The average onset time of a condition is shown (an exact cumulative distribution of onsets is available as an advanced option). Information about side effects associated with treatments include the number of patients reporting each side effect as well as the date of the earliest reporting and the date of the last reporting of the side effect.

**Multiple Projected Points of Time**

Opposed to most decision aids which only provide information about one or two time points (e.g. chance of having cancer at the 5 year or 10 year mark) [16] TreatmentExplorer provides the average time of the outcome (e.g. on average women following this treatment who develop cancer will develop cancer after 8 years on average) or the time range of side effects (e.g. during the 1st six months). To provide that information the horizontal axis of the TreatmentExplorer visualization represents time. Each treatment is represented using a consistent horizontal scale, allowing patients to compare the average time of condition onset as well as the duration of side effects.

**Animation and Guided Narration**

Figure 1. This TreatmentExplorer’s interactive decision aid (as used in the experiment), prepared for patient C. Oswald who has a condition leading to paralysis. She can see that there are 3 treatments (A, B, C) and review each treatment one at a time to compare them. All three treatments were a similar popularity (the 3 bars for A, B and C are of same height). C is currently selected. On the right a visualization summarizes treatment success rates over 10 years i.e. the proportion of healthy patients (green) versus patients who develop paralysis (gray), the average onset of paralysis (4 years), the proportion of patients experiencing side effects (75% of them) and when those side effects are reported (between 0 and 5 years). This complex display is introduced step-by-step using staged animation and text-based narration.
The decision aid framework [16] review of existing decision aids highlights that the majority of decision aids use text or traditional simple graphics (e.g. barcharts) as information format. Interaction is mostly nonexistent, with few exceptions (see [26] for best example). In contrast TreatmentExplorer proposes a novel visual display combining all the information needed to compare treatments. Staged animation and synchronized text-based

Figure 2. Animation steps in TreatmentExplorer. 1) first the proportion of patients who remained healthy after receiving the selected treatment appears, i.e the green color fills the appropriate section of the visualization. 2) next, the remaining patients i.e those afflicted with the paralysis condition appear (i.e. gray fills the remaining portion of the screen). 3) the average onset of the paralysis is added. 4) Treatment side effects appear. Each step is described with a text-based narration. A right arrow button cues users to step through at their own pace.
narration explain elements of the visualization (see Figure 2) and progressively increase the richness of the display. This animation draws from the lessons of [27].

DESCRIPTION OF THE INTERFACE
We now describe the typical experience of users with TreatmentExplorer. As example scenario and data we use the condition and treatments used in the controlled experiment - see later sections of the paper. First the user sees the title information i.e. “Prepared for patient C. Oswald” (who has been told that she has a condition leading to paralysis.) She can see that there are 3 treatments (A, B, C) with a bar chart showing the relative popularity of the three treatments (in the experiment the 3 treatments had similar popularity i.e. the 3 bars for A, B and C were of same height – see left side of Figure 1). A text box invites her to click on one of the treatments.

First treatment: When the user selects Treatment C the axes of the visualization appear on the right, and she is told that she will be guided step by step by pressing the Right Arrow button. When she does the step by step animation begins:

1) A green rectangle fills the appropriate section of the visualization (vertically from the bottom) to match the relative portion of patients who remained healthy after receiving the selected treatment (step 1 of Figure 2). Pressing the arrow again animates to the next step:

2) A gray box fills the remaining portion of the screen to indicate the proportion of patients that became paralyzed (as explained in the dialog box – see step 2 of Figure 2).

3) Next a black bar is animated from the left (treatment time) to the right (here the 4 year mark) to indicate the average onset of the paralysis – which is also described in the text narration.

4) Next a rectangle of brown color is added to indicate treatment side effects (proportion of patient getting them, and average range of reports (i.e. likely start and stop of the effect). In this example side effects partly overlay the healthy patient groups as well as the paralysis patients to show that patients in both sets may experience this treatment-associated side effect. While this example includes a single side effect, multiple such treatment risks can be displayed when appropriate.

Each step of the narration is cued by the patient by clicking in the visualization with the mouse or by a key press so that patients can take as much time to read captions and study the visualization as they want. Each step begins with the display of the text and captions (so users have time to read them) then the animation of the graphic elements begins after a brief delay. When the animation has completed for a step, the prompt appears to alert the patient that they may continue when they are ready.

Once all steps have been completed, the user can either review the same data on his own (i.e. replay the step by step animation, or mouse over elements of the display to re-read the corresponding narration text) or select a different treatment. When a different treatment is selected the visualization is blanked and the information for the new treatment is presented using the same method. When all treatments have been reviewed in detail at least once, the speed of animation is increased to a much faster pace so that the final visualization can be accessed quickly for back and forth comparison. In the fast animation, elements animate from the old position to the new position (e.g. when switching from treatment B to C, the time of onset would animate from the position it had in the B treatment to the position for the C treatment, allowing user to see how the time of onset changes to a point earlier or later).

To the best of our knowledge no other decision aid attempts to present such rich information in a personalized interactive environment. While the final visualization may be very complex we believe that the use of staged-animation and text-based narration enables users to learn the meaning of the visual representation, which they could then use to remember and compare the benefits and risk associated with each treatment.

To evaluate TreatmentExplorer and its suitability for use as a decision aid, our evaluation process involved expert reviews, a pilot study, and two controlled experiments comparing TreatmentExplorer. We discuss these evaluations and the feedback we received in the following sections.

EXPERT REVIEWS
Four experts in the fields of medicine, public health, and risk communication were given a demo of the TreatmentExplorer prototype and interviewed for design improvements in sessions lasting about 1 hour. Suggestions and feedback were then incorporated into the next version (sometime as option) before interviewing the next expert. Feedback on our evaluation plans and adequate baseline for comparison was also gathered.

The demonstration used a synthetic dataset simulating 90 patient records with rates of cancer and side effects mimicking the breast cancer fact sheets available from the National Cancer Institute [28] and the Susan G. Komen Breast Cancer Foundation [29]. Three treatment options were available: routine care (no treatment), medication (based on Tamoxifen), and surgery (based on double prophylactic mastectomy). Both the medication and surgery treatment options carried a risk of some side effect based on side effects reported in the fact sheets.

TreatmentExplorer was received very well and experts said they hoped to get a version of TreatmentExplorer to investigate their own data. Desired improvements centered on improving the visualization and making it more accessible to patients:

- Using color suitable for patients with colorblindness
• Further simplify the animations so that patients only need to follow one moving object at a time
• Include a meta-description of the represented health records (i.e. How many patients are represented and how long their records provide data for)
• Use a single consistent time frame across treatment
• Animate visual elements from the bottom of the chart
• Overlay side effects over both healthy and condition-developing patients in the visual display
• Add an option to show the distribution of condition onsets to show skew or outliers (instead of only the average)

We discussed the possible role of TreatmentExplorer in clinic environments and refined a use scenario where both patient and physician select treatment options together during an initial consultation. The patient could be introduced to TreatmentExplorer and have any initial questions answered. The patient would then be given access to TreatmentExplorer to use on their own time after the consultation so that they could continue to review the information and explore treatment options.

USER STUDIES
Our user studies compared a text-only interface, a static version of the TreatmentExplorer prototype, and the full-featured prototype with animation. A 10 question questionnaire was used to determine knowledge gained by participants.

Research Questions
The study was designed to answer the following research questions:
1. Are there statistically significant differences in the number of times participants must consult the decision aid in order to answer questions accurately?
2. Are there statistically significant differences in the number of incorrect responses to questions participants provide between the three different decision aids?
3. Are there statistically significant differences in the subjective ratings of usefulness?

Pilot Study
A pilot study with 24 participants allowed us to refine our testing procedure and identified usability issues which were corrected before the main study. For example: rapid animation of graphic elements to show the differences between two treatments after both have been reviewed slowly at least once. While we had only recorded the total number of incorrect answers before reaching the correct answer, our observations suggested that participants could recall more information from the first trial with the full interface so we decided to record more precisely at what time participants were providing the correct answer, i.e. either from recall, or after looking at the interface once more, twice, etc. Results also suggested that participants had more difficulties understanding the treatment risks - such as side effects - so additional text labels were added on the visualization and the color legend was refined.

Main Study participants
42 participants were recruited on campus through emails to mailing lists, paper fliers, and verbal advertisements. These participants came from the undergraduate and graduate population and were compensated $10 for their participation.

Data
A synthetic dataset was created for a fictitious condition so that no participant would have prior experience or knowledge of the condition, its treatments, or their side effects (see Figure 1 and 2). Participants would thus need to use the decision aids provided to them in the study to complete the questionnaire. This synthetic data and fictitious condition also eliminated the risk of participants ever developing the condition themselves in the future and drawing on information and experiences from this study as part of their actual personal health decision process. The synthetic dataset consisted of fictitious patient records for 120 patients dealing with the medical condition and three possible treatments, each with a single unique adverse side effect.

Experimental Design
The study followed a 1x3 between subjects design with participants using only one of three possible decision aids to complete a short questionnaire about a fictitious condition. The three decision aids were each presented in the form of a website displayed in a maximized Firefox web browser on a 36x18 inch monitor set to a resolution of 1920 x 1080 p. Participants could use a mouse to scroll through the decision aid and point and click as desired.

Decision Aid 1: Text-Only
A text-only decision aid was created based on the layout and contents of a 2-page summary of Type 2 Diabetes oral medications produced by Consumer Reports Health: Best Buy Drugs [30]. This decision aid was thus a realistic analog of other text-based decision aids that patients would likely consult when trying to choose between multiple treatment options available for a single condition.

Decision Aid 2: Static-graphic TreatmentExplorer
A functionally limited version of TreatmentExplorer was used as a second level designed to provide a limited experience. This version used the same layout and visualization as the full-featured TreatmentExplorer, however the guided narration, animation, progressive disclosure, and reinforcing captions were removed. This eliminated all interactive features of TreatmentExplorer so that participants would have access only to the static visualization and highlights for each treatment option. This level of interface was intended to isolate the effects of interaction and animation on the usability and effectiveness of TreatmentExplorer.
**Decision Aid 3: Full-Featured TreatmentExplorer**

The final level of this evaluation made use of the fully-featured TreatmentExplorer prototype including guided narration, animation, progressive disclosure, and reinforcing captions.

**Dependent Variables**

A short questionnaire of 10 questions was designed to be filled out by each participant while they used one of the three levels of decision aid. This questionnaire was adapted from the “Questions You May Want to Ask Your Doctor” sections of the National Cancer Institute’s guide for breast cancer treatments [31]. Thus, the questionnaire reflected realistic questions that patients in a real-world healthcare situation would likely need to have answers to. Time was recorded as the cumulative sum of times a participant needed to consult their decision aid. Researchers noted whether or not participants made full use of their time exploring their decision aid or requested the questionnaire early. The cumulative sum of incorrect questionnaire responses provided by participants was recorded. After each chance to complete the questionnaire, researchers marked all responses as correct, incorrect, or unanswered. Responses were either completely correct or incorrect, no partial credit was given. Participants were allowed to provide new answers after exploring treatment options with the decision aid. These new answers were also marked and added to the cumulative sum of responses.

**Procedure**

The following procedure was used for each participant in this study:

1. Participants read the imaginary scenario (in short “you have just been diagnosed with Crimson Blot Syndrome and need to learn about the three possible treatments”). They received no training at all about the interface.

2. Participants were taken to the website of their decision aid and given 3 minutes to investigate their treatment options. Participants did not have access to the questions during this interaction period.

3. The decision aid was taken away and participants were given their questionnaire and a maximum of two minutes to answer as many questions as they could. Participants did not have access to their decision aid during the questionnaire period so answers were based on recall.

4. After the 2 minutes researchers marked all answers on the questionnaire as correct, incorrect, or unanswered.

5. Participants were asked for their subjective rating (from 1 to 10) of how well prepared they felt after their first interaction with their decision aid.

Steps 2-4 were repeated until the questionnaire was completely and correctly filled out.

6. Participants were debriefed on the nature of the study and given the opportunity to ask any resulting questions. Two subjective debriefing questions were also asked:

   i. What feature of this decision aid did you find the most or least helpful?

   ii. What other additional information would you want to see as a patient making a treatment choice with this decision aid?

During the session, participants were not allowed to use calculators, consult any other materials, or take notes of their own. Participants did not have access to both their decision aids and questionnaires at the same time. They were given the same questionnaire with their previously marked answers to continue working from.

**Results**

One participant using the full-featured interface lost focus of the task. He kept commenting on the interface rather than paying attention to the content after the first trial. His performance with 13 errors meets Chauvenet’s criterion as spurious and was removed as an outlier.

A one-way ANOVA (three treatments) assessed the total number of incorrect responses for all ten questions. The means were 7.21 for text-only, 6.00 for static-graphic, and 3.77 for the full-featured interface (p<0.05, F(2,38)=3.74).

We split the ten questions in two groups: five questions about the main symptom of the condition (i.e. paralysis), and five questions about the side effects associated with the treatments (i.e. sweating, headache and rash). For the first group (questions about the main symptom), there were significant differences among the number of incorrect responses across the interfaces. The means were 2.21 for text-only, 2.43 for static-graphic, and 0.69 for full-featured (p<0.01, F(2,38)=8.32). For the second group (questions about side effects) there was no significant difference. The means were 5.00 for text-only, 5.37 for static-graphic, and 3.08 for full-featured (p=0.05, F(2,38)=2.10). Looking in more details at the individual questions related to side effects revealed significant differences in participant’s total number of incorrect responses for the question which asked to list all the possible side effects of all treatments (1.21 for text-only, 0.64 static-graphic, 0.31 full-featured, p<0.05, F(2,38)=4.46) and for the question asking to name the most common of the three side effects (1.93 text-only, 0.79 static-graphic, 0.85 full-featured, p=0.01, F(2,38)=5.31).

The mean number of correct responses after the first trial (i.e. based entirely on recall) were not significant with means of 4.50 for the text-based interface, 4.29 for the static-graphic interface and 5.3 for the animated one (p>0.05, F(2,38)=2.92).

Similarly one-way (three treatments) ANOVA analysis found no significant differences between interfaces for the number of times participants had to consult the decision aid to answer all questions correctly. Participants’ subjective
ratings were also not significant (note that participants only used one of the three versions).

Discussion
Results suggest that participants using the full-featured TreatmentExplorer interface demonstrated better knowledge gains about the treatments’ risks and benefits than participants using the static or text-only interfaces. Since there was no significant difference in term of time or preference rating it can only be said that using the fully-featured version did not negatively impact participant completion times or experience. That is, walking through a series of guided narration did not slowdown participants when compared with participants using familiar formats like text.

During the study participants seemed more engaged when using the full-featured and static interface compared to the text-only interface. For example, participants using the text-only interfaced seemed to lose interest and get bored after reviewing the information once and waited for the trial time to end, especially after the first one or two sessions. On the other hand participants who used the full-featured or static interfaces went back and made full use of the trial time. Seven participants who used the text-based interface recommended that they would like to see the data in tables or charts. Text-based interface users who had time to see the full-featured interface after the experiment all reacted positively to it – mentioning that it would be a more efficient solution.

Additional observations and feedback from participants suggest possible differences between versions. For example, three of the static and full-featured interface participants commented that they could easily remember how the proportion of healthy and paralysis-afflicted patients resized between treatments. One participant even drew the final visualization for a treatment on the questionnaire when he struggled to answer a question. During the debriefing six participants spontaneously made comments about the usefulness of color, e.g. “the color coding helped to distinguish elements”, or “the green and grey were always balancing each other but the yellow-ish (i.e side effect) colors changed from one treatment to the next”.

A number of comments referred to the complex graphics used in the static-graphic and full-featured interfaces. Six out of thirteen participants using the static-graphic version said that the graphic was either “confusing” or “not-friendly”; one commented that while the proportion of healthy and paralysis-afflicted patients was obvious, other features of the (static) visualization were more confusing. However, negative comments about graphics from participants who used the full-featured interface were both fewer in number and more specific. One participant said that there were “too many numbers (and that) she panicked”, another that it was “tricky to understand” and a third one said that it was “confusing at first but made sense later”; but other comments focused on smaller aspects e.g. about the axis labels.

For the full-featured interface, there were positive comments about the animation being helpful. For example one commented that terminology like “onset” was confusing at first but that it became clear once they watched the bar representing onset animate between treatments. Two participants commented that they appreciated the text description of the visualization as it animated and felt they learned it quickly with the text-narration to guide them. We saw participants who used the full-featured version gesture with their hands and mimic the animations while they were trying to remember what they had seen.

While the full-featured version with visualization and narration seemed useful there are still many aspects of the design that could be improved. Showing a view of all three treatments side-by-side will also help patients compare among alternative treatments. Other possible improvements include: changing the color of the graph from green to gray at the average onset of paralysis rather than at the time zero; moving the notification text-box to the side so that it does not overlap any part of the graph; animating the graph elements from left to right to give a sense of time continuity.

We observed participants clicking rapidly at the end of the first session to fasten the animation in order to finish in time. Further studies may want to vary the speed of animation to determine the best possible timing so that users do not need to wait but can easily link graph items with the annotations.

Another area of improvement could be the terminology used. Several of the participants had problems understanding the definition of ‘first onset’, ‘average onset’ and ‘last onset’. Furthermore, when participants were asked to identify the side effects, nineteen among the forty-one participants initially included the symptom of the condition (i.e. paralysis) as a side effect. A brief description of these terms may help patients to make better comparison among treatments. Finally, comments suggested that treatment cost would be an important issue for comparing treatments.

GUIDELINES FOR ANIMATION AND INTERACTION
Feedback from experts and observations during the user studies encourage us to propose the following guidelines for designers. They complement existing guidelines for use of animation and text-narration (e.g. [19]).

• Begin Each Narration from the Same State: Patients are perceptive of subtle unintended differences and will question the meaning of inconsistencies when they appear. To avoid this accidental confusion, all narrations should begin from the same clear starting state.

• Follow a Consistent Narration Order: Patients who watch a series of narrations will learn the order of information and begin to expect data points to follow the
prescribed sequence. This not only helps them to structure how they think about the data but also provides a way to help patients compare data points across narrations. Patients who are learning a narrative interface will also be less intimidated if their expectations about what will happen next are not defied.

- **Provide Both Replay and Skip-Through Options:** Some patients will find value in re-watching a narration if they feel they missed some information in prior viewings. Other patients will be quick to absorb information and may only need a quick reminder of the final state of a narration later. Providing options for patients to quickly skip to the end of a narration will allow them to refresh their memory while full replay options will support both sets of patients.

- **Explain One Data Point per Stage:** Stages of narration should be simple to allow patients to focus their attention to one data point at a time. Related data points should be grouped into a sequence of stages.

- **Give Users Control Over the Flow of Narration:** Patients interacting with a guided narration will be trying to understand the information presented to them and learn from the narration. Some patients may need to re-read text or re-think accompanying animation several times to feel comfortable. Patients should then have the control to decide when the narration advances to the next stage.

**CONCLUSIONS**

TreatmentExplorer is a prototype decision aid allowing patients to explore their treatment options and educate themselves alone or with their physicians on the benefits and risks of each treatment. To the best of our knowledge no other decision aid attempts to present such rich information in a personalized interactive environment (they typically present only one or two measures, at a single point in time, and simply list possible side effects). While the combined visualization of all the needed information may appear to be complex we believe that the use of staged-animation and text-based narration enables users to learn the meaning of the visual representation, which they can then use to remember and compare the benefits and risk associated with each treatment, to ultimately make more informed decisions.

Our empirical evaluation demonstrated that with the use of animation and guided narration TreatmentExplorer users with no prior experience were able to learn about treatment options, and they did so more easily than with a text based or static decision aid. Further refinements and in-depth studies with real patients learning about real treatment options will be needed to fully understand the role of novel interactive decision aids such as TreatmentExplorer, and how they might be best integrated into the clinical workflow. We hope that this early work will inspire further research on medical decision aids, and ultimately help patients get a greater understanding of their options in order to take greater responsibility for their care and to assert their preferences.

**ACKNOWLEDGMENTS**

We would like to thank Beth St. Jean, Linda Aldoory, and Min Qi Wang for their expertise and advice during the development of the TreatmentExplorer prototype. This work is supported in part by Grant No. 10510592 for Patient-Centered Cognitive Support under the Strategic Health IT Advanced Research Projects Program (SHARP) from the Office of the National Coordinator for Health Information Technology.

**REFERENCES**


[31] National Cancer Institute, "What You Need to Know


