ABSTRACT

Title of Document: DESIGNING INTERACTIVE DECISION AIDS FOR MEDICAL RISK COMMUNICATION AND EXPLORATION OF TREATMENT OPTIONS.

Lyndsey Franklin, Master of Science 2013

Directed By: Professor Ben Shneiderman, Department of Computer Science

Medical treatments carry unique benefits and risks which patients must understand in order to decide which of their options is best for them. Prior research has demonstrated that patients are ill-equipped to understand both medical terminology and the statistical information presented to them through standard decision aids. Patients are unable to use the information about treatments to make decisions and as a result make poor choices with regards to their healthcare. The contributions of this work are 1) a multi-dimensional model for describing the content of decision aids; 2) TreatmentExplorer, a prototype interactive decision aid designed to communicate treatment risks and benefits through the use of visualization, animation, and guided narration; 3) an evaluation of TreatmentExplorer with four experts in health communication; 4) a preliminary usability evaluation comparing the performance of TreatmentExplorer against design alternatives, and 5) guidelines for interactive decision aids based on the results of these preliminary user evaluations.

DESIGNING INTERACTIVE DECISION AIDS FOR MEDICAL RISK COMMUNICATION AND EXPLORATION OF TREATMENT OPTIONS.

By

Lyndsey Franklin.

Thesis submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Master of Science
2013

Advisory Committee:
Professor Ben Shneiderman, Chair
Dr. Catherine Plaisant
Professor Jon Froehlich
Professor Beth St. Jean

Dedication

For my dad, Lynn, who gave me the push and kept me going and my mom, Patricia, who gave everything to make my dreams possible.

Acknowledgements

"If needed," said the document template. Ha.

I am very blessed to have nearly 30 years of friends, family, and mentors to thank. This is a very short and incomplete list. If I've missed you here, it just means that I look forward to thanking you in person.

Thank you to my committee members who gave their time and guidance. Ben Shneiderman and Catherine Plaisant: I read your work and it inspired me to give graduate school a try. It has literally been a dream come true to work with you and I cannot thank you both enough. Beth St. Jean: this was all a vague idea until our first meeting. Thank you for coming through with just the right idea at the right time. Jon Froehlich: I've been excited about your work since the first brownbag. Thank you for getting up so early.

Thank you to my other expert advisors and reviewers: Seth Powsner, Min Qi Wang, Linda Aldoory. I learned much from you all and this could not have been done without you.

Thank you to my family. Aunt Yvonne and Uncle Don: you rescued mom and me from certain doom more times than I probably realize. My super siblings Ryan, Kari, Jana: I love you and your families more than words can express and I only hope to make you proud. Laura, Charlene, and Steve Hall: you've raised me as your own and you will always feel like home to me. Lynsey & Rob: my clones who kept me going with phone calls and care packages and visits and more support than I could have ever hoped for. Thank you.

Thank you to my new family. Kelly and Scott Boardsen: this was a strange and scary place until you came along. Thank you for being the brave, generous, loving people you are. I've learned so much from you.

Thank you to Mona Leigh Guha and Beth Foss. You were the first to take me seriously and invite me along. MapleSchool Fridays were the best part of that semester and it has been a joy to work with you. Charley Lewittes and Anne Rose: thank you for always having time for a friendly word and for being amazing problem solvers when I needed it most. Thank you of course to all of the HCIL: you were worth the drive east to be a part of. Thank you Rose Kirby: you were the first friend I made in Maryland and that made all the difference. Thank you Atif Memon: you are a remarkable educator to have time for every one of your students and simply having you to talk to was a gift.

Thank you to my co-working family back home: Andrew Cowell, Brett Didier, and Tim Seiple. I can hardly wait to be of service again. Thank you to my friends who recommended me: Rick Riensche, Scott Butner, Frank Greitzer, and Mark Whiting. I will continue to make myself worthy of your vote of confidence. Thank you, Jack Hagemeister. A tremendous amount of my ability was learned in your class or that big comfy chair in your office you always had time to let me sit in.

Finally, I need to acknowledge my dog, Spectre, who reminded me throughout this whole ordeal that food, sleep, and playing outside were worthwhile pursuits.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

1.1 Problem Definition

Medical patients are increasingly required to make difficult decisions about their treatment while under stress. They must become knowledgeable about their possibly complex conditions and unfamiliar treatment options, sometimes in a short period of time. Along with coming to an understanding of their condition, medical patients face treatments that carry unique benefits and risks. Evidence-based medicine promotes the ideal that patients should be given data about their treatment options so that they can come to a logical decision based on treatments that have the greatest efficacy.

Unfortunately, patients are often unable to use the information presented to them about treatments to make decisions and as a result make poor choices with regards to their healthcare. Patients impose their own preferences on their medical care in terms of lifestyle-impacting side effects which they are willing to cope with as part of their treatment [1]. 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 [2] [3]. Patients face other obstacles to their rational decision making beyond their emotional state. It is widely noted that patients are ill-equipped to understand both medical terminology and the statistical data presented to them in decision aids. Quantitative information presented in tables and graphs is a barrier to patients with low numeracy skills. Even educated patients

have difficulties reading and understanding textual presentations of treatment information [4]. Interpretation of both numeric and textual information is a frequent hazard [5].

The increasing availability of laptops, tablets, smart phones, and the widespread use of the internet makes the problems facing patients and their information needs more important than ever. Patients are taking the initiative to learn more about their conditions from online sources [6]. This implies an increased expectation by patients to have access to relevant, reliable health information on demand. Work is needed to understand how to best leverage the affordances of online technologies and personal computing devices such that patients can both access and use health information to make better health decisions. Decision aids must be prepared that can support patients on demand as well as overcome patient skill deficiencies. Thus, the design of a useful decision aid faces challenges of risk communication and health literacy as well as other challenges such as the placement of a decision aid within a medical workflow. The following sections outline unanswered research questions from these three related fields which have driven the design of a prototype medical decision aid, TreatmentExplorer.

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. The study of risk communication within the medical domain strives to identify the factors which make medical data easier to understand when presented

to the average patient. Much of this work examines the role of graphics and the features of graphics that increase patient knowledge. Unanswered questions from the study of medical risk communication include [7] [8]:

- Which graphical formats are patients able to accurately read?
- How does the presentation of equivalent measures such as frequency or percentage affect patient understanding (i.e. 5 out of 10 patients vs. 50% of patients)?
- Does the order of information produce a bias towards or against a treatment?
- How does framing of risk (incremental vs. absolute, gain vs. loss) affect its interpretation?

Health Literacy

Patients do not have medical expertise and cannot easily understand medical terminology, treatments, or their options. Basic terminology and standard medical procedures are daunting onsets of information overload, especially when a patient is also emotionally burdened with a recent diagnosis of a serious medical condition. Before a patient can begin to understand the evidence supporting treatment options, they must first be able to understand what information they have and be comfortable navigating through decision aids and other informative materials. Health literacy is a problem for otherwise literate patients and compounds the obstacles faced by non-literate patients. Supporting the health literacy of patients is critical as low health literacy contributes to health disparities and is associated with poor outcomes

including increased hospitalization, higher mortality rates, and difficulties with disease management [9]. Design concerns for low-literacy patients include [4] [8]:

- What literacy level should be targeted by a decision aid?
- How can the needs of low-literacy patients be met without compromising data presentation?
- What features of a decision aid are most adversely affected by low literacy or numeracy?

Medical Decision Aids

Medical decision aids have served their purpose if they convey one data point and leave the burden of assimilating and understanding to the patient. Despite this additional cognitive burden, decision aids are consistently shown to reduce patient anxiety, reduce passivity, promote realistic perceptions of treatment benefits and harm, reduce negative emotions, and increase patient knowledge about risks and treatment options [10] [3] [11] [12]. Other work has shown that more guidance from trained medical professionals such as physicians or nurse practitioners improves patient understanding of risks as well as patient satisfaction with treatment options [10] [13] [12] [14]. This suggests there is room for improvement in decision aids with respect to their use as communication tools for patients and physicians. Information needs of patients and physicians need to be better understood as well as where communication interventions belong within the healthcare process. Open questions dealing with the use of decision aids include [14] [15] [9]:

- How can patient-centered communication at the physician or clinic level improve patient understanding of their treatment options?
- Where in the workflow of healthcare is the most appropriate place for evidence-based treatment information?
- What are the information needs of a patient and how do they differ from the medical professionals supporting them in their care?

1.2 Design Goals

An information visualization approach to the preceding challenges would begin with the following design goals:

Support Multiple Treatment Options

Patients often learn about their treatment options in isolation. The burden of consolidating the information about each treatment falls to them. Varying information sources provide varying types of information. A more ideal solution would support the comparison of multiple treatment options in the same space by making the same information available about all possible treatment options.

Go Beyond Treatment Success Rate

Included in the most basic information patients typically receive about a treatment option is its chance of succeeding. While this is important information, it does not provide patients with a complete understanding of what undergoing a particular

treatment option would entail. It is desirable for patients to be able to compare treatment side effects, costs, and other lifestyle factors for their treatment options.

Capture Changes Over the Course of Treatment

As a patient follows a care plan, their circumstances may change. Some treatment options may have a cumulative effect on their condition (such as regular exercise on cardiovascular health). Other treatments may offer immediate benefit but have diminishing returns over the course of a lifetime. This information needs to be available to patients so they can anticipate how they may be impacted over the entire length of their care but is often missing from most healthcare information patients encounter.

Leverage Real Outcomes from Real Patients

Apart from scientific evidence and the results of clinical trials, the testimonials and support of other patients is an important source of information for some patients. In fact, entire communities have arisen to provide just this type information [16]. However, this also results in an additional burden of finding credible community resources. Patients can benefit from knowledge of the treatment outcomes of real patients similar to themselves, but need a way of accessing credible information.

1.3 Contributions

The first contribution of this work is an analysis of the design space of medical decision aids based on the content existing decision aids provide. This analysis,

based on a literature review, yields a multi-dimensional model which can be used to assess decision aids relating to a wide variety of conditions and demonstrates the information gaps patients must overcome when reconciling the information they gather from multiple sources. It serves as a starting point to investigate questions about the impact of decision aid content on risk communication and the information content of medical decision aids.

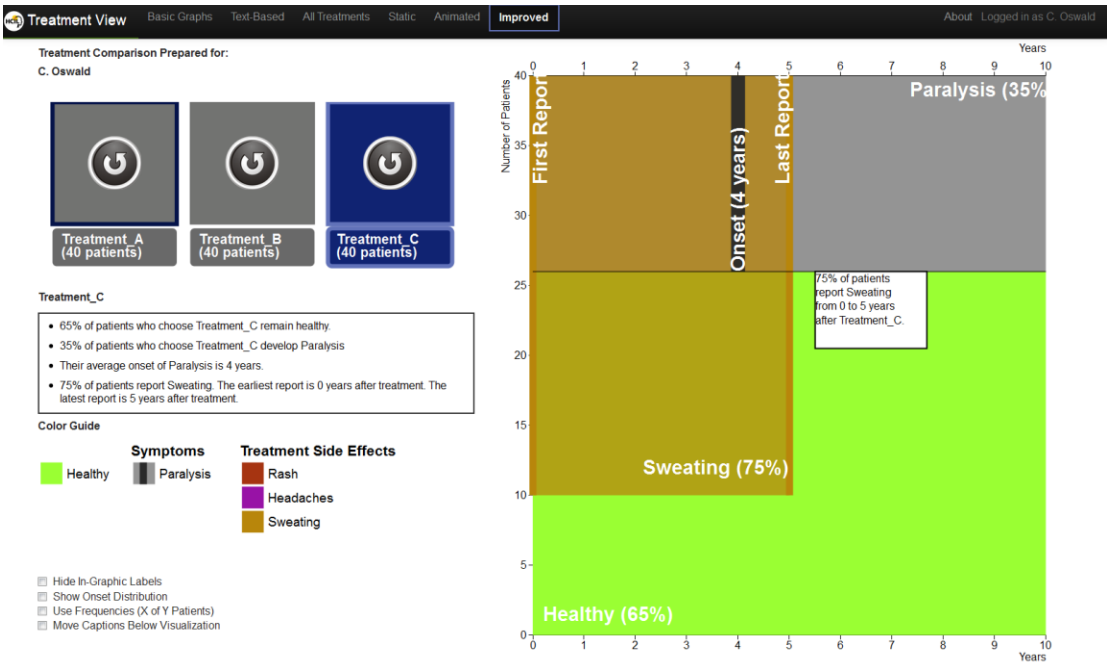


Figure 1. The TreatmentExplorer prototype for medical risk communication and treatment option exploration.

The second contribution of this work is a prototype a medical decision aid, TreatmentExplorer, which has been designed to address the information gaps identified through analysis using the dimensional model (see figures 1 and 2). TreatmentExplorer follows the best practices which have been put forth by related research and builds on that research with advanced features such as:

- The display of information for multiple treatments

- The representation of likely treatment outcomes for multiple points of time
- The inclusion of multiple data points for each treatment
- Personalization to patient-specific conditions
- Guided narration of treatment highlights with animation

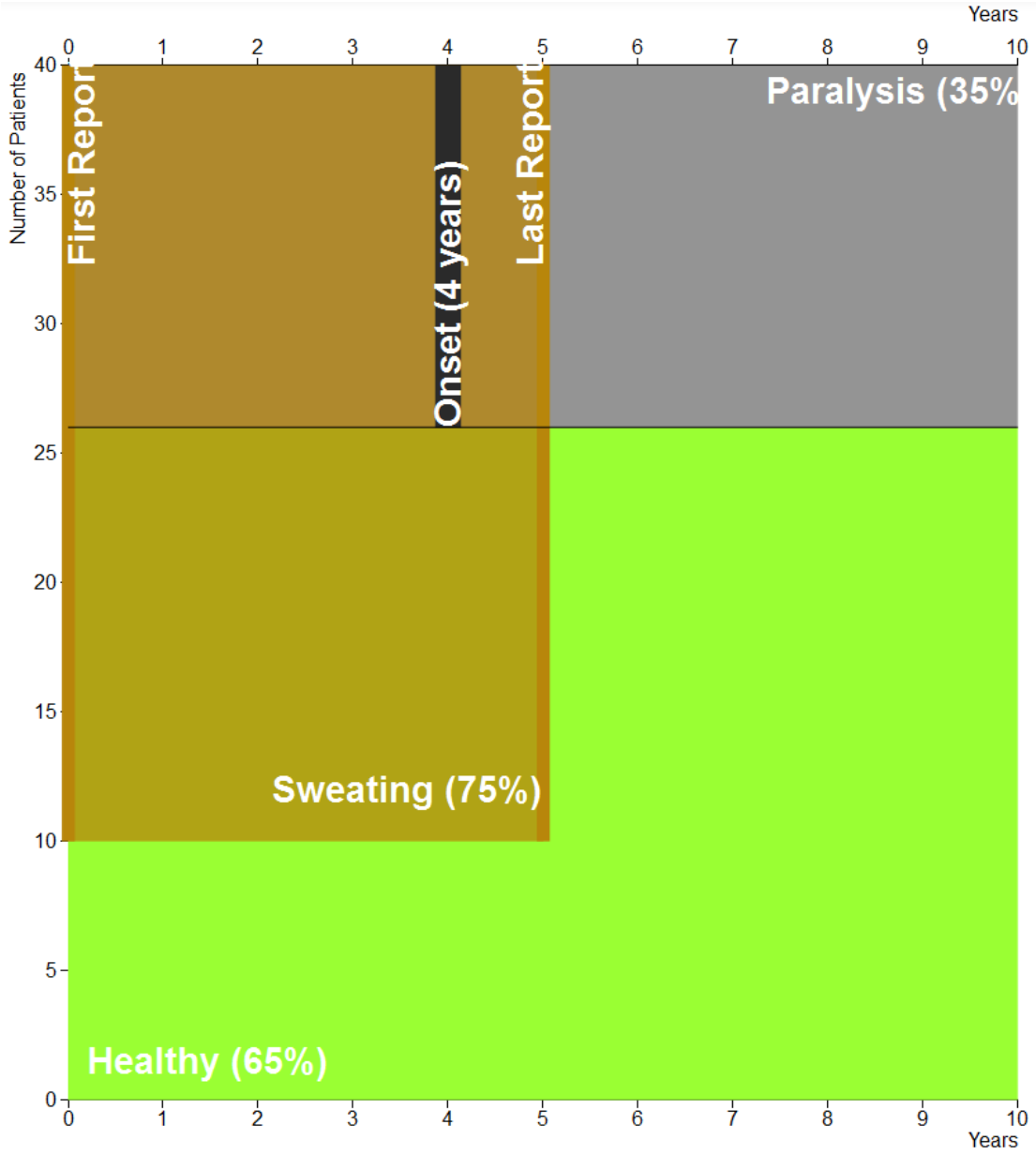


Figure 2. Visualization portion of the TreatmentExplorer which shows the overall success of a given treatment, the average onset of the primary symptom, and a side effect associated with the treatment.

TreatmentExplorer also provides a platform through which the preceding research questions from risk communication, health literacy, and medical decision aids may be answered experimentally with case studies and controlled experiments.

As part of the design of TreatmentExplorer, experts in health and risk communication have evaluated early versions of the prototype and provided feedback on its features and ease of use. This co-design process forms the third contribution of this work and resulted in additional best-practice recommendations for future decision aids.

The fourth contribution of this work is the evaluation of TreatmentExplorer with a preliminary user evaluation. This evaluation compares the knowledge gain participants experience when using TreatmentExplorer with the performance of participants using a baseline analog decision aid.

Finally, the fifth contribution of this work is a set of guidelines for the use of text, graphics, animation, and guided narration derived from the results of the TreatmentExplorer preliminary evaluation, pilot studies, and expert reviews.

Chapter 2: Related Work

TreatmentExplorer builds from the research of several fields including theoretical concepts of sense-making and thought visualization and empirically tested lessons from risk communication, health literacy, and decision aids. It also draws from research in the domains of information visualization and storytelling.

2.1 Sensemaking and Thought Visualization

Dervin's Sense-making theory describes the process by which individuals organize information in their attempt to reach an understanding (figure 3). Emphasis is placed on sense making and unmaking as an active "verbing" process with knowledge considered to be the sense made of information at a particular point in time-space by someone [17]. In the process of sense making and unmaking, individuals "bridge the gap" between their current knowledge and new information based on their history and experience resulting in new knowledge as well as functions and dysfunctions [17]. TreatmentExplorer responds to the call to support patient learning as an active process by providing a consistent representation of treatment options and probable patient outcomes. Patient-users are able to explore treatments and leverage an organized presentation to help them make sense of the otherwise overwhelming amount of information. This particularly addresses a tenet of sense-making theory which claims that interfaces designed dialogically with contributions anchored in material conditions and verbs will result in the higher capacity for understanding [17].

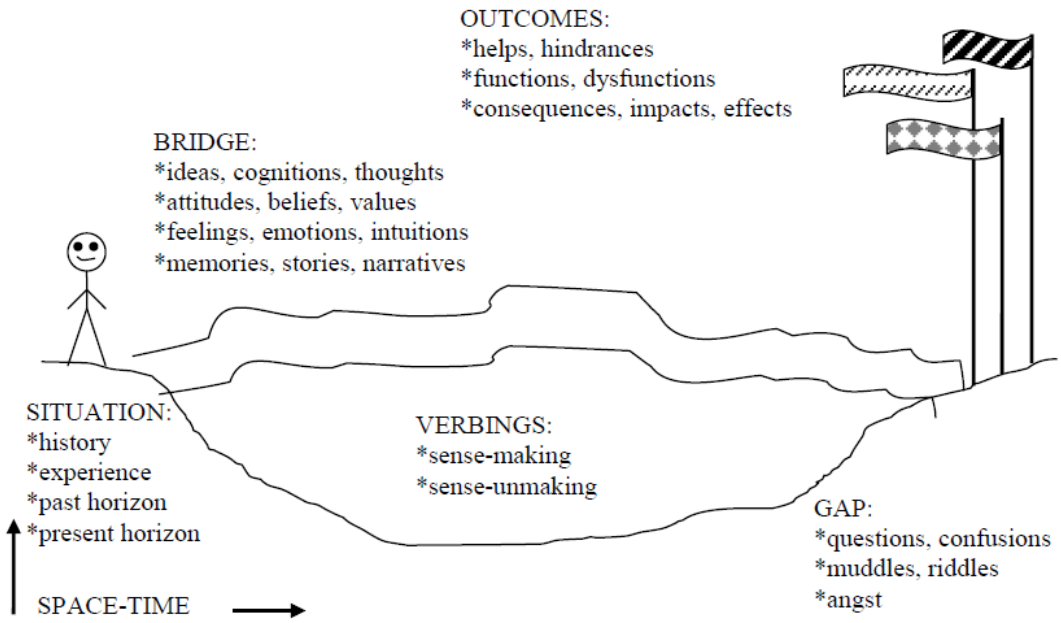


Figure 3. Dervin's Sense-Making Metaphor [18, p. 238].

Consistencies in visual communications provide insight into how people think and should be used to guide design [19]. Tversky discusses research supporting the use of spatial actions creating meaningful patterns such as groups and hierarchies [19]. Other uses of space indicate that vertical dimensions are preferred for graphics with values that can be evaluated such as sums with larger/higher values to higher spatial positions. The horizontal dimension is preferred for neutral concepts such as time [19]. Bars in diagrams are frequently interpreted as containers which separate their contents from everything else and nested bars or frames can be used to indicate hierarchy among contents [19]. TreatmentExplorer builds on research into effective visual communication. Patient groups and their size are represented along a vertical dimension as recommended while time is expressed along the horizontal axis in the TreatmentExplorer visualization. Glyphs are avoided in favor of simple visual indicators such as bars and lines. Progressive disclosure of visual elements with

explanatory text supports patient learning and creates shared mappings as patients make use of TreatmentExplorer.

2.2 Risk Communication

Research efforts in risk communication seek to understand the most effective mechanisms for communicating probabilities of uncertain and possibly hazardous events to non-experts. When applied to healthcare, the risks are often risks of disease, side effects, complications, and possibly death. Information about these risks typically is provided by medical professionals attempting to guide non-expert patients through treatment decisions.

Bunge, Muhlhauser, and Steckelberg surveyed risk communication literature for quality criteria in an effort to compile the evidence for the supported criteria [20]. They found that criteria established for the presentation of numerical data, verbal presentation of risks, diagrams, graphics, and charts were based on evidence. However, the content, loss-/gain-framing, and patient oriented outcome measures were based on ethical guidelines. Little research support was available for backing criteria on the quality of evidence, pictures or drawings, patient narratives, cultural aspects, layout, language, and development process. In this work, the focus is on making use of such established risk communication guidelines for presenting treatment options in a graphical way.

Levy et al. surveyed cancer risk calculators found on the internet to review their content and consistency [21]. Their results show that most risk calculators did not provide information to assess their credibility and that each calculator varied in

the factors it used to assess cancer risk. They warn that the potential to misinterpret cancer risk by using such sites puts patients at risk of making inappropriate medical decisions. In a similar study, Waters et al. found that risk communication formats varied between websites with few health care industry affiliated sites providing comparative risk information while the majority of government affiliated sites did so [6]. Use of formats to reduce bias and facilitate comprehension varied widely. In this work, the ability to explore multiple treatments and understand the details of each is a priority. Comparison of treatment efficacy is facilitated through both textual and graphical representations which remain consistent across all treatments. Content is designed to be provided through medical records providing credibility.

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 [22]. They note that frequencies are preferred for providing information about absolute risks and/or highlighting changes between levels of risk (figure 4). They also recommend repeatedly drawing attention to the time interval over which a risk occurs and the inclusion of graphs and summary tables. In this work, frequencies are available and the visualization builds on elements of the recommended pictographs. A summary of treatment highlights is always present with the visualization as well.

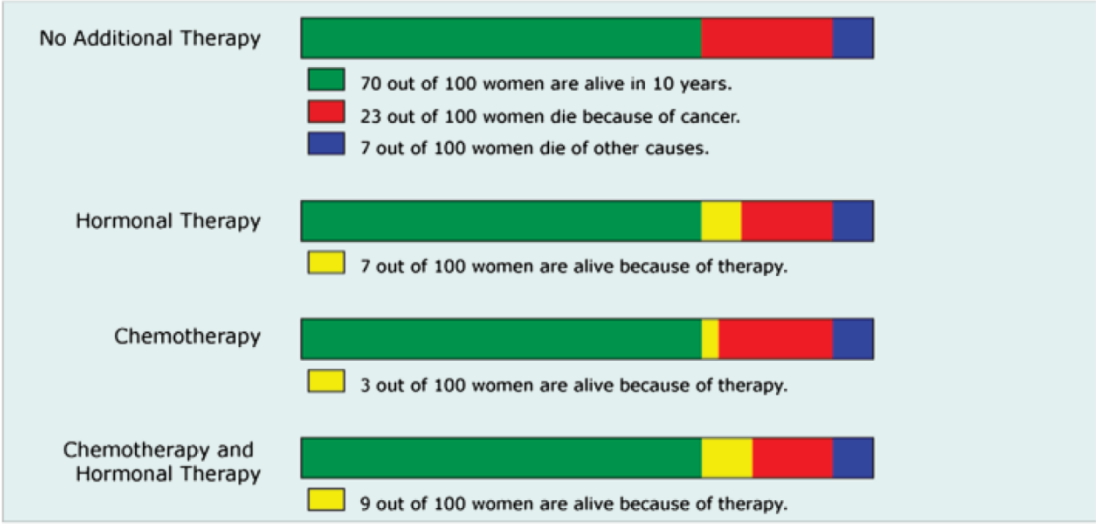


Figure 4. Example bar graph format from [23] [22] which provides survival rate information for multiple treatment options and highlights treatment differences using both frequencies and color.

Other work in risk communication suggests that even physicians need assistance with evidence-based risk communication. Factors affecting a physician's test-ordering tendencies have been shown to include their tolerance of uncertainty and both physicians and patients alike have shown difficulties understanding certain statistical measures such as numbers needed to treat [11]. Han et al. have also 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 [24]. Follow-up work endeavored to produce novel visualizations capable of representing randomness and its effect on uncertainty [25] (see figure 5). The work on TreatmentExplorer does not yet address uncertainty. Rather, it bases the data it presents on the information in Electronic Health Records (See Chapter 4). Thus, risks presented represent actual reports of medical incidents from patients. TreatmentExplorer also avoids use of the statistical measures which this prior work

suggests are confusing to patients and physicians. This avoidance of measures to focus on other critical statistics also follows one of the recommendations put forth by [22] which is to present only the most critical information, even at the expense of completeness.

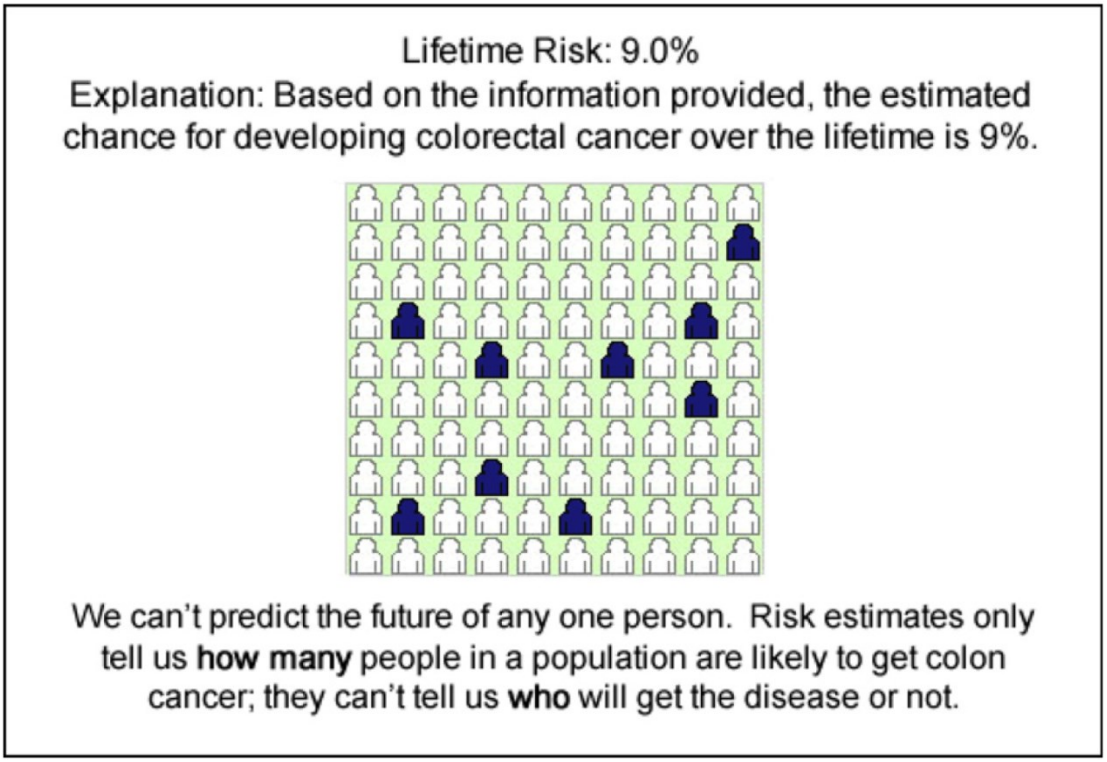


Figure 5. Novel visualization from Han et al. [25] to depict uncertainty within risk estimates.

2.3 Health Literacy

Healthy 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 (as well as patients without such deficiencies) are able to access. There is not the same emphasis on presenting a specific number to patients, rather summary information and relative ratings are included. Brown et al. are among several researchers who report work

confirming that numeracy, graphicacy, and health literacy are correlated [4]. Particularly, numeracy predicts graphicacy even after factors such as education are controlled for [4]. Other work supports the idea that factors such as numeracy affect both gist and verbatim knowledge of treatments in patients and that such knowledge is associated with medically superior treatment choices [26].

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 such as figure 6 [27] [20]. Other research suggests that the presence of graphics increases the believability of information [28]. Verbal expressions such as "low", "medium", and "high" have been shown lead to misunderstandings between physicians and patients and interfere with patient understanding [5]. This work builds on these finds by making use of graphics to represent the high-level properties of treatments. Where text is used, simple, numeric expressions of important information are used instead of easily misinterpreted phrases.

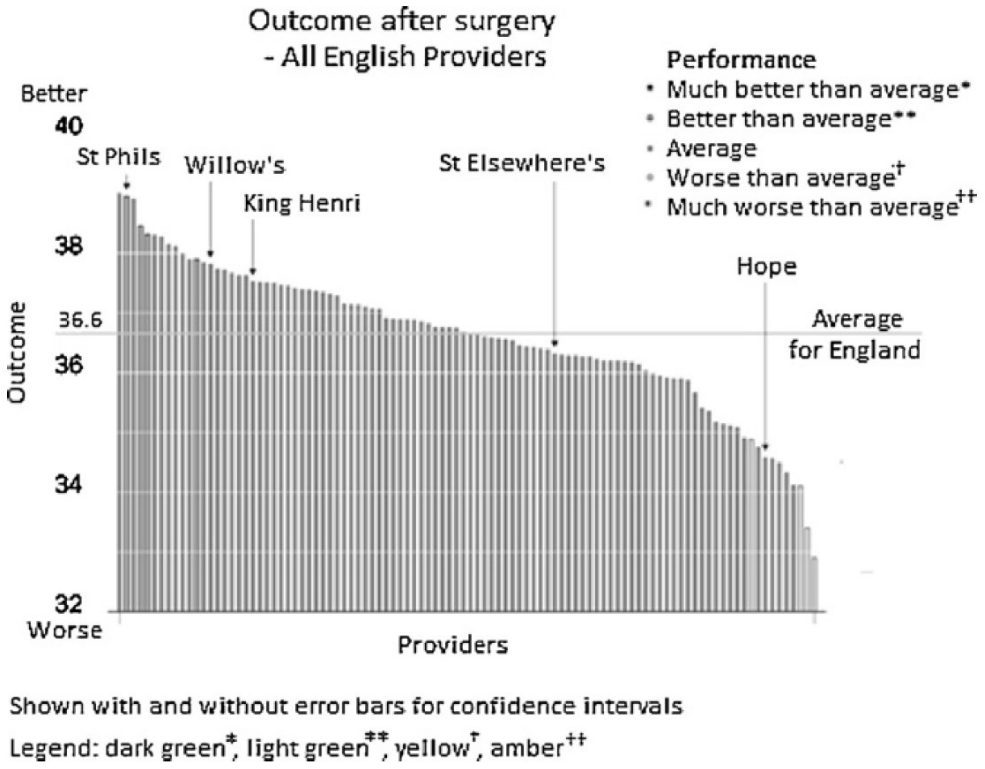


Figure 6. Alternative provider rating visualization studied by Hildon, Allwood, and Black in [27].

Other health literacy related research attends to basic human-computer interaction issues such as navigation through information. Menu structures were the focus of Chaudry et al. and their studies of chronically-ill, low literacy patients [29] (figure 7). Their recommendations for menu systems are to use large widgets and to include "home" and "back" navigation options to enable quicker navigation. Linear menu systems with breadcrumb-like trails were preferred as well as starting every task from the same location. TreatmentExplorer follows these recommendations through use of a large navigation widget where each treatment is represented as both a labeled button and as part of a bar graph. All the guided narrations in TreatmentExplorer are started from this single navigation widget and begin with the same empty screen.

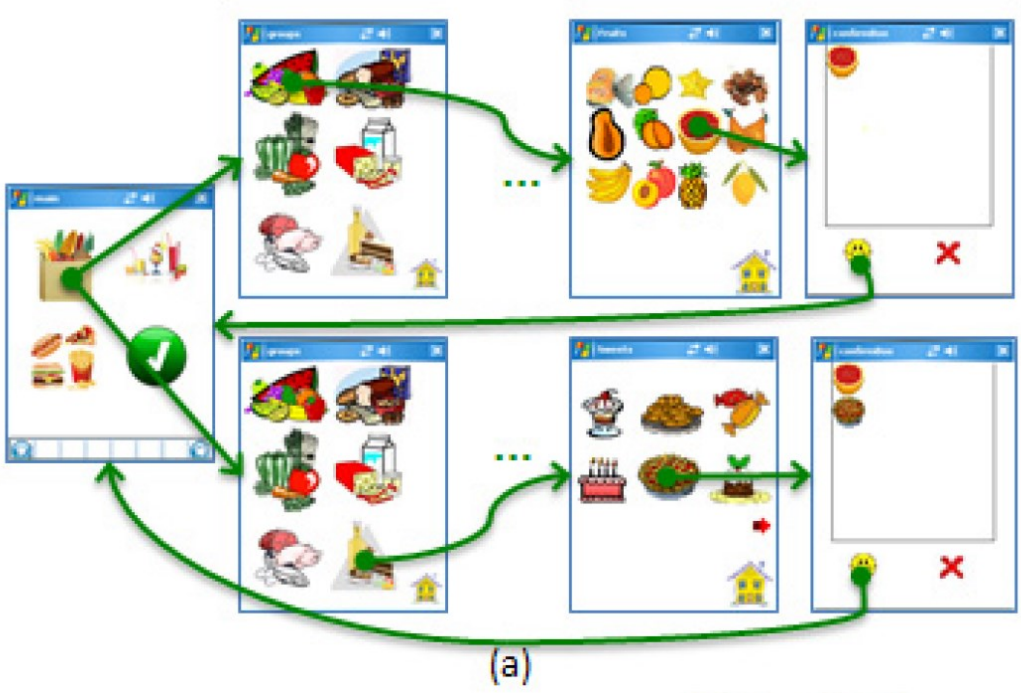


Figure 7. Example linear navigation and menu style studied in [29] for suitability for low-literacy patients.

2.4 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. Studies have explored decision aids in the role of support for medical professionals [30], physician-patient communication [13] [3], and patient education [15] [10] [31] [14].

Decision aids invoked in the role of physician-patient communication aid have shown a number of effects. Studies have shown that emotions such as anxiety interfere with a patient's 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]. Guidance on important questions and the ability to educate family members about conditions are noted as important benefits of decision aids [15]. Figure 8 shows an example of one such question answering system from [32]. Visual representations provide memory prompts which reduce the cognitive load of patients during decision-making [14]. Study results suggest that patients who use decision aids are able to express more emotions, use more cognitive terms, and verbalize decision-relevant feelings [14]. TreatmentExplorer has been designed to fill this role in the patient care process: by educating patients and providing cognitive support for remembering treatment details, patients using TreatmentExplorer should be able to communicate more effectively with their physicians and participate more in their healthcare decision making processes.

Figure 8. Screen captures from a prototype decision aid for cardiac patients to facilitate question asking and answering with their physicians and care providers (from [32]).

Along with decreased decisional conflict and easier communication, decision aids are associated with increased patient knowledge with the greatest gains from patients reporting low baseline knowledge [10]. Incorrect responses to survey questions dropped dramatically after patients had access to decision aids in [10]. Leroy discusses in [33] how patients access information through a variety of sources

now to educate themselves and use the information they find in making healthcare decisions. Patient education is one of the intended uses of TreatmentExplorer. Credible information derived from aggregated patient outcomes is provided to patients so that they can familiarize themselves with their treatment options.

Medical professionals can also benefit from decision aids and support systems. Lee and Bakken [30] report of work which provided nurse practitioners decision aids designed to support adherence to clinical practice guidelines for patients managing obesity. Their results indicate that such assisted professionals were better able to adhere to clinical guidelines resulting in improved care of patients. Patient goals were also captured with greater regularity and included as part of consultations. This emphasis on patient-centered interventions has been shown to increase the effectiveness in the nursing process [30]. TreatmentExplorer does not address a specific condition; rather, it supports medical professionals by providing a platform to communicate treatment information in an unbiased manner. This support included unbiased representations of clinical evidence backing treatments. Medical professionals benefit from the completeness of information in TreatmentExplorer as well as being able to use it as a starting point for patient-centered care.

2.5 Information Visualization and Storytelling

Information visualization research has always sought to establish the best practices for presenting data in the most salient ways. Information analysis typically requires interactive visualizations through which users are able to gain insight into their data. As noted in [34], design principles connect the visual design of a visualization with

the viewer's perception and cognition of the information conveyed. Important information can be emphasized and lesser relevant details de-emphasized through visual techniques. TreatmentExplorer draws from such visual communication principles that have been studied in information visualization research.

Kosara and Mackinlay in [35] argue that with the design space of information visualization well defined, suitable techniques can be found to represent most datasets. However, they also argue that best practices for communicating data are lacking. Storytelling is thus an ordered sequence of visualization steps which can include text that roughly correspond to time or causality. Several storytelling scenarios exist within visualization research including self-running presentations for large audiences, live presentations, and individual or small-group presentations. TreatmentExplorer follows the design mold of these individual presentation scenarios. Requirements for such individual presentations include flexibility and user control greater than that afforded by simple slideshows [35]. TreatmentExplorer explicitly follows this recommendation by providing a user driven guided narration in which the user controls the pace of the narration and has multiple animation options for transitioning between treatments.

In [36], visualizations from online journals, blogs, and instructional videos are analyzed for their techniques of storytelling with data graphics. Salient features which provided value to narratives included annotations, visual highlighting, and consistent visual platforms. In addition, it was noted that the use of single-frame interactivity helped encourage users to explore visualizations and provided a tacit tutorial of the available information and interactions. TreatmentExplorer takes

advantage of all this storytelling techniques by using a consistent frame for all narrations and annotating individual changes in the visualization as the narration progresses. TreatmentExplorer also provides animated transitions which have been credited with making these changes clear to users [36].

Automated systems are desirable when it comes to providing users with narratives to explain their data. Hullman, Diakopoulos, and Adar report on work which aggregates and summarizes stock behavior [37]. These annotations are chosen to provide additional context to support user interpretation of information. TreatmentExplorer combines this idea of visualization annotation with the ideas of data-driven exploration of care plans for patients as described in [38]. In [38], medical histories are mined for similar patients and outcomes of their care plans. TreatmentExplorer differs from this work by providing a patient-friendly visualization and focusing on the information needs of patients, rather than physicians. TreatmentExplorer also advances care plan visualization research by representing more than a single derived metric in its visualization.

2.6 Summary of Related Works

TreatmentExplorer is inspired from the results of research in risk communication, health literacy, and medical decision aids. Patients needing to make treatment decisions have been shown to be under-prepared to use the empirical evidence supporting their treatment options. Often, patients are not health literate and do not have the statistical background to make sense of the information. Complicating this are emotional factors which impede a patient's logical decision making. Graphics are

one option for making statistical information easier for patients to understand. Research has demonstrated that some graphic formats are better suited for communicating probabilities to patients than others. Other work has shown that even physicians struggle with statistics and graphics from time to time. There is room for future investigations to explore the use of interactivity in decision aids and their role as patient-physician communication aids. Techniques from information visualization and storytelling research are needed to provide patient-supporting experiences which can guide patients to an understanding of their likely health outcomes.

Chapter 3: A Multi-Dimensional Model of Medical Decision Aids

This chapter presents an information-centric, multi-dimensional model for describing the content of medical decision aids. This model was developed as the result of an extensive literature review and examining the content of decision aids with respect to the variety of information available and its presentation. It differs from traditional methods of evaluating decision aids by avoiding simplistic ratings of patient gist and verbatim knowledge produced by a decision aid. It also avoids addressing healthcare business workflows which require examining factors irrelevant to the decision aid itself such as how information is updated.

3.1 Decision Type

The types of decisions that need to be made in a healthcare scenario range in complexity from the very simple to the very complex depending on the treatment options available. The severity of the condition a patient faces will impose emotional and social complexities and dramatically affect the time frame in which treatment decisions are made. However, in the context of this model this does not impact the classification of the decision presented by a single given decision aid. That is, this model is designed to be condition and stage-of-condition independent. The risks, implications, and consequences of treatment choices to deal with an aggressive cancer are severe, yet from an information perspective the types of decisions required share

similarities with the decisions of a patient with a seasonal allergy: multiple treatment options may be available, some treatments will be more or less effective than others, and repeated attempts to manage the condition may be needed before ultimate success is found. Table 1 summarizes the classifications within the Decision Type dimension.

Table 1. Types of Decisions Supported by Medical Decision Aids

Decision Type	Description	Example
Binary	A patient may either opt for a treatment option or take no action.	A patient may choose to have an abnormal growth biopsied or wait and observe it for changes first.
Multi-Option	A patient has more than one treatment option available, but some or all of them may be mutually exclusive.	A patient must decide how to best reduce the immediate risk of cardiovascular disease through diet, medication, or surgery.
Combination Options	A patient has more than one treatment option, several of which may increase the likelihood of others providing benefits.	A patient looking to reduce a high risk of lung cancer decides to stop smoking as well as improve dietary and exercise habits.
Continuous Management	A patient has made critical decisions about their care and now needs assistance in maintaining current treatment plans.	A diabetic patient must learn how to plan meals and improve food choices.

Binary

Binary decision aids are the most common in both research literature and healthcare industry use (e.g. figure 9). They are the least complex and often the decision to be made is whether or not a patient elects to undergo the presented treatment option [7] [39] [5] [40] [41]. Studies indicate that framing effects of personalized risk and comparisons with the risk of an average population have a significant effect on

whether or not participants opt for treatment [42]. The order in which risks and benefits of a treatment are presented has been shown to influence whether or not participants felt positively about treatment though the presence of contextual information eliminated this bias [43]. Edwards et al. [44] report on a randomized controlled trial where diabetes patients chose between a "treatment as usual" management plan and a "tight control" plan based on information presented in either text, graphical, or text and graphic formats. Their findings suggest that the format of the decision aid had no significant effects on the reduction of decision conflict, however participants preferred simple graphics which avoided anchoring information and induced information overload. Other work, however, suggests that intention to undertake recommended lifestyle changes was in fact influenced by graphical formats but only in participants also receiving high-threat communications [27] [28]. Graphical decision aids have also been shown to impact the emotional response [45] of participants and decrease their passivity in counseling sessions [14].

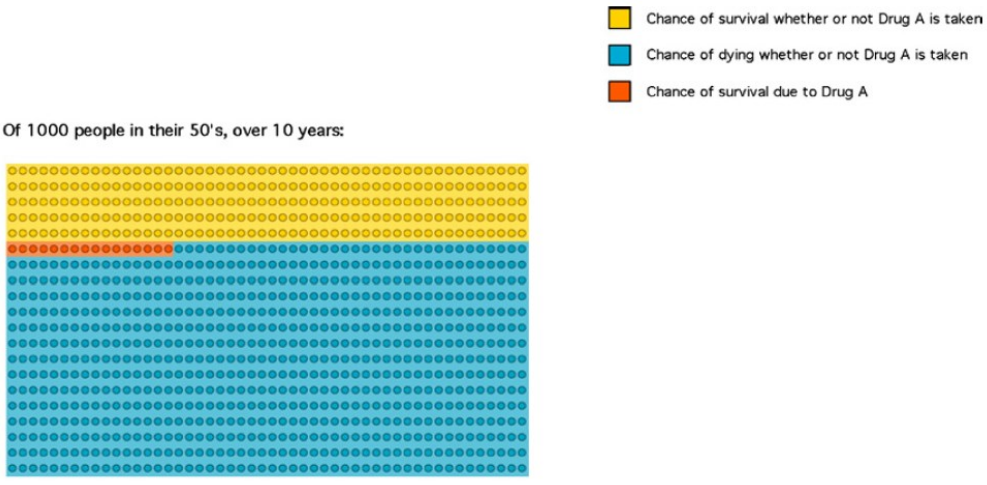


Figure 9. An example from Price, Cameron, and Butow [40] of a graphic supporting a binary decision.

Multi-Option

Multi-option decision aids are often used when there are multiple treatments a patient can consider at once that for varying reasons are mutually exclusive of each other (see figure 10). For example, an early stage prostate cancer patient may be given a decision aid to help decide between surgery, radiation, chemotherapy, or simply a watch-and-wait strategy [1]. Or, a patient at high risk for breast cancer may need to choose the extent of preventative surgery [4]. A key difference between multi-option decisions and binary decision is the emphasis of the decision aid: for a multi-option decision the information must support comparison between options through its format. This is different from a collection of binary decision aids which each individually tell a specific story about a single treatment. The collection may not provide the same information about each option nor will each aid necessarily provide measurements which can be compared across decision aids. Comparison is a difficult information process to support in some formats and studies have found that while participants prefer simple, familiar graphics, their knowledge as measured by question-answer accuracy is higher when more sophisticated graphics are used [1]. When used as part of consultations, multi-option decision aids have been shown to improve knowledge and reduce decisional conflict [15]. When choosing from amongst treatments, participants tend to prefer symbols to numbers to represent strength of the recommendation or evidence for the treatment and incremental risks are consistently perceived as lower in text-only decision aids [20]. Other work has shown that both verbatim and gist knowledge are significantly associated with medically superior treatment choices [26].

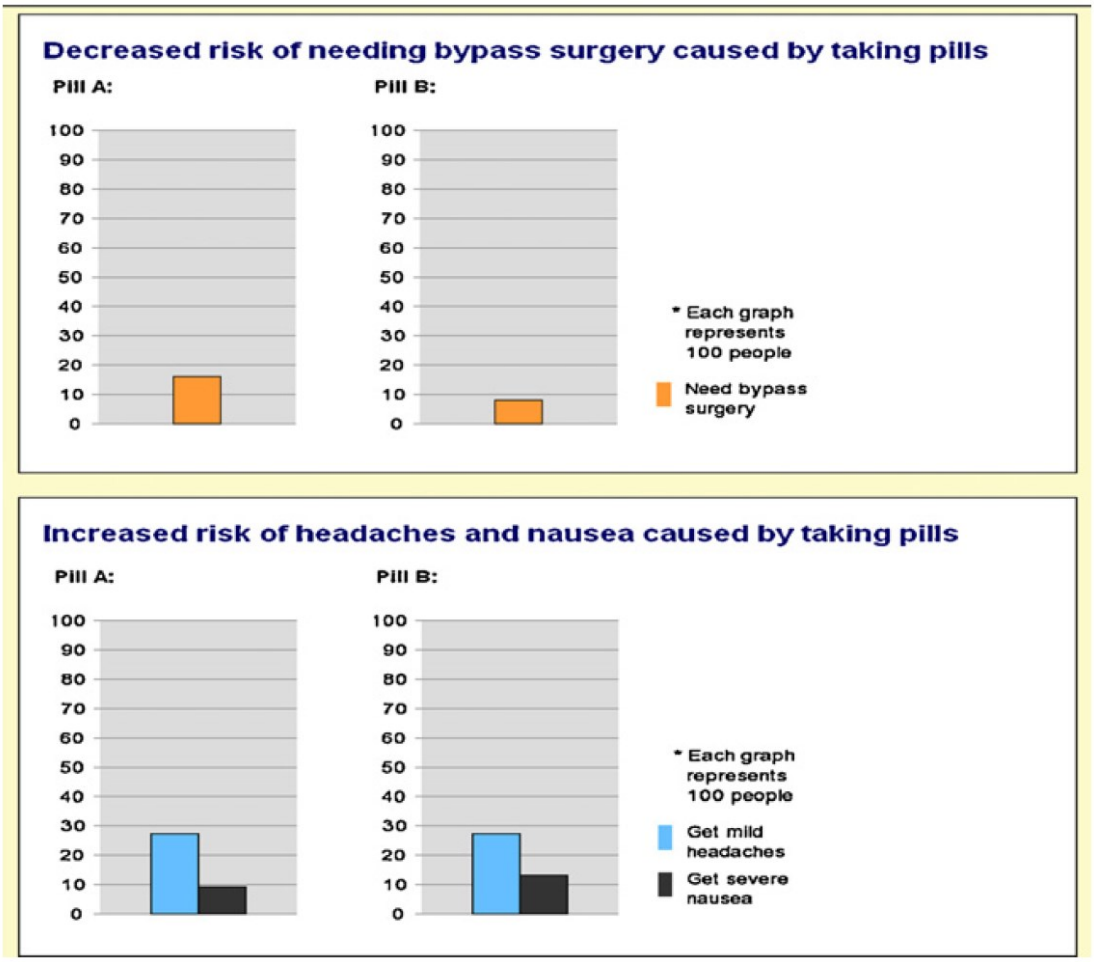


Figure 10. An example of a graphic supporting a decision between multiple treatment options (from [26]).

Combination

Decision aids supporting the combination of treatment options are uncommon in the literature. This may be due to the fact that binary decision aids are much simpler to develop and when treatments can be combined to even greater benefit, it seems unnecessary to spend time explaining how their interaction will magnify their effects. But, they can be particularly effective in communicating the effects of lifestyle changes on long-term conditions and helping patients determine their priorities.

Jones et al. [13] report on work where 90% of patients in a clinic environment were able to make lifestyle adjustments to address their risk of cardiovascular disease with the help of a model-based decision aid (see figure 11). Other systems have been successfully built to recommend lifestyle options to patients at risk of cardiovascular disease that are in agreement with clinical guidelines and practices [46].

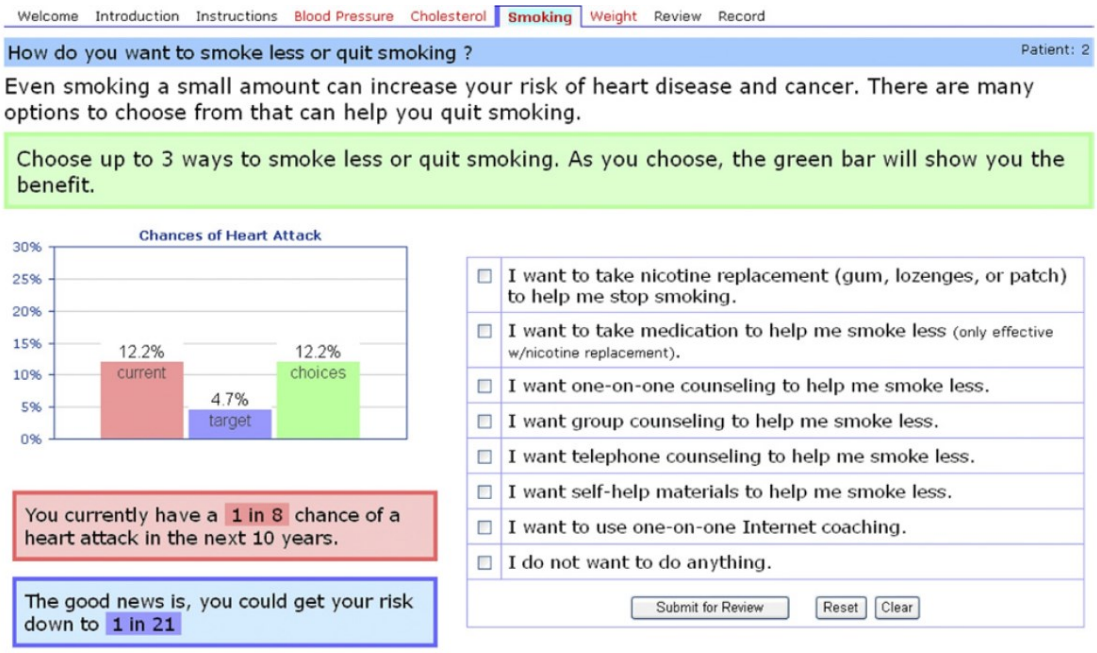


Figure 11. An example decision aid which supports exploring a combination of lifestyle and treatment options from [13].

Continuous Management

Finally, decisions aids supporting continuous health management are often used in situations where patients have chronic conditions necessitating continuous management. Continuous management decisions differ from other types of decisions in that each single decision has little impact on a patient's overall condition. However, over the long term of managing a patient's health, the cumulative effect of

these smaller, continuous management decisions has a determining impact on the outcome of a patient's condition. In these cases, critical decisions have been made and a patient has developed a care plan but must now incorporate those decisions into daily life. For example, Chaudry et al. [29] report on a graphics based aid designed to help low-literacy diabetic patients make healthier food choices. Others have addressed the needs of informing cardio-vascular patients in learning medication side effects and understanding quality of life factors [32]. Adherence to clinical guidelines has been found to increase when nurse practitioners are assisted with decision aids for patients managing obesity-related health conditions [30]. The intention of diabetic patients to adhere to their care plans increased while they interacted with game-like decision aids [47].

3.2 Timescale

Decision aids often include an explicit time frame in which the outcomes they communicate are expected to occur. Because the majority of decision aids are non-interactive, these timescales typically do not vary and provide a single snapshot of a patient's risk in an unchanging way. The impact of timescales on the effect of decision aids is an understudied space in the literature.

Single Projected Point

In decision aids supporting a single projected data point, a patient's risk is provided for a single point in time regardless of factors that may change during the span of time between the current time and the projected point (see figure 12). For example, a

cancer risk calculator may provide a patient's risk of developing pancreatic cancer within the next 10 years without showing the effects of lifestyle changes on that risk. There is some evidence to suggest that short timeframes are best for achieving risk reduction through behavior change [48]. Many decision aids or risk communications do not supply an explicit time frame for their data [6]. For example, the National Center for Chronic Disease Prevention and Health Promotion reports that reducing blood pressure reduces the risk of major cardiovascular events by 50% [49]. While useful to know, this gives a patient no indication of the immediacy of the risk reduction. Instead, it is provided as a data point at an implicitly singular future point of time.

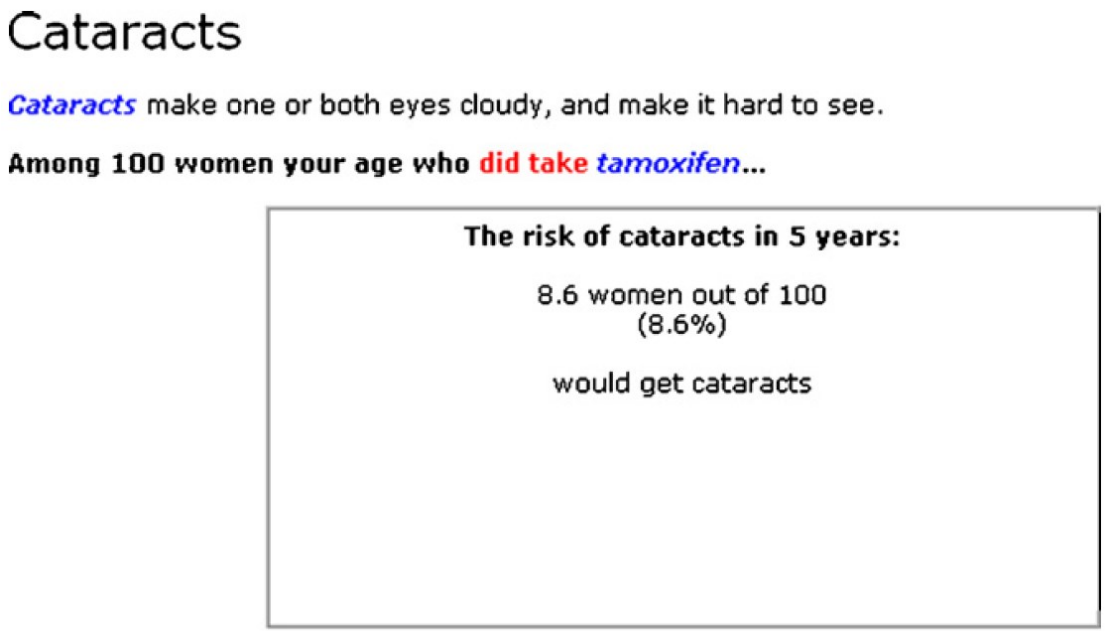


Figure 12. An example decision aid which reports statistics at a single projected point from [50].

Multiple Projected Points

Decision aids supplying multiple projected points of time present a patient's risk at several distinct points of time, usually chosen to be equidistant from each other or

represented as a continuum. For example, a patient's risk of cardiovascular disease might be presented as a table with rows reporting the patient's risk at 5 years, 10 years, and 15 years form the current date (see figure 13). Continuous measurements, such as mortality rates, are often represented graphically as curves or lines with the cumulative measure expressed on the vertical axis and time represented along the horizontal axis. Most often, multiple projected points are used to communicate the changing of treatment effects or the evolution of a patients risk over time should no actions be taken [20] [10].

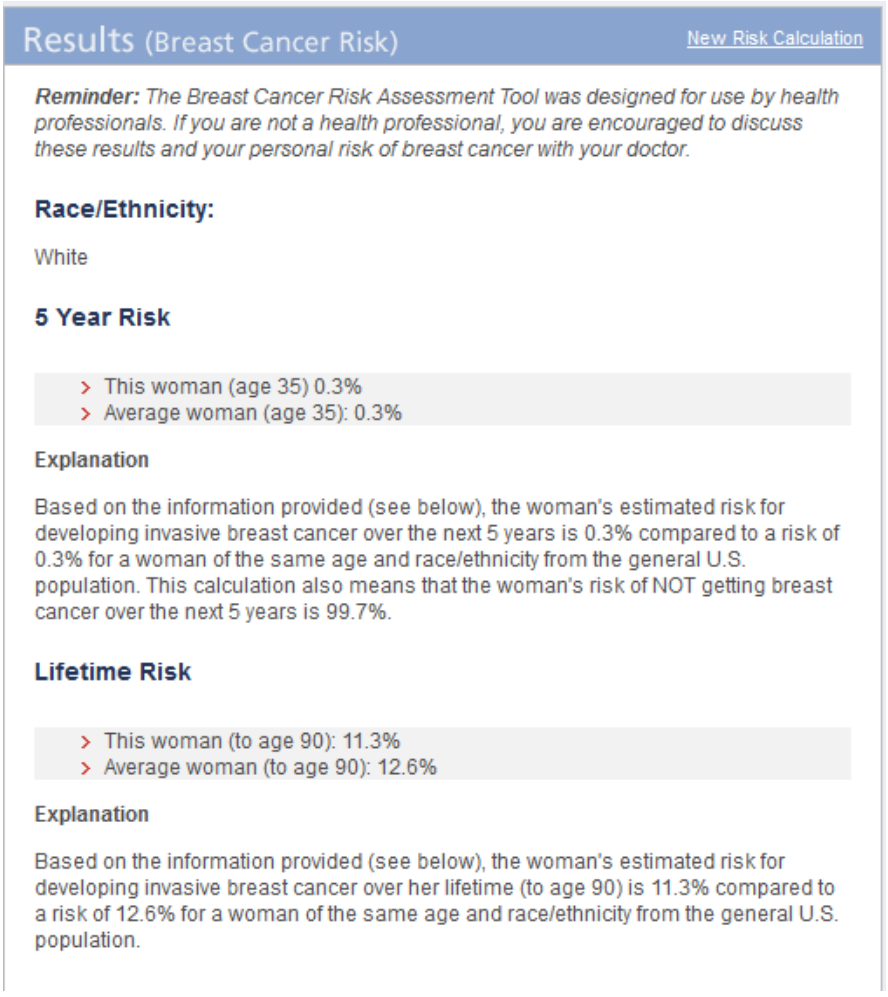


Figure 13. An example risk calculator which reports statistics from two projected points of time (from [51]).

3.3 Measurement Types

A decision aid is designed to convey information to patients in order to help them chose their best treatment options. In order to do this, a variety of measurements may be provided so that patients may determine which factors are important for themselves. Measures are provided without priority and it is often the task of the patient to choose which are relevant based on personal preferences (see literature on patient preference elicitation). Evidence-based medical practice involves the consideration of a great many points of data and each is unique to a patient's personal condition. This dimension is included in the model as an indication of the design complexity of a decision aid. With all other dimensions identical, a decision aid needing to reflect numerous salient measures will be more difficult to design than a decision aid detailing a single measurement, say the risk of experiencing headaches as a side effect of an oral medication. Some common measurements such as numbers-needed-to-treat are found to be easily misinterpreted by patient and physician alike [11]. Verbal expressions of risk and other measurements are known to have a wide degree of interpretation between physicians and patients [5]. Measurement presentation has also been shown to have an effect on interpretation with measures shown in familiar formats such as bar charts misinterpreted [27].

Survival/Mortality Rates

For a great many conditions, a patient's choice of treatment can be the result of a single measurement expressing the risk of dying from their given condition or their risk of their condition worsening significantly. For many patients, a treatment is only

as successful as its resulting survival or mortality rate. Framing effects have been consistently found when equivalent measures are studied. Patient treatment preferences have been shown to vary based on whether patients were presented with survival rates or median survival times [1]. Supplying comparative measures such as the risk of an average population has been shown to increase the likelihood that a patient opts for treatment if the patient's risk is above average [42].

Lifestyle Impacts

The lifestyle impact of a treatment is often overlooked within a decision aid. Side effects may be mentioned, but depending on the source of the decision aid the incidence of a given side effect may be unreported or expressed in ambiguous terms such as "very low." Depending on the severity of the original condition, adverse side effects may be considered irrelevant and a treatment choice must be made despite them. For relatively minor conditions however, the risk of a moderately impactful side effect such as fatigue or headaches might be a patient's deciding factor between opting for treatment or not.

3.4 Data Source

The information presented in a decision aid should come from credible and verifiable sources. The widespread use of the internet by patients in their research puts them in danger of encountering inaccurate and unsupported information. Waters et al. [6] report that only 53% of cancer risk assessments found online provide information about the statistical model or the peer-reviewed literature that was used to calculate

the risk estimate. In order to reach the best supported healthcare decision that meets their needs, patients must have reliable information. In this dimensional model, three sources of information are accepted.

Literature Summaries

Summaries of medical literature may be the easiest data sources to obtain and redistribute making them a popular choice for data for any decision aid. Scientific studies have the advantage of publications which make providing reference information to patients simple. Clinical trials and other forms of scientific study are often summarized and the findings relevant to a particular patient are distilled into the content of a decision aid. Educated patients are free to follow references to the original trials and studies for exact information. While they are relatively easy to produce, literature summaries are time consuming and only as accurate as the reports they summarize. They are susceptible to bias if the selected research is not balanced or flawed. They are rarely personalized to any particular patient and are not easily updated. Authors of decision aids may also be forced to reconcile data from conflicting studies and risk imposing an interpretation of study data that is unsupported. However, literature summaries are an important information source for decision aids as the broad scope of information makes them an ideal starting point for patients who have little to no knowledge of a particular condition and need to begin their search for information.

Models

For some decision aids, a scientifically developed model accounts for one or more condition-specific parameters and classifies a particular patient's risks based on the patient's personal expression of model parameters (see figure 14). Breast cancer decision aids frequently make use of the Gail Model for patient-specific risk estimates [20] [21] [43]. Similarly, cardiovascular disease can be modeled by several systems including expert systems using ARIC data [46], the UKPDS risk engine [52] [48], and the Framingham Risk model [52] [13] [48]. An advantage of model-backed decision aids is that small changes can provide feedback to patients using interactive decision aids as demonstrated by Jones et al. [13] and others [28] [45]. Not all interactive decision aids invoke models, however. Ancker, Weber, and Kukafka [53] report on the use of interactive graphics for communicating a static value of risk to low-numeracy participants and base their drawings off a simulated literature summary.

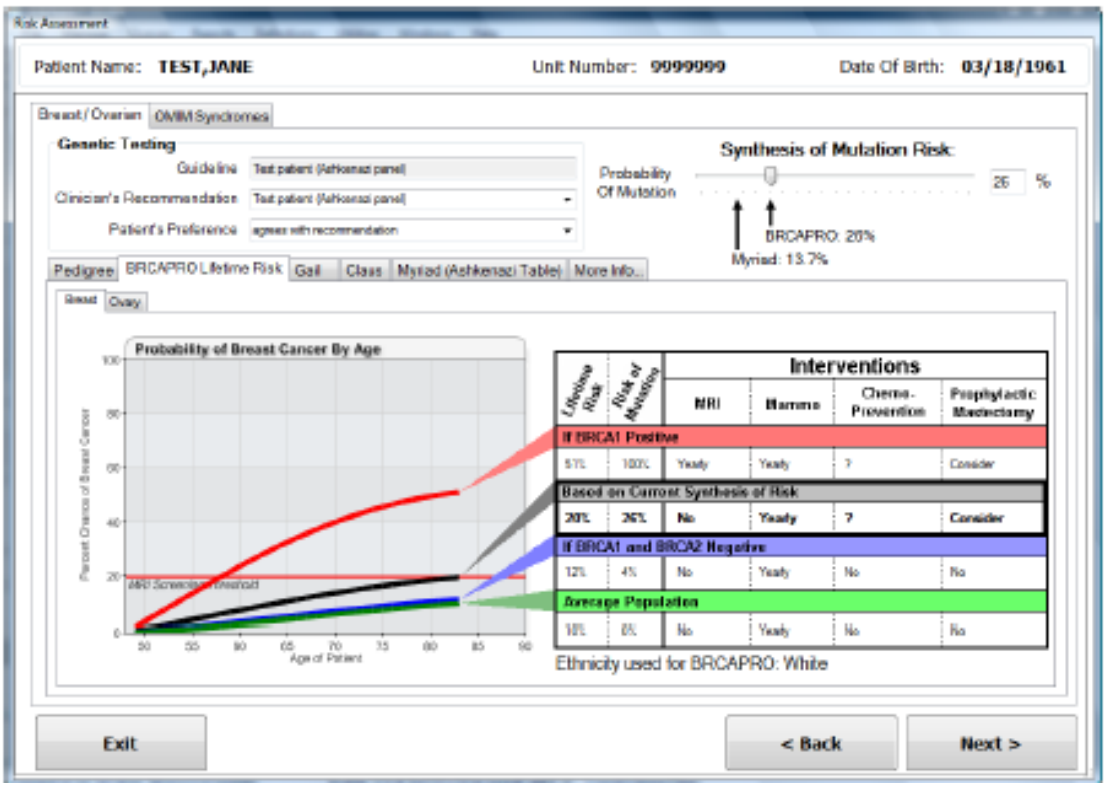


Figure 14. A model-based system for exploring patient breast cancer risk (from [54]).

Electronic Health Records

In large clinic settings, a patient's personal records may be compared against the records of other patients and measures reported in a decision aid are a reflection of the outcomes of other patients with similar health records as the given patient. For example, the Hughes riskApp uses EHR data to identify high risk hereditary breast and ovarian cancer patients and model their risk of developing cancer across their lifetime [55]. Kharrazi [47] reports on the development of an interactive system for children with diabetes which uses parent-input reports of compliance as records to drive an in-game reward system. EHRs have the potential to provide the same interactivity as scientific models which make them a promising source of data for decision aids provided a large enough number of records are available for use. With

too few records, a decision aid would suffer from the same lack of reliability as clinical trial with a small or insignificant sample size. In this regard, a scientific model may be needed to estimate data when real accounts are not available. With access to enough records, however, EHR based decision aids will be able to account for a fine level of personalization and capture subtleties of a condition that even a finely tuned model may have to abstract away. Health records may also be patient-input or reported as in the case of the online health community Patients Like Me [16] (see figure 15).

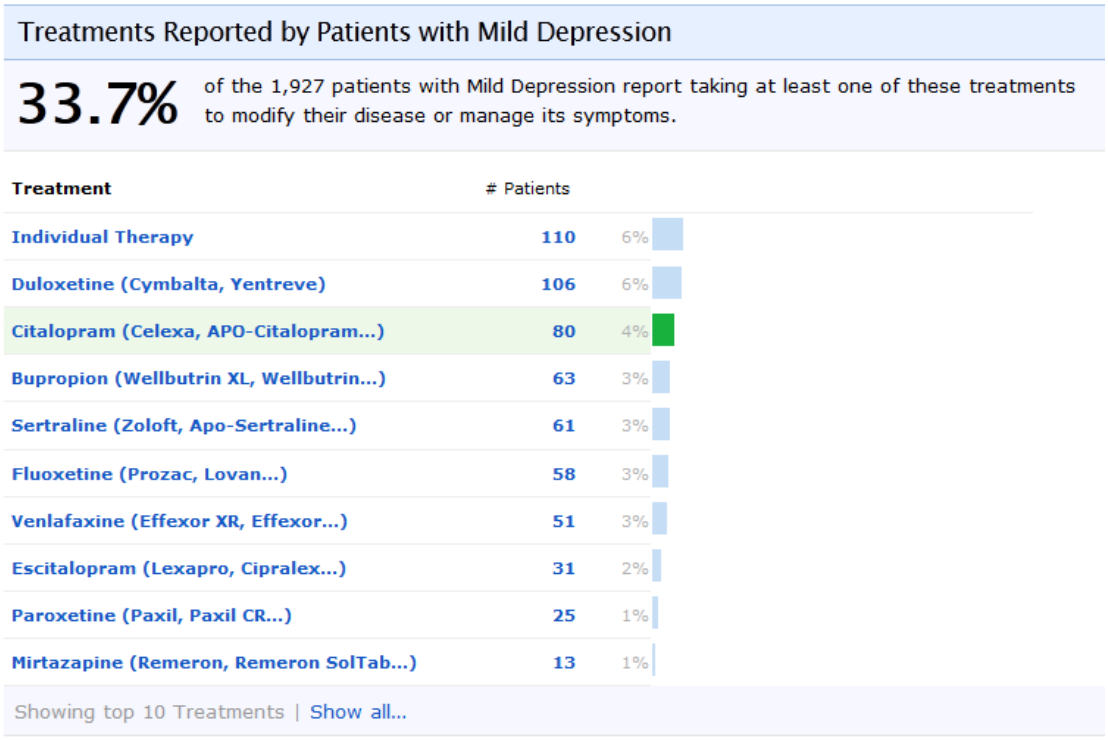


Figure 15. Screen-capture of information from Patients Like Me [16] which draws its data from patient-input health records.

3.5 Personalization Level

A decision aid's relevance is related to its ability to capture a patient's unique circumstances which is significant as the increased relevancy of a decision aid has been shown to make health communications more effective [56]. The level of personalization supported by a decision aid can range from a series of predetermined options a patient may select from to the capacity to include a patient's entire medical history. This dimension is a measure of the personal relevance of the presented information in a decision aid. With all other dimensions identical, a decision aid personalized with a detailed medical history of a particular patient will be more relevant to the patient than a decision aid customized on a subset of that patient's history. In addition to personalizing a decision aid to a patient's medical condition, decision aids may be personalized to reflect a patient's preferences for treatment. Treatments a patient might deem unacceptable for any number of reasons might be excluded from consideration entirely. It could be argued that even a collection of entirely generic decision aids is actually personalized as it reflects the options the patient who collected the information is willing to consider.

A patient's specific risk of a condition is the most common level of personalization. This patient-specific risk can be derived from a wide range of factors such as height and weight [30] [46], medications [32] [52], diet plans [32], exercise regimens [32] [28], treatment preferences [1] [15] [31] [46], and other lifestyle factors [46] [52] [13] [21]. In some cases, these personalizing data are gathered automatically from a patient's electronic health record [13] [47] but are most

commonly entered either by the patient from their own knowledge or with a medical professional's assistance.

3.6 Information Format

The most widely studied dimension within this model of decision aid content is the information format of a decision aid. With the utilization of the internet, a broad range of media has become available for the communication of patient treatment options. Rather than attempting to capture all the possible media now used for the production of decision aids, this dimension will focus on broad categories of presentation. Decision aids may belong to one of several categories summarized in Table 2. The same data may be presented in multiple formats and a number of decision aids may be designed to support patients making the same decision. A patient may be given multiple decision aids of differing formats in order to make the information as clear as possible or to leverage the advantages of some formats over others. The classifications within this dimensional model refer to the format of a single decision aid as a stand-alone product.

Table 2. Information Formats Used by Decision Aids

Format	Description	Example
Text	Treatments, risks and outcomes are expressed in written formats without graphics or augmentation.	A patient is given a report to use in determining whether or not to undergo a medical procedure.
Graphics	Treatments, risks, and outcomes are expressed in a graphical format such as bar charts, pie charts, pictographs, etc.	A patient is given a booklet of infographics which portray the risks of experiencing side effects of a possible medication as a series of bar charts. Each chart represents a side effect and each bar of each chart represents a year of treatment.
Text + Graphics	A decision aid contains both text (as described above) and graphics (as described above)	A patient is given a report with particularly salient study results called out in a table and risks communicated through pictographs.
Animation	A decision aid uses graphics which are animated to reflect changing measures or guide patients in understanding one graph's relation to another.	A patient is given a video which narrates a smoker's cumulative risk of acquiring lung cancer as a series of pictographs. Each step in the animation alters the data in the pictograph by one year at a time.
Interactive	A combination of text, graphics, or animation is available which patients may manipulate through a series of controls and observe the effects on relevant data.	A patient's risk of cardiovascular disease is assessed based on current lifestyle factors and then the patient selects a number of lifestyle adjustments to observe how those adjustments affect projected risks.

Text

Text-only decision aids are commonly found in research literature where they are very often used as a control condition in a randomized trial. There is evidence to suggest that patients prefer other formats, particularly those that provide immediate feedback on questions [31]. This may be due to a difficulty in interpreting statistics

which has been shown to hinder both patient and physician alike [11]. Some work has even found that numeric text alone produces low knowledge in comparison to pictographs [50]. Other studies have suggested that the believability of data was perceived as greater in decision aids which contained graphics instead of just text [28] and that risk presented as text-only data is often overestimated [20].

Graphics

The study of graphics as decision aids has provided evidence that features which support the accurate or correct interpretation of data are different from those that prompt behavior modifications [7] (see figure 16 for commonly studied graphics). Numeracy and graphicacy have been repeatedly shown to affect the accuracy of patient understanding [4]. Pictographs have been shown to help patients attain higher risk comprehension, particularly those with low-numeracy [26]. Factors such as horizontal layout and shading have been investigated for their impact on graphic understanding [40]. Familiar graphics are often preferred based on qualitative reports but can also lead to less accurate knowledge [44]. There is danger in applying unfamiliar graphics such as funnel plots which allow patients to apply their own, possibly incorrect, interpretation to data [27].

The combination of text and graphics in a decision aid has some mixed results. Participants have reported information overload when risk communicated through graphics is augmented with additional information as text [44]. Other work has suggested that tables, which combine text with graphical layouts, are associated

with higher verbatim knowledge in patients but at the cost of lower gist knowledge [26].

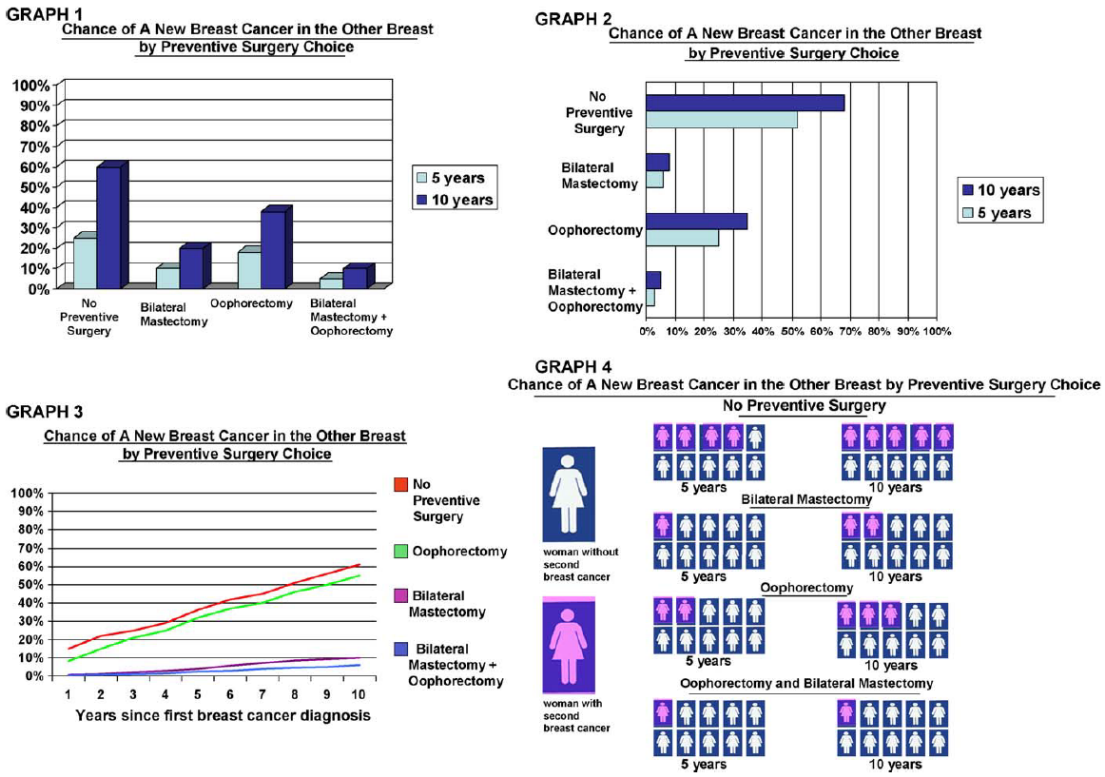


Figure 16. Multiple graphic formats evaluated by Brown et al. in [4]

Animation and Interaction

Animated and interactive decision aids are less prevalent in research literature but their effect has been encouraging. Low-numeracy participants have been found to report higher risk-feelings than high-numeracy patients except when using interactive graphics such as figure 17 [53]. Uncertainty of cancer risk was effectively communicated through a dynamic visual format by Han et al. in [25]. Interactive decision aids have also consistently led to more expressed emotional responses including relief about small risks, concern over large risks, and feelings of

empowerment [45]. Participants not making use of interactive features report lower intentions to make lifestyle changes or adhere to care guidelines when compared with participants in interactive conditions [47] [13]. One barrier to the adoption of interactive decision aids is studied by Xie, Watkins, and Huang [41] who indicate that the controls used by interactive decision aids are frequently non-intuitive to target populations such as older adults. Some evidence has also suggested that interactivity can distract patients from understanding relevant information [57].

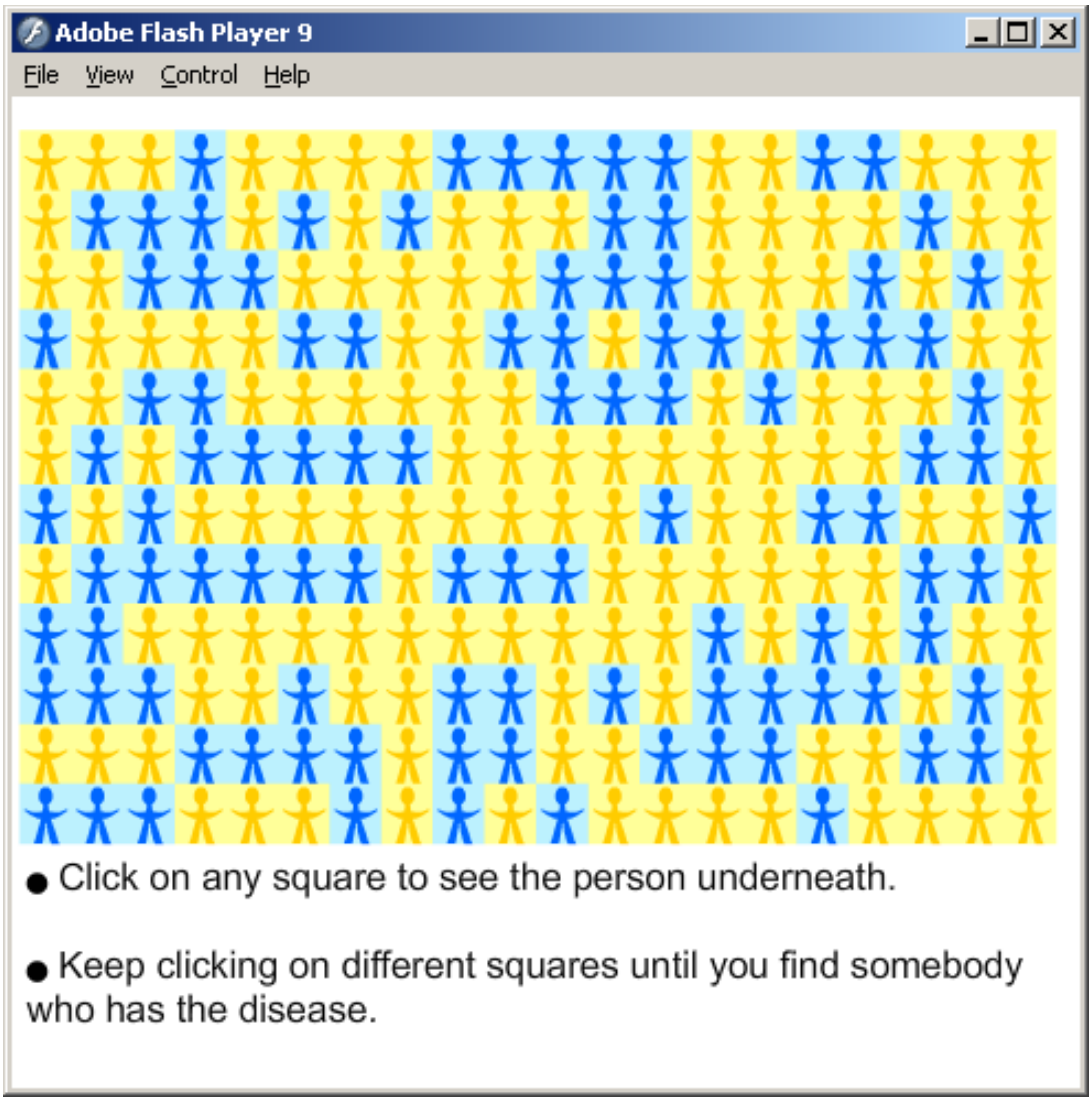


Figure 17. An Interactive decision aid designed to communicate randomness in health risks [53].

3.7 Multi-Dimensional Model Summary

The content of medical decision aids can be described through six independent dimensions: decision type, timescale, measurement types, data source, personalization level, and format. The most common combination of dimensions in published research is that of binary, single projected point of time, single measurement decision aids derived from literature summaries and without personalization presented in a text or basic graph format. This prevalence of simple decision aids does little to address the information needs of patients who may have more complicated decisions to make over the course of longer timescales or when multiple data points need considering. Electronic health records are becoming more available as data sources which could enable a greater degree of personalization for patients which in turn, would increase the relevancy of decision aid content. Finally, animation and interactivity is grossly understudied. The internet provides a platform which supports interactivity and with more patients turning to the internet to search for information, research is needed to understand how such interactivity and animation can be used to better support patients and their information needs.

Chapter 4: TreatmentExplorer Design and Expert Review

After analysis of the design space and the gaps exposed by the Multi-Dimensional Model several observations were made. It became clear that decision aids capable of supporting the exploration of multiple treatment options were lacking. Further, many decision aids did not address issues of patient risks and outcomes over time. Single measurements were the most common and side effects of treatments were rarely quantified when mentioned. Much of the research reviewed did not discuss effects of personalized decision aids apart from the consensus that personalized decision aids were more relevant to patients. These observations inspired the TreatmentExplorer prototype (demo and code can be found at http://www.cs.umd.edu/hcil/treatmentexplorer/).

4.1 Design Goals

As discussed in Chapter 1, the challenges faced by patients choosing a treatment option present several design goals which include:

1. Multiple-Treatment Options: the prototype should support patients comparing multiple treatment options.
2. Multi-Measurements: the prototype should go beyond representing only a treatment's success rate and express other details about a treatment such as its popularity and side effects.

3. Multiple-Projected Points of Time: the prototype should capture the changes in patient outcomes that can be expected over the entire course of treatment, rather than at a single sampled point of time.
4. Personalized Outcomes: the prototype should represent personalized outcomes that are most likely to be experienced by a given patient. This includes leveraging the outcomes from other patients to make the decision aid reflect the most realistic outcomes possible.

4.2 Early Prototypes

Early designs for the TreatmentExplorer prototype were more ambitious in the scope of their visualizations. In particular, they attempted to show information for all available treatments at once. Ultimately, simpler visualizations depicting details for a single treatment where chosen to avoid risking information overload.

Basic Graphics

A clear starting point for TreatmentExplorer involved the use of several basic graphs commonly used in all forms of media that patients would already have exposure to and be familiar with. There is extensive research to support the efficacy of certain basic graphs in isolation and in combination with text-based content. A partial prototype is shown in figure 18.

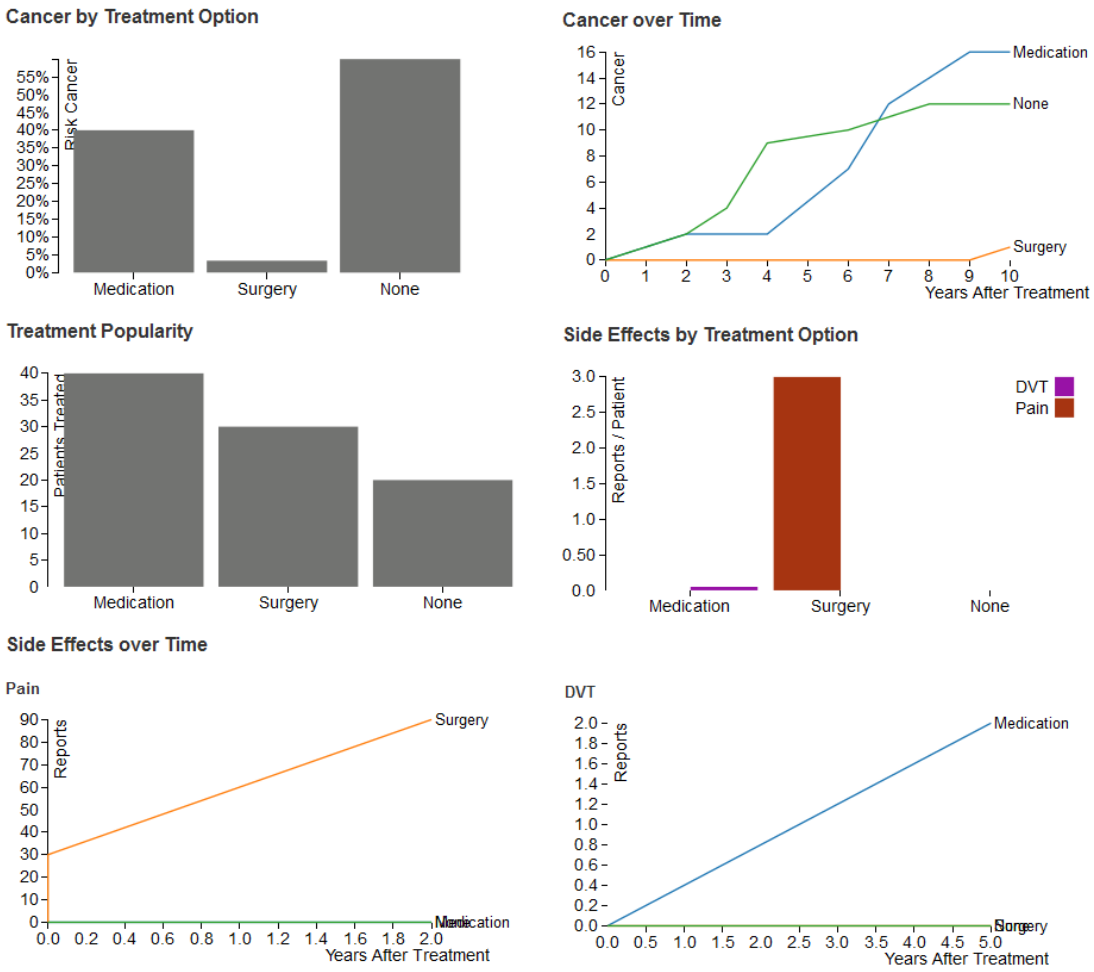


Figure 18. A partial, early decision aid prototype using multiple basic graphs to communicate cancer risk and side effects for three treatment options.

Figure 18 shows six basic graphs, both bar and line, which communicates the risk of cancer for three hypothetical preventative cancer treatments and the associated side effects. These graph choices have been studied in risk communication literature for their ability to convey varying measurements to the average patient. This partial prototype illustrates several design challenges, however.

To begin with, multiple graphs are needed to convey multiple measurements. Bar graphs are ideal for conveying the difference in magnitude of a single measurement between multiple options. They are therefore appropriate for showing

measurements such as treatment popularity and a single risk calculation for a single point of time. A patient could compare these two single measurements across three treatments easily. However, bar graphs are not generally able to convey time as well as other types of graphs (such as line graphs). To communicate how cancer rates vary across time, another graph such as a line chart or augmentation with other visual embellishments such as animation is needed. Line charts are excellent for showing time trends of a single measurement across time. But multiple line charts are required to show the trend in onset of cancer and side effects across time. Using a single line chart would impose a great deal of clutter and confusing multiple instances of a treatment in a single graph.

A second design challenge is that it is difficult to obtain a clear understanding of a single treatment as a whole. This series of graphs is designed to support comparing multiple treatments against each other one metric at a time. The details of a single treatment are spread to multiple locations. A patient using such a decision aid must assemble various pieces of information from many places to understand a treatment as a whole.

All Treatments

Initially, all information about all treatments was available to a patient in a single visualization. Comparisons could be made across treatment representations and coordinated interaction would highlight the same data points across all treatments at once so patients could easily find a data point's representation. A partial prototype of this design is shown in figure 19.

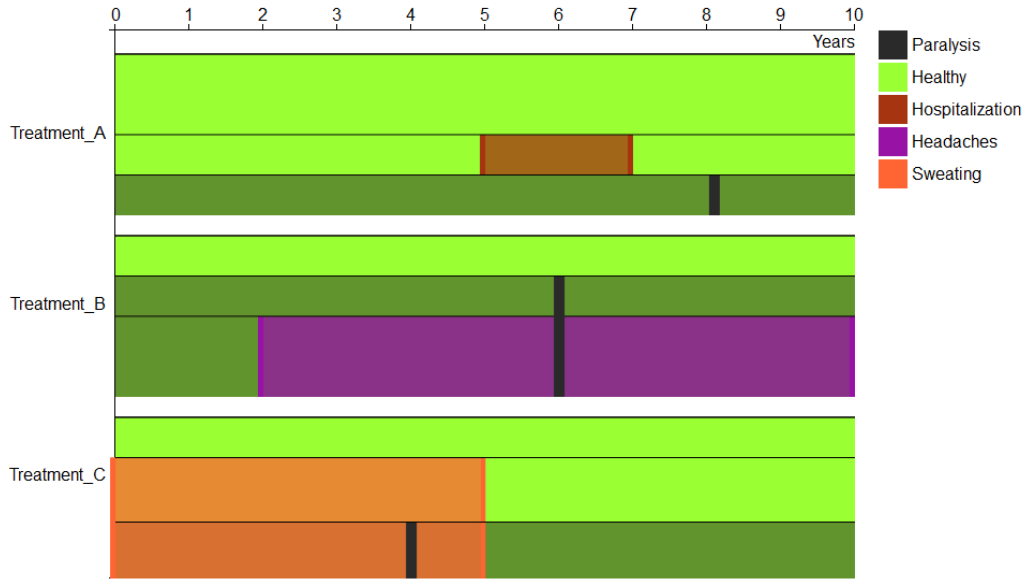


Figure 19. An early prototype displayed all treatments simultaneously for rapid comparison across treatments at a glance.

This prototype has the design benefit of enabling comparison across treatments rapidly as well as providing a complete description of a single treatment which overcomes the two drawbacks of the basic-graphs prototype.

This partial prototype is not without its own drawbacks, however. Because all treatments are represented in the visualization at once, there is little room for treatment details. For example, the vertical sizing of each treatment bar represents the relative popularity of the treatment. However, because of the vertical stacking of each treatment this is difficult to compare when all treatments are similarly popular. Further, there is no room for a vertical axis to report exact numbers of patients who experience a condition or side effects for each treatment.

A second drawback to this design is the real potential for information overload. Research has shown that many patients face numeracy and graphicacy deficiencies [4] [9]. With all available information presented at once, it could easily

be difficult for a patient to not know where to start reading this novel visualization. Vertical sizing shares some properties with bar charts; however the additional horizontal representation of time could confuse patients. Interactivity could help by illuminating and reinforcing which measurements are related across treatments but would still not be enough to help an overwhelmed patient know where to focus first.

4.3 Description of TreatmentExplorer

The final TreatmentExplorer prototype reuses the treatment representation from the All Treatments prototype but scales down the visualization such that a single treatment is displayed at a single time (see figure 20). The vertical and horizontal axis can thus be repurposed to display treatment-specific details such as the exact number of patients choosing a treatment and the timescale most relevant to the treatment. Animated navigation between treatments preserves the ability to compare treatments even though the information of only one treatment is displayed at a time. The TreatmentExplorer prototype thus fits into the following dimensions of the multi-dimensional model discussed in Chapter 3.

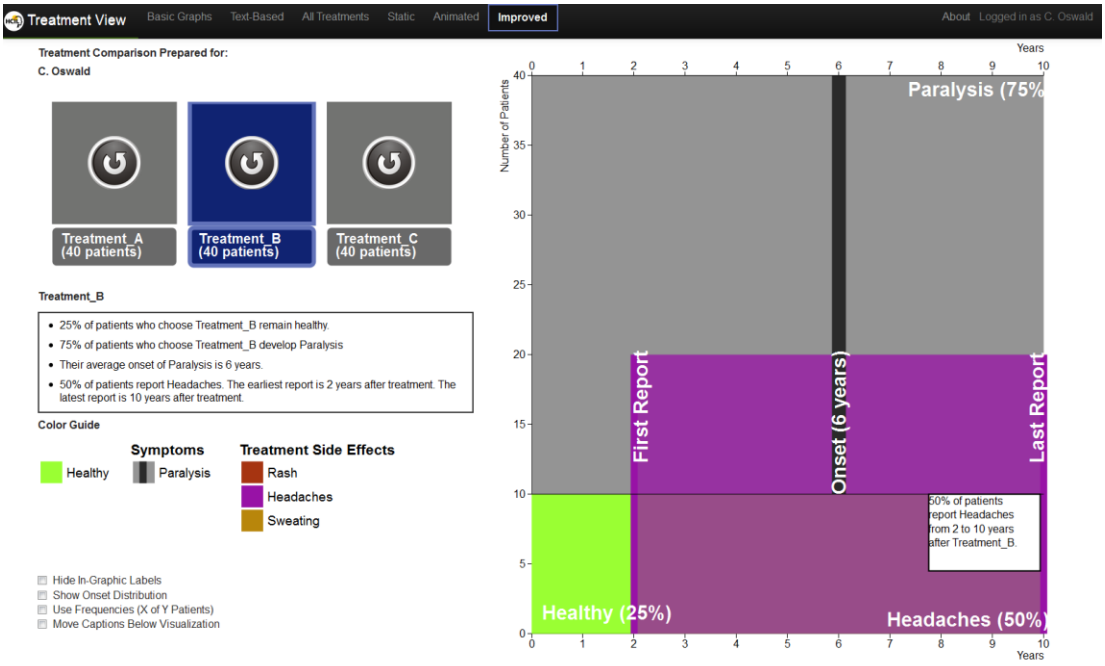


Figure 20. The final TreatmentExplorer prototype includes multiple measurements for each of multiple treatments across multiple points of time. Interaction and guided narration support patient understanding of treatment information.

Electronic Health Record Data-Sourced

The TreatmentExplorer prototype was developed using synthetic Electronic Health Records in a format similar to the records used in EventFlow [58]. By opting for EHRs as a data source, the TreatmentExplorer prototype can be easily updated to reflect the most current data relevant to a condition and patient. It also enables TreatmentExplorer to support a great variety of conditions. Deployed in a clinic environment, a fully-operational system would be useful to all patients. For ease of development, hand-built synthetic datasets were used and designed to exhibit specific characteristics such as varying treatment popularity, varying reports of side effects, varying onset of symptoms, and varying success rates. These synthetic datasets were

either based on published statistics about the prevalence of certain conditions (such as breast cancer [59]) or completely fictitious.

Patient-Specific Personalization

Because EHRs are used to drive the content of TreatmentExplorer, a physician/patient team can select only the most relevant patient records to visualize. A search interface would allow for the custom input of patient age, gender, and vital information. Alternatively, direct input from a patient's own health records could be used and similar patients found through algorithms such as those in [38]. The resulting treatment visualizations will thus display only the most relevant outcomes to a patient.

Multi-Treatment Option

TreatmentExplorer supports the visualization and comparison of multiple treatment options. This is a relatively unexplored part of the design space as revealed by the dimensional model discussed in Chapter 3. By supporting multiple treatment options and their comparison, patients using TreatmentExplorer are relieved of the cognitive burden of assimilating data from multiple different decision aids. Patients are able to explore the relative differences between treatments as well as learn the details of individual treatments as a whole. Navigation between treatments is enabled via a series of buttons in a graphical navigation panel as shown in figure 21.

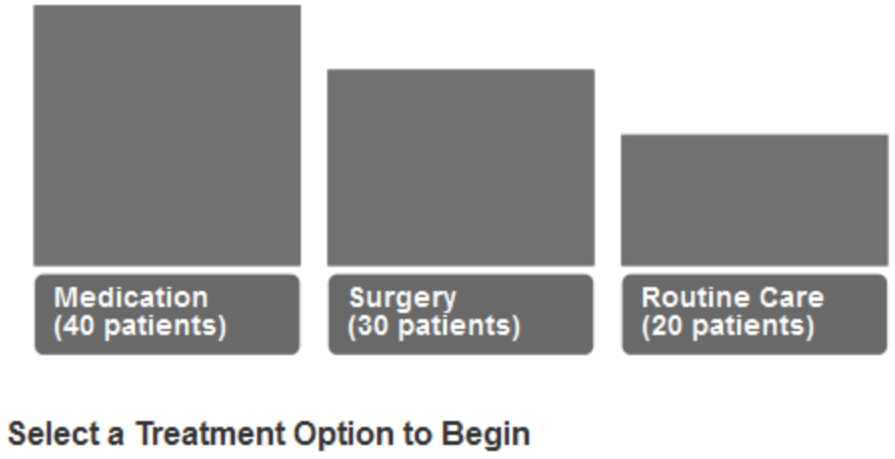


Figure 21. Navigation in TreatmentExplorer is enabled through buttons. The navigation panel also displays information about the relative popularity of each treatment.

Initially, when the TreatmentExplorer prototype is loaded in a web browser, neither the visualization nor the treatment specific highlights are visible. Patients are instead prompted with a mostly-empty page to select from one of their treatment options. This overcomes one of the design challenges of the previous prototype by providing patients with a clear starting point for using the visualization. Selection of a treatment is shown with a selection color and a patient's exploration of a treatment marked by an animation-replay button once the patient has viewed the narrative for a selected treatment as shown in figure 22. Additionally, the navigation panel graphically presents the relative popularity of treatments within the represented EHR data. The accompanying bar graph displays popularity on the vertical axis as well as the exact number of patients choosing that treatment in the button label. Animated transitions between treatments, both rapid and with guided narration, enable the comparison between treatments even though the details of one treatment are displayed at a time.

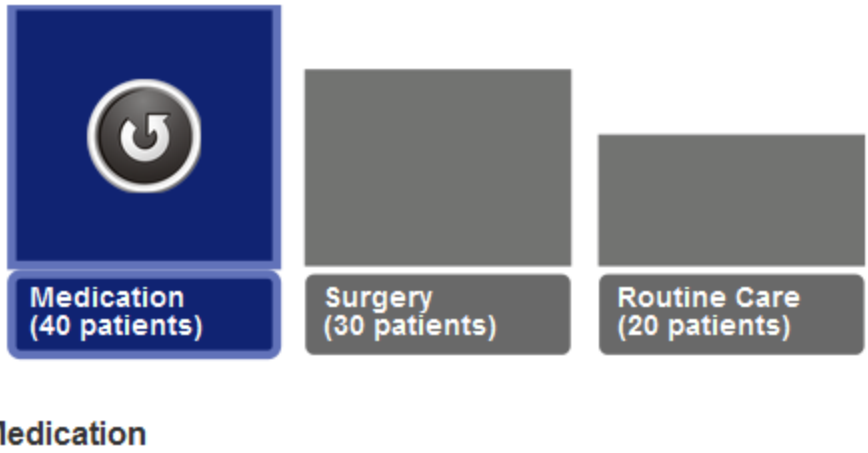


Figure 22. A selected treatment is visually distinguished from the other available treatments with color and labeling.

Multi-Measurement Display

TreatmentExplorer also represents multiple relevant measurements and details of a treatment which further explores the design space outlined in Chapter 3. The proportion of patients remaining healthy after treatment, experiencing a specific condition, and afflicted by side effects are all represented in the visualization. Additionally, the average onset time of a condition is available and 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. An example visualization is shown in figure 23.

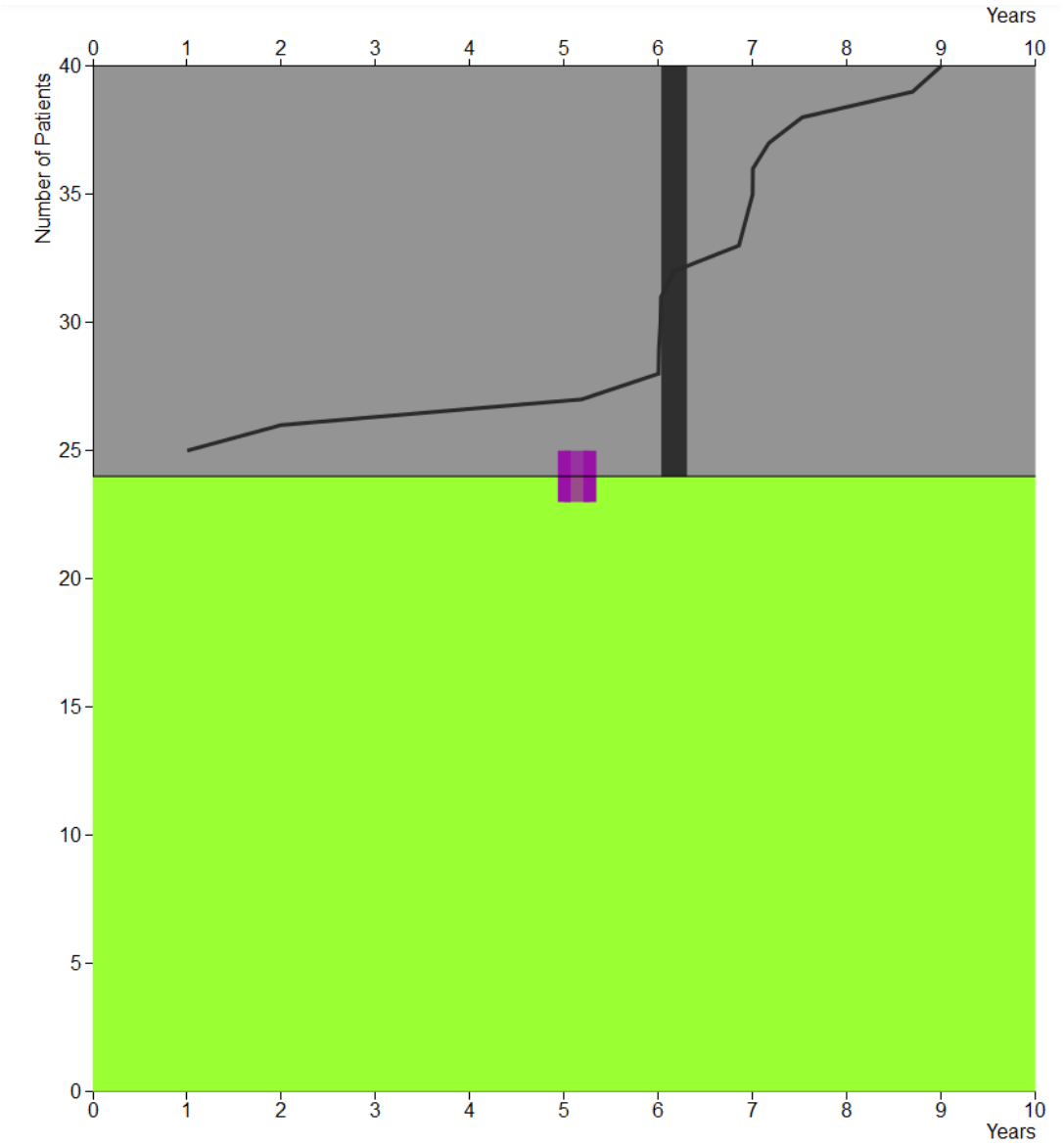


Figure 23 The TreatmentExplorer visualization displays the relative portion of healthy vs. condition-developing patients who have chosen a given treatment as well as the average time to a condition's onset and any side effects reported.

The number of patients choosing a given treatment is represented along the vertical axis of the visualization. Time is represented along the horizontal axis with duplicate axis labels to assist patients in making judgments about the timing of depicted events. A consistent, bright green color is used to represent the proportion of patients who remain healthy and do not develop a given condition at the end of the represented

time period. A grey color and accompanying dark grey solid bar represent the proportion of patients who eventually develop a condition and the average time at which they develop a condition. The exact cumulative distribution of condition onsets is available as a line overlay in this region as an advanced option. These measurements were selected to provide information relating directly to success and/or mortality rates of conditions and describe the overall efficacy of a treatment.

Side effects are also shown in the visualization and appear as overlays using an orange-purple color scale to distinguish them from the healthy and afflicted groups. Side effects are included in the visualization to provide information on potential lifestyle-affecting complications associated with a treatment. Because both healthy and afflicted patients may report side effects, side effects are positioned equally atop the healthy and non-healthy regions. Accompanying text highlights reinforce the measurements represented on the visualization and in-visualization text tips further support patients' interactive exploration of the represented data.

Multiple Points of Time

The horizontal axis of the TreatmentExplorer visualization is derived from the length of time captured by the EHRs. Each treatment is displayed using the same horizontal scale based on the length of the longest record available in the visualization. This allows patients to compare the anticipated time of condition onset as well as the reporting of side effects along a consistent axis between treatments. It also provides a consistent context for patients to understand how well outcomes of patient with their specific condition can be predicted and for how long.

Animated Graphics and Text with Guided Narration

One of the under-studied spaces identified by the dimensional model was that of interactive decision aids. Because the TreatmentExplorer visualization is a novel design, an animated, guided narration has been included to both teach patients how to interpret the visualization and to explain the treatment outcome data itself. This guided narration makes use of synchronized text captions and transition of visualization elements to explain each piece of the visualization (see figures 24 through 29). The narration begins when a patient selects a treatment to explore and includes the following steps:

1. A 'Start Screen' appears and explains that the visualization will update to show the outcomes of patients in the treatment group and how long the records of those patients extend into the future. (Figure 24)
2. The proportion of the visualization representing patients of the treatment group who are healthy appears and is sized appropriately. (Figure 25)
3. The proportion of the visualization representing patients who eventually develop the given condition appears and is sized appropriately. (Figure 26)
4. A solid bar appears in the region representing the patients who develop a condition. This bar shares the height of the region but is a fixed width. This bar represents the average onset of the condition and moves to the appropriate place along the horizontal axis. (Figure 27)
5. A uniquely colored, semi-opaque region representing a side effect of the given treatment appears. This region moves vertically such that half the region lies over the condition-afflicted patients and the other half lies over the healthy

patients. The width of the region is such that the left edge aligns with the point along the horizontal axis representing the first report of the side effect in the treatment group. The right edge of the region aligns with the date of the last report within the treatment group. (Figure 28)

6. A concluding screen confirms that the narration has finished and prompts the user to replay the same narration, explore the current treatment, or select a new treatment. (Figure 29)

While the visualization plays through its animations, the treatment highlight list is kept synchronized. The explanatory text reporting exact statistics begins with the exact statistics of the healthy patients (step 2, Figure 25) and provides accompanying captions through all side effects (step 5, Figure 28). Additionally, captions appear in the center of the visualization as the animation plays. The in-visualization captions display the same text as the treatment highlights and can be moved below the visualization through a toggle available below the treatment highlights list.

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 of the narration begins with the display of the text highlights and captions and the animation of the step begins after a brief delay. When the animation has completed for a step, a prompt appears to alert the patient that they may continue when they are ready. Figures 24 through 29 display this animation sequence.

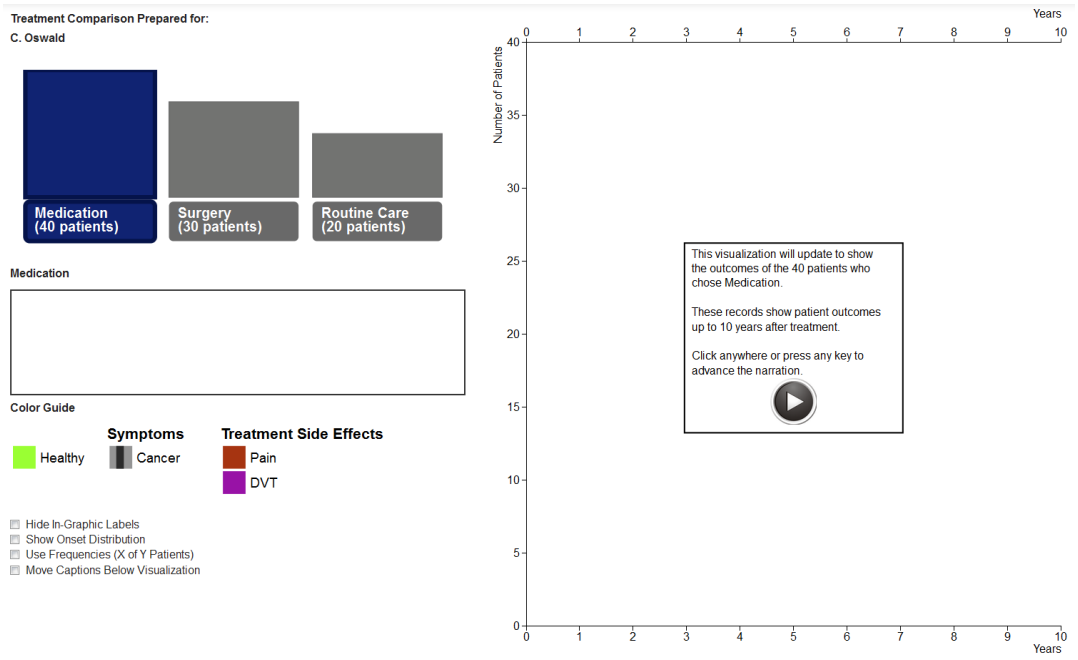


Figure 24. Animation Step 1: A patient selects a treatment from their options and is shown a start screen explaining the visualization update and describing the extent of the treatment data.

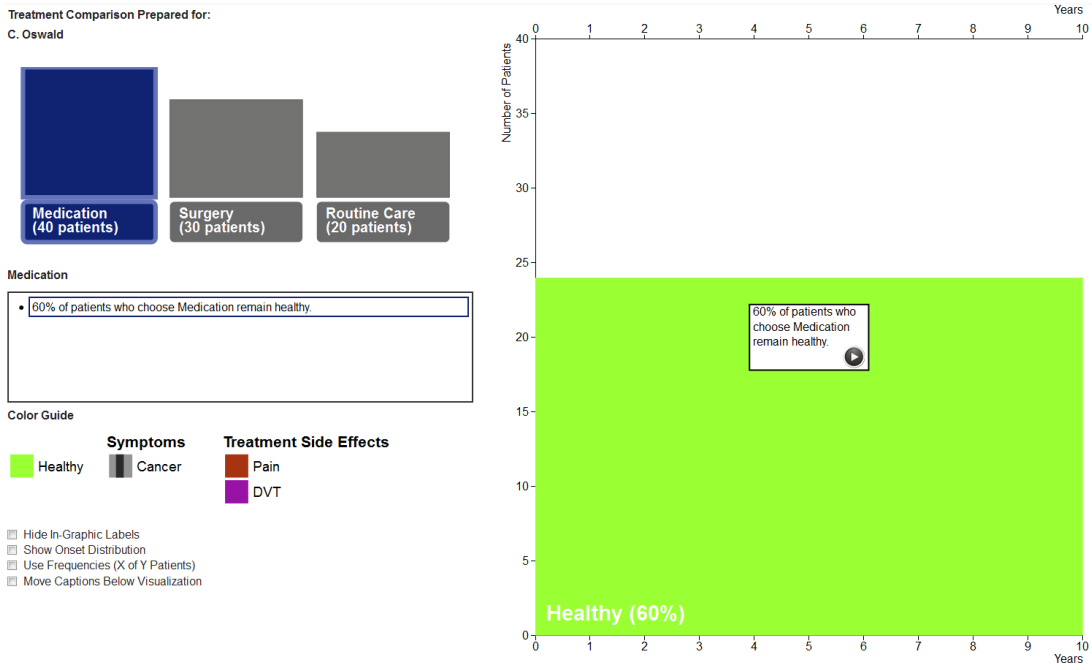


Figure 25. Animation Step 2: The proportion of patients who remain healthy after a treatment is animated onto the visualization with synchronized explanatory text.

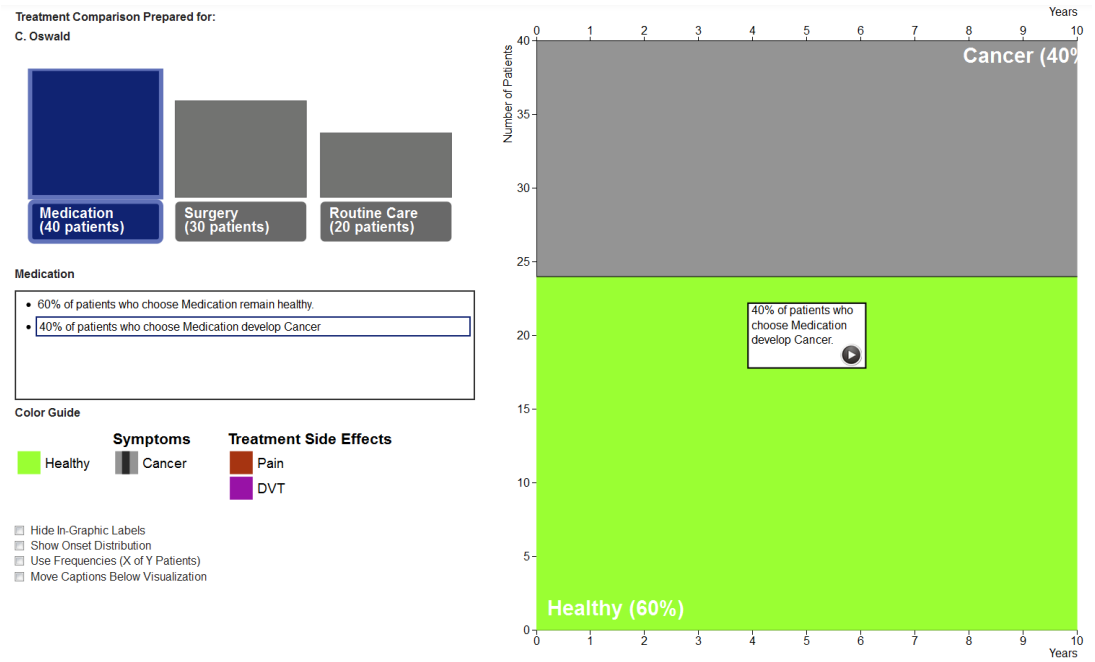


Figure 26. Animation Step 3: The proportion of patients who eventually develop a condition is animated onto the visualization with synchronized explanatory text.

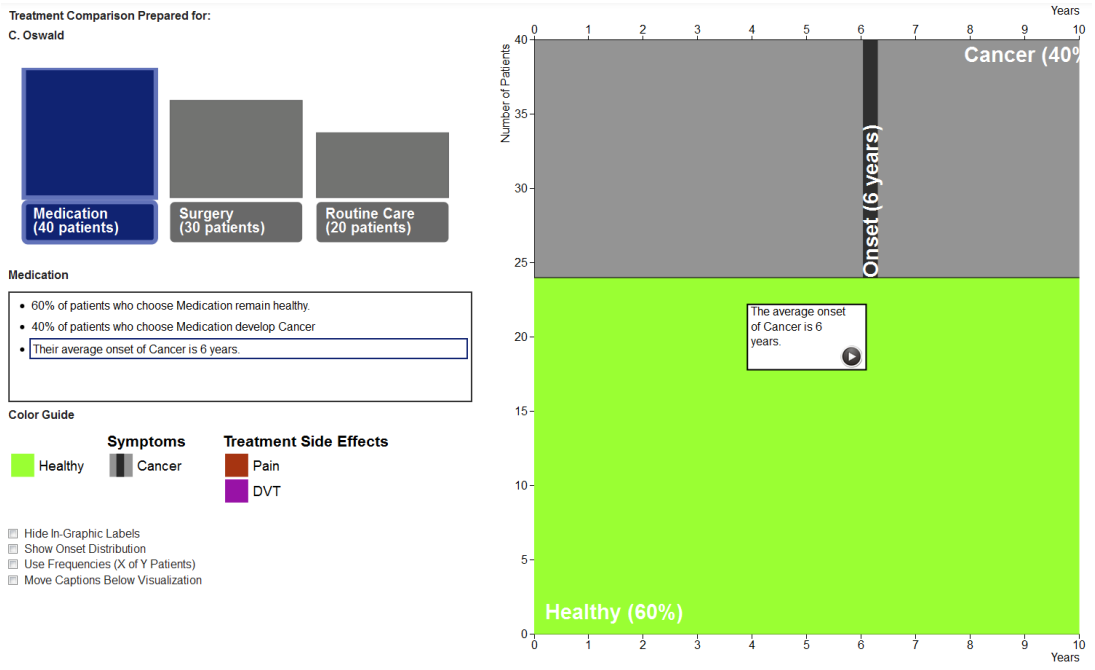


Figure 27. Animation Step 4: A solid bar representing the average onset of a condition appears at the left edge of the visualization and animates to the right with synchronized explanatory text.

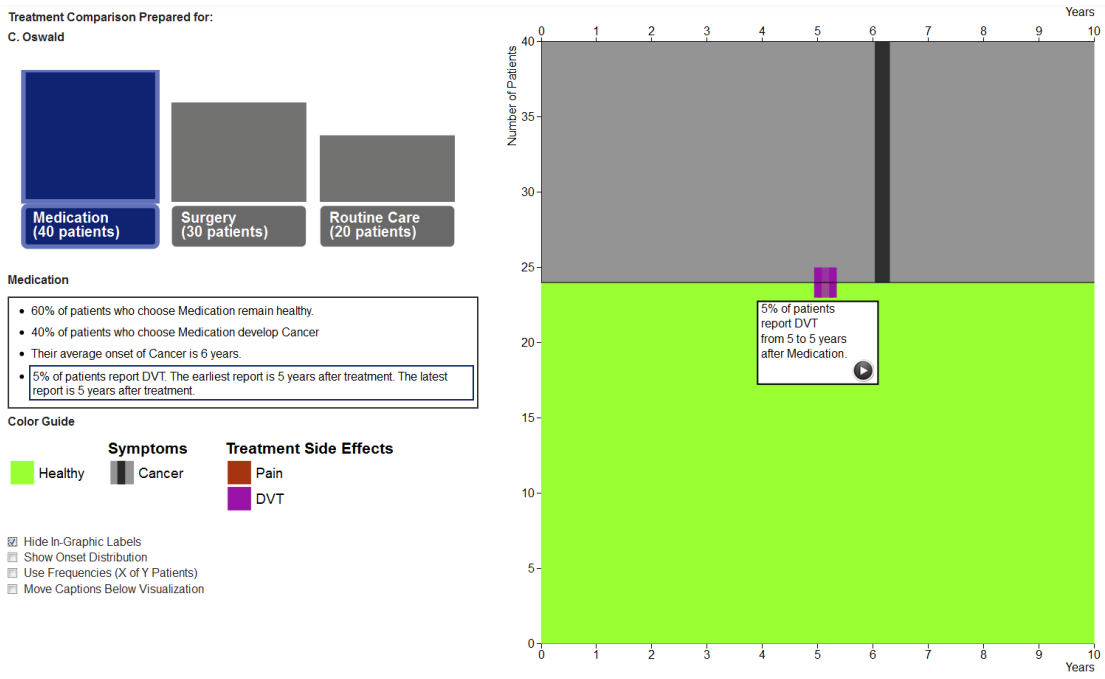


Figure 28. Animation Step 5: A rectangular region representing the side effects of a treatment animates onto the visualization over the range of time that reports of the side effect appear and sizes vertically by the number of patients reporting the side effect.

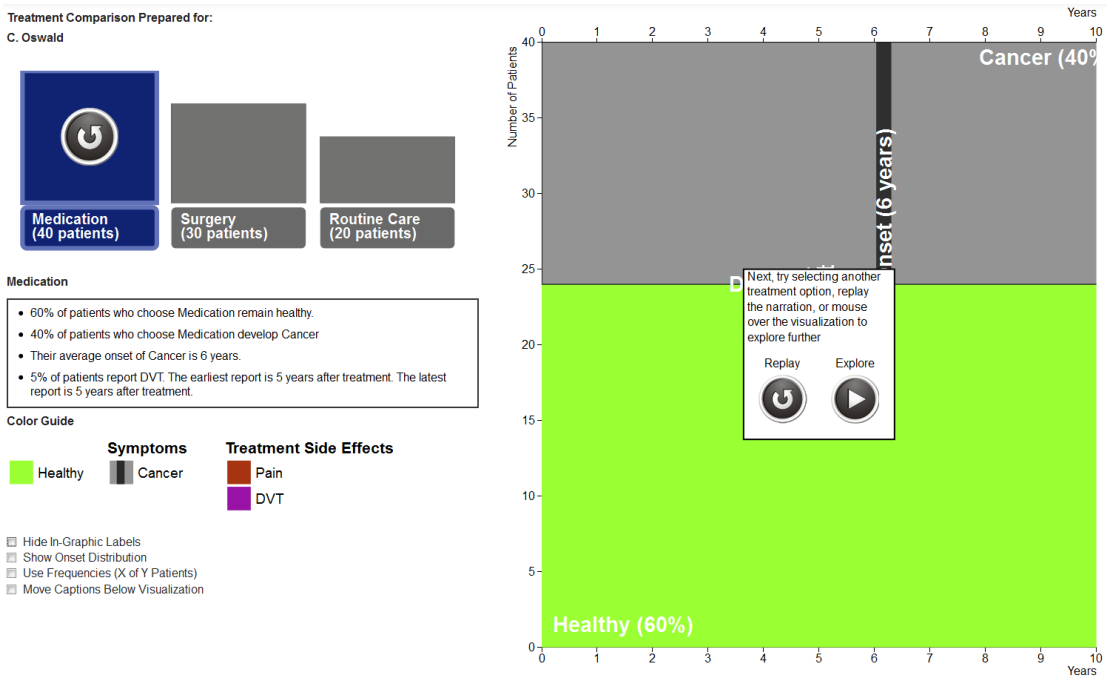


Figure 29. Animation Step 6: An ending screen informs the patient that the guided narration has finished and prompts them to replay the narration, explore the visualization, or to choose a different treatment option.

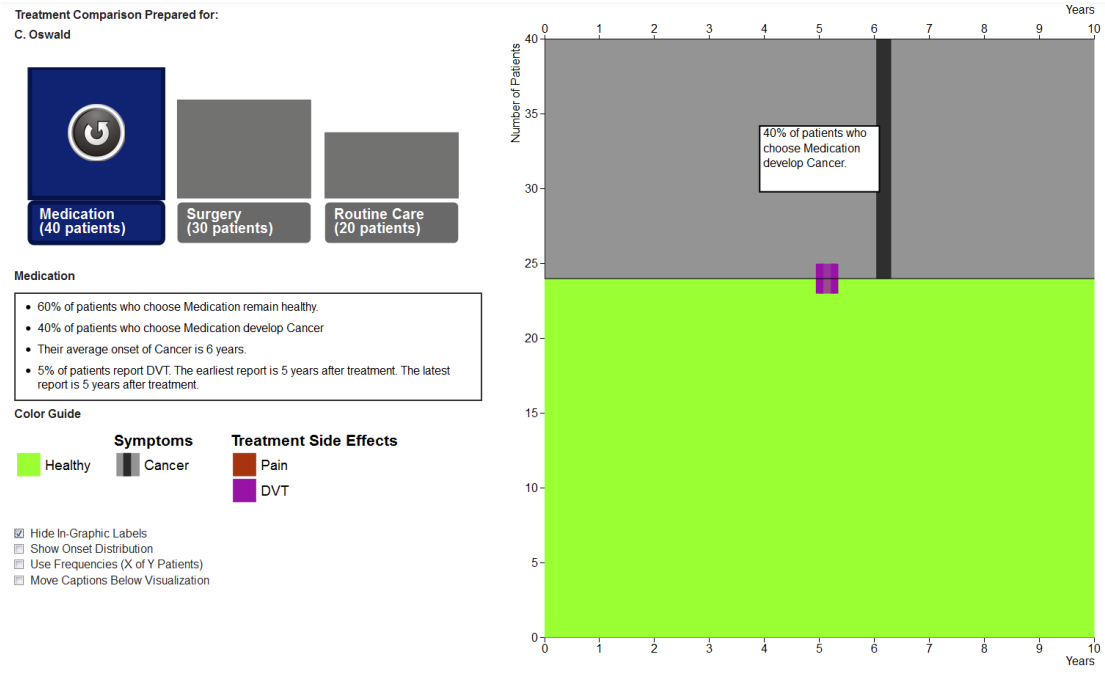


Figure 30. Once the animation has finished, patients can explore the visualization with the mouse. Pop-ups with explanatory captions from the guided narration appear as patients mouse-over visual elements..

Rapid Transition for Faster Comparisons

During pilot testing of TreatmentExplorer, almost all of the participants making use of the fully featured, animated version of TreatmentExplorer made the same request: after commenting on how useful the guided narration was, they asked for a way to quickly move to the end of the animation so that they could refresh their memory faster. They indicated that once they had seen the full narration the first time, it would be more useful if they could quickly move between treatments without having to pause for the narration.

In response to these requests, a shortened version of the animated transition between treatments was developed and replaced the default transition between treatments. Initially, when a patient first selects a treatment, the full, interactive,

guided narration plays. Subsequent selections of that treatment use the shorted version. The shortened narration follows steps 2-5 of the full narration. That is, the healthy region is resized, the afflicted region is resized, the average onset bar is moved, and then the side effects are transitioned. However, in the shortened transition, these steps play as a sequence without pausing for user interaction. The duration of the animations is also shortened such that the entire sequence completes in one third of the time of the full animation (not counting pauses for user interaction). Further, the text highlights for a treatment appear at once after the shortened animation sequence plays. In-visualization captions are not shown during the rapid animation, but after the transition the in-visualization captions become available as patients explore the treatment.

Advantages of TreatmentExplorer

To summarize, the TreatmentExplorer prototype offers the following advantages towards addressing the design goals of Chapter 1:

- Multiple Treatment Options are Shown: Patients are able to explore multiple treatment options and compare them in the same decision aid. Because treatments are visualized individually, there is space for a complete treatment picture which includes success rate, onset of symptoms/condition, and individual side effects associated with a treatment.
- Personalized by Relating to Similar Patients: By drawing its data from EHRs (presumably from local clinic systems), patients are presented with information which reflects realistic outcomes based on the information from

similar patients receiving treatment from similar clinics. This provides a credible source of information but also entertains the possibility of relating patient testimonials and community support to patients making healthcare decisions.

- Data is Organized with Narration: Patients in the process of learning about their health situation and their treatment options are supported in organizing the relevant data. Not only are the important facts about a treatment introduced in a structured way, but their visual representation is explained as it appears in the visualization.

4.4 Software Engineering Design to Support Flexibility

TreatmentExplorer is intended to be the patient-user facing interface to a larger clinical system with access to large numbers of electronic health records. As such, there are several technical challenges that needed to be overcome in its design.

Input Source

Ultimately, TreatmentExplorer will have access to a database of patient health records. To simulate this source of individual records during development, two text files serve as input for the system. The first file simulates individual patient health records and is based on the format used by the EventFlow system [58]. Patient records appear as a series of events with one event on each line of the file and information about the file tab delimited. Event details include the patient's id, the name of the event, the start date of the event, and an optional end date of the event (in

case of interval-based events). The second input file is a meta-information file. This meta-file contains the name of an event and the category of the event. This category information instructs TreatmentExplorer on the interpretation of the event. This interpretation affects how events are counted and used in the statistical calculations performed by TreatmentExplorer and how the event is represented in the resulting visualizations. The categories used by TreatmentExplorer are summarized in table 3.

Table 3. Event Categories used by TreatmentExplorer

Category	Interpretation
Diagnosis	Beginning of patient records. Used to determine the starting date of a patient's records and events.
Treatment	Name of a treatment option represented by patient records. This is an interval event representing how long a patient underwent the named treatment.
Adverse	Name of an adverse side effect of a treatment. This is a complication associated with a treatment that is independent of any primary symptoms or conditions a patient has been diagnosed with.
Condition	Name of a primary symptom or primary condition a patient has been diagnosed with and is seeking treatment fore.

By using text files, the TreatmentExplorer prototype can be used to visualize any arbitrary condition with any number of treatment options. This makes TreatmentExplorer a readily adaptable and generalizable prototype.

Data Structures

Associative arrays are used to store patient health records. These arrays are first accessed by patient id and entries are also associative arrays keyed by event start date. Supporting associative arrays map conditions and adverse side effects to the ids of patients who report those events in their records.

Visualization Supporting Algorithms

The data structure containing health records and keyed by patient id and the structure associating treatments with the ids of patients undergoing that treatment is built through the following algorithm (in pseudo code):

```
for each Patient Record as {id, event, start, end}:
        if PatientRecords[id] is undefined
                PatientRecords[id] = {}
        if PatientRecords[id][start] is undefined
                PatientRecords[id][start] = []
        push {id, event, interval, start, end} to PatientRecords[id][start]
        if EventPatients[event] is undefined
                EventPatients[event] = {}
        if EventPatients[event][id] is undefined
                EventPatients[event][id] = []
        push {start} to EventPatients[event][id]
```

In this algorithm, 'interval' is a Boolean value used to indicate whether or not a valid end date for the event was part of the record as it was read from the event text file. Point events have undefined end dates in the input text file and will not have an 'end' value pushed to the PatientRecords structure. This first data structure building loop is needed to collect patient records in order of starting dates which is not required or enforced in the input file. Date normalization is performed as part of the algorithm which builds the structure that supports the TreatmentExplorer visualization. This

date normalization converts all dates to a difference between the beginning of a treatment and the event. This allows TreatmentExplorer to represent patient records equally when they may come from very different periods of time. Date-based data points represented in the TreatmentExplorer visualization are thus relative to the beginning of treatment.

The visualization-supporting data structure is built through the following algorithm (in pseudo code):

```
visData = []
for each Treatment in Treatment List
    count = 0
    datum = {}
    TreatedPatients = EventPatients[treatment]
    for each PatientId in TreatedPatients
        count = count + 1
        first = null
        last = null
        patientConditions = []
        patientAdverses = {}
        for each Date in PatientRecords[PatientId]
            if Date is after last
                last = Date
            else if Date is before first
                first = Date
            for each Record in PatientRecords[PatientId][Date]
                if Record.event is Condition
                    push Record to patientConditions
                else if Record.event is Adverse
                    push Record to
                    patientAdverses[Record.event]
        difference = last – first
        push difference to datum.durations
        for each Record in patientConditions
            difference = Record.start – first
            push difference to datum.conditions
        for each Record in patientAdverses
            difference = Record.start – first
            push difference to datum.adverses[Record.event]
    datum.patientCount = count
    datum.conditionOnset = Avg(datum.conditions)
```

```
        datum.durationMax  = Max(datum.durations)
        push datum to visData
```

Interface Implementation

In order to provide patients with a readily accessible decision aid, the TreatmentExplorer prototype has been implemented as an interactive website using HTML and JavaScript. The D3 JavaScript library [60] is used to provide SVG based visualization support. By building TreatmentExplorer as a web-based decision aid, it allows patients to access the decision aid from convenient locations during and after consultations with their physicians and healthcare providers. It also eliminates the need for software installation and maintenance duties to be forced upon patients. TreatmentExplorer can access remote record systems and provide patients the most important information from a wide range of sources.

Technical Limitations

TreatmentExplorer has been designed to be adaptable and generalizable to a wide variety of conditions. However, in its early development state, there are some simplifying assumptions made by the data structures and algorithms. It is assumed that there is only one primary symptom or condition of concern to a patient. While TreatmentExplorer can represent an arbitrary number of treatment options, it is assumed they all relate to the treatment of a single condition. Treatments are assumed to be mutually exclusive and modifications would be needed for both the algorithms and visualization to accurately represent patient records of patients undergoing multiple treatments. Treatments are assumed to have at most a single

adverse side effect in the guided narration though this is not enforced in the current data structures or algorithms. Future research will need to explore options for presenting multiple side effects and conditions on a single visualization.

4.5 Expert Review Process

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. Experts were recruited through mutual contacts based on their research focus and working experience. These demo and interview sessions were conducted at the experts' work location and lasted an hour. Suggestions and feedback were then incorporated into the next version of TreatmentExplorer before it was shown to the next expert. Consensus between experts in feature suggestions and improvement was used to decide which improvements to make. Along with interface and feature suggestions, experts were asked about the possible role of a fully developed system in a professional medical setting. Evaluation strategies were also discussed for evaluating the efficacy of such a system.

Expert Review Scenario

For demonstrating TreatmentExplorer to expert reviewers, the following scenario was given as a typical use-case:

> A patient, Donna, is 60 years old and has a family history of breast cancer. She's also recently been tested positive for the BRCA1 gene mutations so she's at a high risk of developing cancer herself. She's

now deciding whether or not she should undergo surgery or take medication to reduce her risk.

The demonstration dataset was comprised of synthetic data. It simulated 90 patient records with rates of cancer and side effects derived from publically available breast cancer fact sheets available from the National Cancer Institute [51] and the Susan G. Komen Breast Cancer Foundation [61]. Three treatment options were available: routine care (no treatment), medication (based on Tamoxifen), and surgery (based on double prophylactic mastectomy).

4.6 Improvements Resulting from Expert Reviews

Each expert received TreatmentExplorer very well with positive comments ranging from "an excellent start" to "this is gorgeous". Frequently, experts asked for a version of TreatmentExplorer which they could use to investigate their own data. In the resulting discussions, several desired improvements were made clear. These suggestions centered primarily on improving the visualization and making it more accessible to patients. The improvements with the greatest usability impact were immediately implemented. Several other feature requests from experts are in consideration for future development. The immediately implemented improvements include:

Add a representation of the distribution of condition onsets to the display as an advanced option instead of only providing the average.

One expert who was a practicing medical professional indicated that physicians would not make use of a measure such as the average onset of cancer. Rather, they

would insist on having the distribution of cancer reports from the entire treatment group available. This request was made by other experts though others with experience in public health and communication suggested that such detailed information be kept as an advanced option. Alternatives to a distribution were discussed such as a median mark with quartile range markers similar to a standard box-and-whisker plot or a line graph representation showing the cumulative reports of a condition throughout the represented range. After discussion with the experts and the drafting of possible representations, it was decided that the line graph showing the cumulative reports of condition onsets was the best solution. This line would overlay the section of the visualization representing afflicted patients and accompany the bar marking the average onset time when in use.

Provide color scales suitable for patients with colorblindness

Several experts viewed an early prototype of TreatmentExplorer and remarked that the colors used to represent condition onsets and side effects were too similar. One expert was himself colorblind and had difficulties perceiving the range of time that certain side effects were reported in the demonstration scenario. The color scheme of the prototype was immediately revised to use more contrasting colors suitable for colorblindness and acceptable when converted to grayscale.

Use simple animations so that patients only need to follow one moving object

Early versions of TreatmentExplorer's guided narration transitioned the sizes of the representation of healthy patients and condition-afflicted patients when switching from one treatment to another. This was intended to provide a means of comparison for patients and to reinforce the differences between the outcomes of patients in

different treatment groups. Experts remarked, however, that with accompanying text transitioning as well that there were too many moving parts to the visualization and that patients might feel overwhelmed at first. To reduce the barrier to use, animations accompanying the guided narration were simplified so that all treatments began from the same, empty visualization area. Each component of the visualization was removed at the start of narration and re-appeared one at a time as the narration progressed.

Include a description of the limitations of the represented health records

To accompany the guided narration, it was suggested that a beginning text screen provide a high level disclaimer and description of the records the visualization represented before the narration began. This information would help patients understand what the visualization was capable of reporting and what the limitations of the dataset were.

Report all data from same maximum-scaled time-frame

Early versions of TreatmentExplorer ended the representation of a treatment group's records at the average onset of a condition. This was intended to reinforce the difference in onset times between treatment groups by horizontally sizing the groups' representations. Several experts indicated that this might be confusing to patients for several reasons. It could be interpreted as missing information or as extremely adverse outcomes such as deaths in the treatment groups. It was also a representation contrary to traditional risk communication which often provides data for specified and consistent time periods such as 10 years after treatment. To avoid confusing patients, all treatment groups extend to represent the full range of patient records in a

dataset and an additional vertical bar is used to indicate the average onset time of a condition.

Grow visual display from bottom of the chart, rather than the top

One expert pointed out that the graphics patients are most experienced with and are found in most decision aids are oriented with their y axes beginning at 0 at the bottom of a graph and increasing vertically towards the top. The original layout of the TreatmentExplorer visualization had both the x and y axes beginning with 0 located in the top left corner. This was intended to avoid confusing patients with different values of axes located close together. However, it was decided after discussion that re-orienting the y axis to its customary 0-at-the-bottom would be less confusing. Additionally, a duplicate x axis was added to the bottom of the visualization to reinforce the difference in scales. As the guided narration progressed, then, vertical space in the visualization was filled from the bottom and growing towards the top. In this way, a growing portion of healthy patients could be associated with a better treatment success rate.

Overlap side effects so that they overlay both health and condition onset patients in the visual display

Initially, the reports of side effects of a treatment were handled as a group of patients separate from those patients who eventually developed a condition. One expert indicated that this would be misleading to patients who might assume that treatment associated side effects were not reported by patients who developed a condition. To avoid misinforming patients, the drawing of side effect representations was revised such that an equal portion of the side effect representation would overlay the

representation of patients who developed a condition. In this way, side effects and their reports would be shown as independent of whether or not patients developed a condition and associated with the treatment group as a whole instead.

4.7 TreatmentExplorer Role in Decision Making and Evaluation Strategies

Among the improvements to the visualization component of TreatmentExplorer, several experts discussed its possible role in a clinic environment. It was frequently remarked that the TreatmentExplorer visualization by itself could be too overwhelming for the average patient. However, several of TreatmentExplorer's features improved the accessibility. Particularly, the guided narration served to both teach patients about how to use the visualization as well as inform them about their treatment options. The replay option would further support patients by allowing them to watch and re-watch the narration until they felt comfortable with all the information. The experts were well in agreement that the ideal use case for TreatmentExplorer would begin with both the patient and their physician setting up the treatment options together in an initial consultation. During such a 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 view the guided narrations and explore their chosen treatment options.

In discussions about evaluations, the experts suggested a few possible alternatives. One possible evaluation plan would be to sit with patients and

physicians in a mock-consultation setting for a small series of case studies. Another alternative suggested was to focus on a single condition using a real dataset of anonymized patient records of patients with a specific heart-related condition and to discuss TreatmentExplorer with experts from advocacy and support groups relating to that condition. Both these evaluation plans would provide valuable feedback from real patients and potential users of TreatmentExplorer and future work will likely involve such case studies. A third evaluation option discussed involved the controlled study of TreatmentExplorer and the comparison of patient performance on a basic knowledge test using TreatmentExplorer with patient performance when using a comparative baseline decision aid. This third option would allow for quantitative data to be collected about which features were responsible for patient understanding (or misunderstandings) of treatment options. A preliminary evaluation following the intent of this controlled user study is described in the next chapter.

4.8 Improvements Resulting from Participant Feedback

A pilot study of eight participants (see chapter 5) was conducted to elicit feedback about usability issues before a full evaluation of TreatmentExplorer. These participants were recruited through a convenient sample of fellow HCI researchers but none had prior experience with TreatmentExplorer. They were asked about which features they found most helpful during a debriefing at the end of their pilot study session. Participants using the full-featured version of TreatmentExplorer all made the same request at some point in their debriefing leading to a subtle but critical design change for TreatmentExplorer.

Allow patients to quickly scan their treatment options without viewing the entire guided narration.

Originally, the default behavior of TreatmentExplorer was to transition between treatments by guided narration. This animation served two purposes: 1) to teach users how to read the visualization and 2) to emphasize each important measurement related to a treatment. A replay feature was included early on to allow patients to re-watch the entire narration as many times as they wished. However, pilot participants felt the full narration interfered with their ability to quickly compare treatments. In order to preserve the benefits of animation in emphasizing treatment differences, a faster animated transition was introduced to be the default transition between treatments. When a treatment is first viewed, the complete narration with pauses for user control and in-visualization text or captions is used to introduce the patient to both the treatment and the visualization. The next time the same treatment is selected, the sped-up transition was used. The sped-up transition followed each step of the full animation but took 1/3 the duration. Additionally, pauses for user control were removed and in-visualization text or captions were hidden. With this sped-up transition, patients would still see each treatment difference emphasized as they explored their treatment options. The replay button remained available so that patients can always replay the guided narration. This sped up-transition is only enabled on a treatment after a patient views the full, guided narration once. In this way, patients still experience the benefits of the guided narration but rapid comparison of treatments is better supported.

4.9 TreatmentExplorer Design and Expert Review Summary

The TreatmentExplorer prototype is an interactive, web-based decision aid which utilizes the D3 JavaScript library [60] to provide guided narration and visualization support to patients. It is capable of representing an arbitrary number of treatment options with at most one adverse side effect. The current design represents a single treatment at a time so that details about a treatment can be represented in the visualization. Transitions between treatments enable comparison across treatment options. This design overcomes the challenges presented by both basic graphs and design alternatives which attempt to represent all treatments at once. Four experts were interviewed individually for an hour and provided feedback on the role of TreatmentExplroer in a clinical environment, its evaluation, and features.

Chapter 5: Preliminary User Evaluations

A preliminary usability evaluation was conducted to evaluate the efficacy of the feature-complete TreatmentExplorer prototype. This study was designed to measure patient knowledge gains when using TreatmentExplorer over a text-only decision aid similar to those found online for comparing treatment options. The study procedure was inspired from discussions with experts in the field of public health and medicine as well as the controlled studies conducted in the research reviewed as part of the earlier literature review. A pilot study of eight participants was first run to determine the best procedure for measuring the dependent variables. The results of this pilot study were used to refine the procedure for a full, follow-up evaluation.

5.1 Pilot Study

Because of the preliminary nature of this research, a pilot study was used to determine the best procedure for a full evaluation. Two potential procedures differed in how the dependent variables of time and accuracy would be recorded. The purpose of the pilot study was to determine which of the two possible procedures would provide the best opportunity to measure differences between TreatmentExplorer and the text-only decision aid. Feedback from pilot study participants was also used for identifying usability issues early so they could be corrected before a full evaluation.

Pilot Study Participants

A total of n=8 participants were recruited from a convenient pool of graduate students and staff from the Human Computer Interaction Lab at the University of Maryland but none had prior experience with TreatmentExplorer as either developers or advisors.

Pilot Study Materials

Three interfaces were used during the pilot study. The first interface was a text-based web-page with the information needed to answer the questionnaire questions written in paragraph form with both percentages and frequencies for data points. The second interface was a static version of TreatmentExplorer which showed the textual highlights of each treatment and the final visualization. This static version had no narration, animation, or text-tips. The third interface was the fully featured version of TreatmentExplorer complete with narration, animation, progressive disclosure of highlights, and text-tips in the visualization. A questionnaire with 10 questions about the risk of developing a condition and the side effects of three treatment options was used to elicit knowledge gains from participants.

Pilot Study Procedures

Each interface was pilot tested with two participants with one participant following the rules of Pilot Procedure 1 and the second following the rules of Pilot Procedure 2. Under both procedures the use of calculators, other materials, other decision aids, and note-taking was prohibited. Disallowing note-taking somewhat lowers the external

validity of the study as patients researching treatments would likely take notes throughout the process. However, the intention of the pilot and following evaluations was to measure how well the TreatmentExplorer prototype supported patients with both knowledge and recall as a stand-alone system. Allowing participants to take notes would allow them to organize information for themselves and thus render any conclusions about the utility of the TreatmentExplorer prototype less meaningful. Participants were restricted to the use of their assigned decision aid and the questionnaire.

Pilot Procedure 1

The following partial procedure represents the unique steps of the first potential procedure:

1. Participants read a description of the scenario and their task. They were given an opportunity to ask questions.
2. Participants were taken to the site for their decision aid.
3. Participants were given the questionnaire and timing of their session began.
4. Participants used their decision aid to complete their questionnaire.
5. Participants returned their questionnaire to the researcher when finished or at the end of a 5 minute maximum.

In this first procedure, participants were free to interact with their decision aid while they filled out the questionnaire. Dependent variables were the total number of incorrect responses provided by participants and time. Time was recorded as a continuous running clock started the moment participants were given their

questionnaire and stopped when the questionnaire was complete. Participants would not have the opportunity to correct incorrect responses during the session.

Two participants were recruited and followed this procedure with a higher time limit of 10 minutes maximum. Both participants completed the questionnaire in less than 5 minutes and so the maximum time was lowered to 5 minutes for the rest of the pilot study.

Pilot Procedure 2

The following partial procedure represents the unique steps of the second potential procedure:

1. Participants read a description of the scenario and their task. They were given the opportunity to ask questions.
2. Participants were taken to the site for their decision aid and given 3 minutes to learn the interface and investigate their treatment options (no training was provided at all). They were told that they would be answering questions about the information in the decision aid after the three minutes were over.
3. The browser with the decision aid site was minimized and participants were given their questionnaire and a maximum of one minute to answer as many questions as they could form the questionnaire.
4. Researchers marked all incorrect responses and the participants were given more time to explore their treatment options and correct their answers until they were completely correct.

In this second procedure, participant interaction with their decision aid was limited to at most 3 minutes at a time and the number of times each decision aid was consulted

was recorded. Time would thus be recorded as the number of times participants needed to return to their decision aid. This count reflects the number of consultations the decision aid requires before patients using it become fully knowledgeable about its content. The cumulative total number of incorrect responses provided by participants was counted.

Two participants were recruited and followed this procedure with longer time limits for both the decision aid interaction phase and questionnaire completion phase. These two participants used variations of the TreatmentExplorer decision aid and both were able to complete their questionnaires with high accuracy after a single interaction with the decision aid and their questionnaire in the first attempt. The maximum interaction time with the decision aid was thus lowered from 5 minutes to 3 minutes based on observation of the first two participants' use of the decision aids. Both spent the first 3 minutes intently studying the three treatment options methodically and the remaining 2 minutes rapidly clicking between them. With the shortened time of 3 minutes, participants had enough time to watch each complete animation one or two times. The questionnaire phase originally had a maximum time limit of two minutes. Observation of participants completing the questionnaire in the questionnaire phase suggested that participants spent the first minute answering questions that they were certain of and making guesses for the remaining minute. The questionnaire phase was thus shorted to a one minute maximum for the rest of the pilot study.

Pilot Participant Feedback

As part of the pilot study procedure, all participants were given a debriefing at the end of their session and asked two questions:

1. How well did this decision aid support your answering the questions on the questionnaire? (With 1 meaning not at all and 10 meaning very well)
2. What feature of the decision aid did you find most useful?

Because of the preliminary nature of this evaluation, the two subjective questions asked during debriefing were intended to elicit new design considerations from our participants. Specifically, it was important to find out if participants made use of both the textual highlights and graphic visualization components of TreatmentExplorer equally, or whether they used the textual highlights without consulting the visualization.

The feedback from the pilot study participants was very encouraging. Many remarked that the visualization component of TreatmentExplorer was the fastest way to answer questions and that the textual highlights served as a back-up to confirm what participants read from the visualization. Participants using the fully featured, animated version of TreatmentExplorer remarked that the animation helped them remember the treatment differences and answer the questions of the questionnaire faster. Participants using the animated version of TreatmentExplorer also rated their decision aid higher on the scale of 1 to 10 than participants given a static, non-interactive version of TreatmentExplorer that had no accompanying narration or animation (8/8.5 vs. 7/6). Further, one participant using the static version of TreatmentExplorer commented during debriefing that she would forget numbers in

between answering questions but remember them when glancing at the visualiation. Participants using both versions of TreatmentExplorer remarked that the bright colors of the visualization helped them remember the differences between treatments because "each picture was very distinct from the others."

During the pilot study, the benefit of animation and interactivity became clear from the comments of two participants. One participant remarked how the animating bar representing condition onset "helped [her] figure out that there were two parts to the condition: who got it and when they got it." Another participant who had use of the static version of TreatmentExplorer commented that "[she] didn't figure out that the dark bar meant something else too besides that those people got sick." She followed by adding that she only realized that condition onset information was in the visualization when a question on the questionnaire asked about it. Participants using the static version of TreatmentExplorer also mentioned during debriefing that they would like to have had interactive features, including those available from the fully featured version of TreatmentExplorer. Since participants had no prior experience with TreatmentExplorer, these were spontaneous design requests which lend support to the utility of the full-featured, animated version of TreatmentExplorer as an interactive tool.

While the animation was clearly helpful to participants of the pilot study, a single feature was requested by each participant who used the fully animated version: the ability to quickly skip through the animation to the ending visualization. Once participants had viewed the animations the first time, they wanted to be able to quickly refresh their memory when they next selected a treatment. The

TreatmentExplorer prototype was revised to include a rapid transition to replace the guided narration as the default transition between treatments. This rapid transition was only enabled after participants had completely viewed the guided narration for a treatment once. The rapid transition followed the same steps as the guided narration but without pausing for user interaction or displaying in-visualization text tips. Participants could replay the complete narration at any time through use of the replay button.

Pilot Study Summary

The encouraging results of the pilot study suggest that the three variations of decision aids differ in important ways. Firstly, participants seemed to remember more information from the fully featured version of TreatmentExplorer than from the static, non-interactive version. Participants also subjectively rated the utility of the full featured TreatmentExplorer higher than the static version.

5.2 Usability Study

Following from the results of the procedure pilot study, a preliminary usability evaluation was conducted to evaluate the efficacy of the feature-complete TreatmentExplorer prototype. The procedure of the full evaluation was drawn from the most successful pilot study procedure.

Research Questions

This preliminary usability study was designed to answer the following research questions derived from the observations made during the pilot study:

1. Are there statistically significant differences in the number of times participants must consult their given decision aid in order to complete the questionnaire?
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 between the three different interfaces?

Participants

24 participants were recruited through emails to mailing lists, paper fliers, and verbal advertisements. These participants came from the undergraduate and graduate population at the University of Maryland and were compensated $10 for their participation in a session. The procedure was refined by the results of the eight participant pilot study and was based on pilot procedure 2.

Design

This study followed a 1x3 between subjects design with participants using one of three possible decision aids to complete a short questionnaire about a fictitious condition. A between-subjects design was chosen to both eliminate the need to hand-

build three dissimilar but equally complex sets of conditions/treatments and to lessen the intensity of study sessions on participants (who would have to learn detailed statistical information about three different conditions).

Materials

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 [62] . 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. (See figure 37 in Appendix B).

Decision Aid 2: Static Graphic TreatmentExplorer

A functionally limited version of TreatmentExplorer was used as a second level designed to provide a limited experience with TreatmentExplorer. 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 of TreatmentExplorer. (See figure 38 in Appendix B).

Decision Aid 3: Full-Featured TreatmentExplorer

The final level of this study made use of the fully-featured TreatmentExplorer prototype including guided narration, animation, progressive disclosure, and reinforcing captions. (See figure 39 in Appendix B).

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. Participants would thus need to use the decision aids provided to them in the study to complete the questionnaire given to them as part of the study. This synthetic data and fictitious condition also eliminated the possibility 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 provided fictitious patient records for 120 patients dealing with a single medical condition and three possible treatments, each with a single unique adverse side effect.

Scenario

As part of the instructions for this study, participants read a description of their task in the form of the following scenario:

> "Imagine that you have recently been to your doctor's office because you have been experiencing dry, red skin on your hands and feet as well as a mild

tingling sensation in your fingers and toes. Your doctor examined you and determined that you have developed a condition known as "Crimson Blot Syndrome". Along with uncomfortable dry skin, this condition carries a serious risk of sudden paralysis and needs to be treated immediately. Fortunately, there are several treatment options but each carries a risk of some side effect. You must now make a decision about which treatment option is the best choice for you.

To help you make this decision, you will be given a decision aid which explains the success rate of each treatment and its side effects. Your doctor has also given you a short list of questions for you to answer about each treatment to help you learn about them. You will need to be able to answer these questions from memory and without the use of the decision aid later. You will be able to use the decision aid for 3 minutes and then be asked to answer as many questions from the list as you can in one minute. You will be able to return to the decision aid and repeat this process as many times as you need to complete the questionnaire."

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 (see Appendix A). 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

[59]. Time was recorded as the cumulative sum of times a participant needed to consult their decision aid. The cumulative sum of incorrect questionnaire responses provided by participants was also recorded. After each chance to complete the questionnaire, researchers marked any incorrect response for participants. Participants were allowed to provide new answers after exploring treatment options with the decision aid. These new answers were also marked if they were incorrect and added to the cumulative sum of incorrect responses. Incomplete answers after a questionnaire answering phase were not counted as incorrect.

Procedure

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

1. Participants arrived and signed a consent form. They were told that they would receive $10 for the session and that a bonus was offered for the best performance in their group (fewest incorrect questionnaire answers with the fewest decision aid exploration phases used to break ties)
2. Participants read a description of the scenario and their task. They were given the opportunity to ask questions.
3. Participants were taken to the site for their decision aid and given 3 minutes to learn the interface and investigate their treatment options. They were told that they would be answering questions about the three treatment options once their 3 minute exploration time was over. Participants did not have access to the questionnaire or questions during this decision aid interaction period.

4. The decision aid was taken away and participants were given their questionnaire and a maximum of one minute to answer as many questions as they could. Participants did not have access to their decision aids during the questionnaire answering period.
5. Researchers marked and counted any incorrect responses and the participants were allowed to repeat the treatment exploration/questionnaire answering phases and correct their answers. No partial credit was given, answers were either completely correct or marked incorrect.
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:
    1) How well did this decision aid support answering the questionnaire? (Using a scale of 1 to 10 with 1 being poorly and 10 being very well)
    2) What feature of this decision aid did you find the most helpful?

Steps 3-5 were repeated until the questionnaire was completely and correctly filled out. During the session, participants were not allowed to use calculators, consult any other materials, or take notes of their own. They were also not allowed to switch decision aids. 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. Participants were allowed to ask clarification questions if they felt they were needed to complete the questionnaire.

The debriefing phase (step 6) was used to elicit feedback from participants. Participants' subjective rating of decision aid utility was used to answer research

question 3. Design considerations were also sought for continued TreatmentExplorer development.

Results

A total of 28 participants volunteered to partake in the preliminary user evaluation. Three participants failed to arrive for their selected session. A fourth participant's results were invalidated when they disobeyed study instructions and left their assigned decision aid and began using a different version. This left 24 participants with valid performances for analysis.

Performance time across appeared to be affected by interface option, however the results were not significant after a one-way, independent-measures ANOVA (three treatment levels) due to high within-treatment variance ($p > 0.05$, $F(2,21) = 2.58$, see table 4, figure 31).

Table 4. Means and Standard Deviations of Decision Aid Uses.

	Text Only	Static Graphic	Full-Featured
M	3.625	2.625	2.75
SD	0.92	1.06	0.87

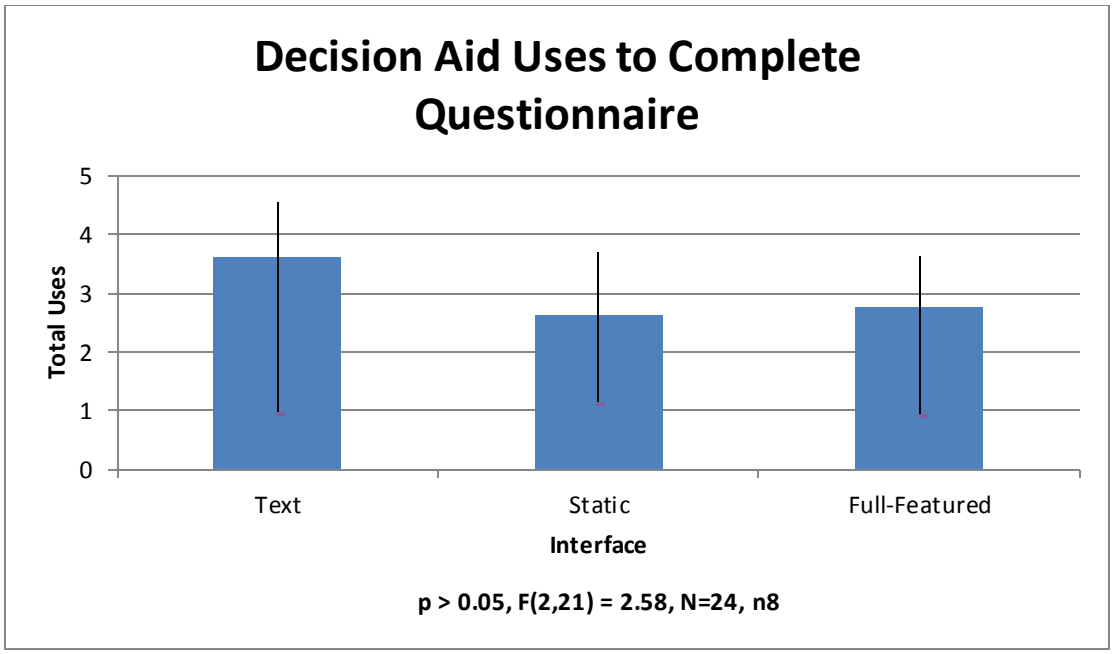


Figure 31. Means and Standard Deviations of Decision Aid Uses

Inaccurate responses were analyzed at several levels. An initial one-way independent-measures ANOVA assessed the total number of incorrect responses participants provided for all 10 questions of the questionnaire but was not significant due to high within-treatment variance ($p > 0.05$, $F(2,21) = 0.95$, see table 5, figure 32.)

Table 5. Means and Standard Deviation for Total Incorrect Responses

	Text Only	Static Graphic	Full-Featured
M	4.25	3.88	3
SD	1.83	2.23	1.41

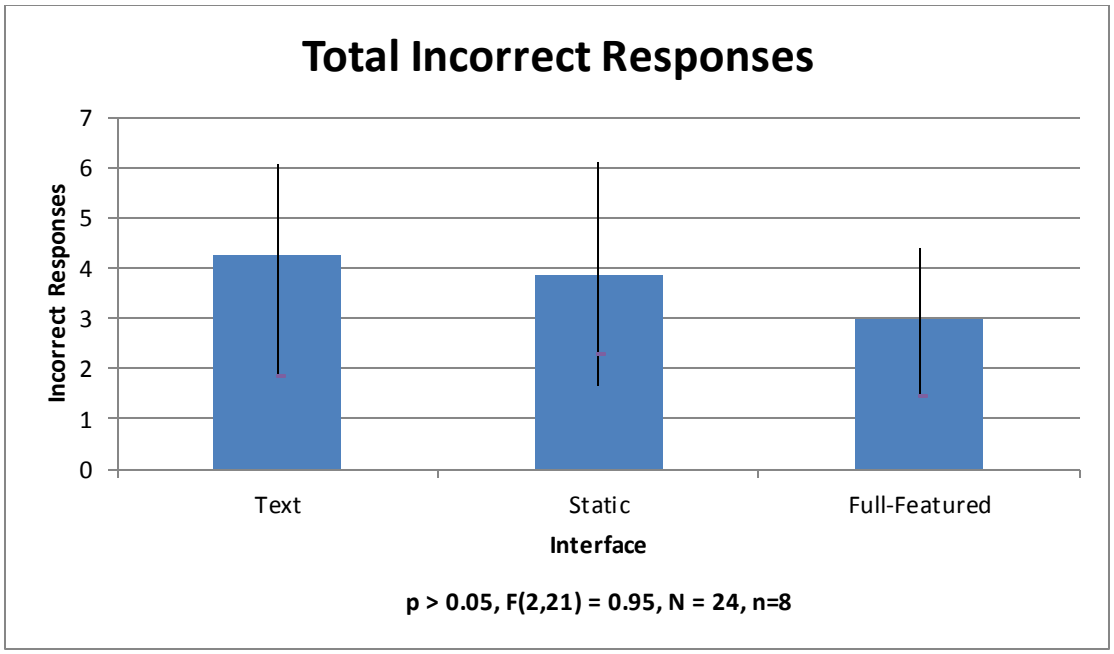


Figure 32. Means and Standard Deviations for Total Incorrect Responses

Follow-up analysis separated the questions of the questionnaire into two groups based on the subject of the question. The first five questions of the questionnaire focused on the prevalence and risk of scenario side effects. A one-way independent-measures ANOVA revealed significant differences in the number of incorrect responses for the five questionnaire questions dealing with scenario side effects ($p < 0.05$, $F(2,21) = 4.97$, see table 6, figure 33). A post-hoc Tukey's HSD test revealed significant differences between the text only interface and the full-featured interface with regards to incorrect responses. Participants using the full-featured interface were able to answer questions about side effects with significantly fewer incorrect responses than participants with a text-only interface.

Table 6. Means and Standard Deviation for Incorrect Responses on Side-Effect Questions

	Text Only	Static Graphic	Full-Featured
M	3.75	3.125	1.625
SD	1.67	1.46	0.92

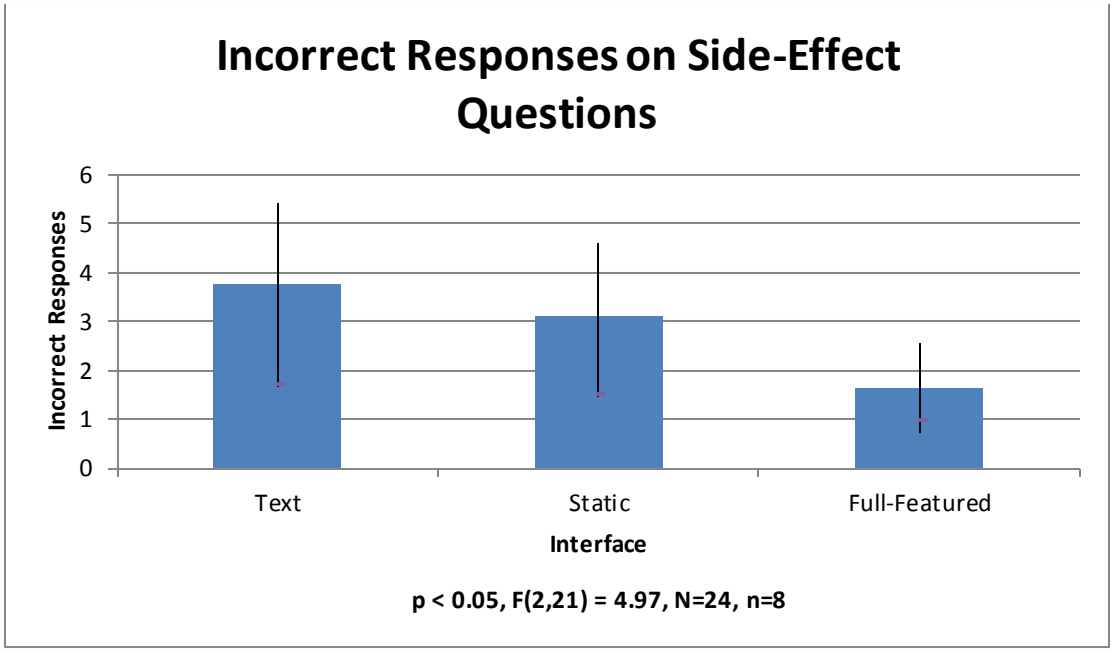


Figure 33. Means and Standard Deviation for Incorrect Responses on Side-Effect Questions

Per-question analysis revealed significant differences on question # 3 of the questionnaire which was "What is my risk of getting headaches if I chose Treatment C?" This question was actually a trick question in that headaches were a side effect associated with a different treatment and the correct response to question 3 was "none" or "zero". A one-way independent-measures ANOVA indicates significant differences in participants' ability to correctly answer this question ($p < 0.05$, $F(2,21)=3.69$, see table 7, figure 34). A post-hoc Tukey's HSD test indicates that

participants using the text-only interface answered question 3 incorrectly more times than participants using the other interfaces.

Table 7. Means and Standard Deviation for Incorrect Responses on Question #3

	Text-Only	Static Graphic	Full-Featured
M	0.875	0.125	0.25
SD	0.83	0.35	0.46

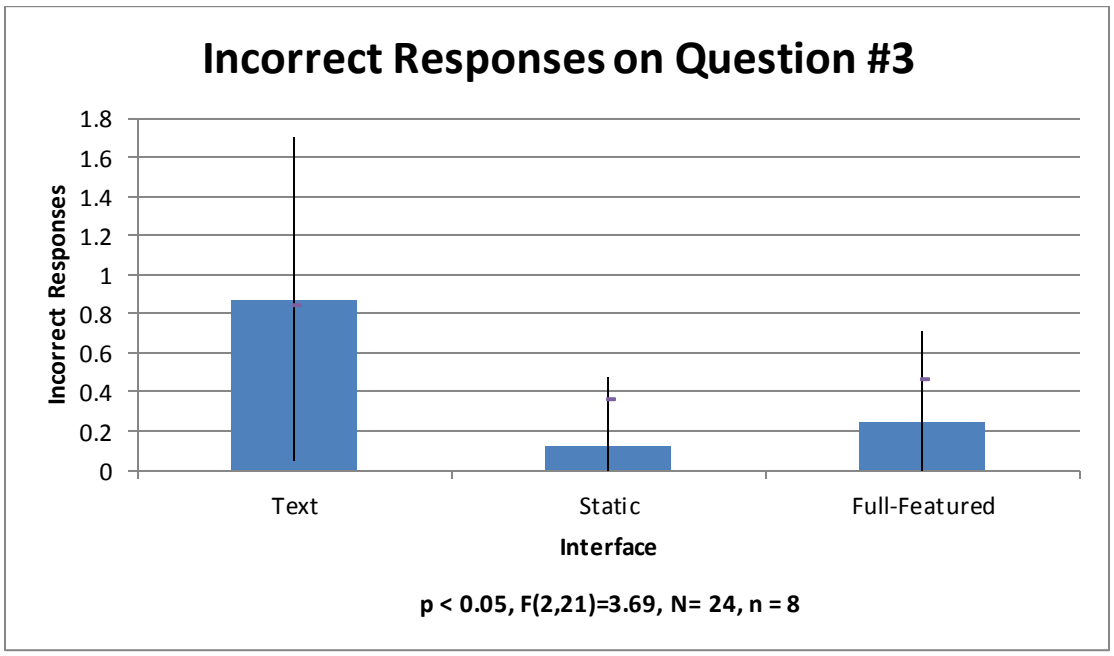


Figure 34. Means and Standard Deviation for Incorrect Responses on Question #3

The second group of five questions of the questionnaire focused on the risk of the primary condition of the scenario. A one-way independent-measures ANOVA did not show significant differences in the incorrect responses participants provided for these questions ($p > 0.05$, $F(2,21)=1.61$, see table 8, figure 35).

Table 8. Means and Standard Deviation for Incorrect Responses on Primary Condition Questions

	Text-Only	Static Graphic	Full-Featured
M	0.75	1.5	1.625
SD	0.71	1.31	1.06

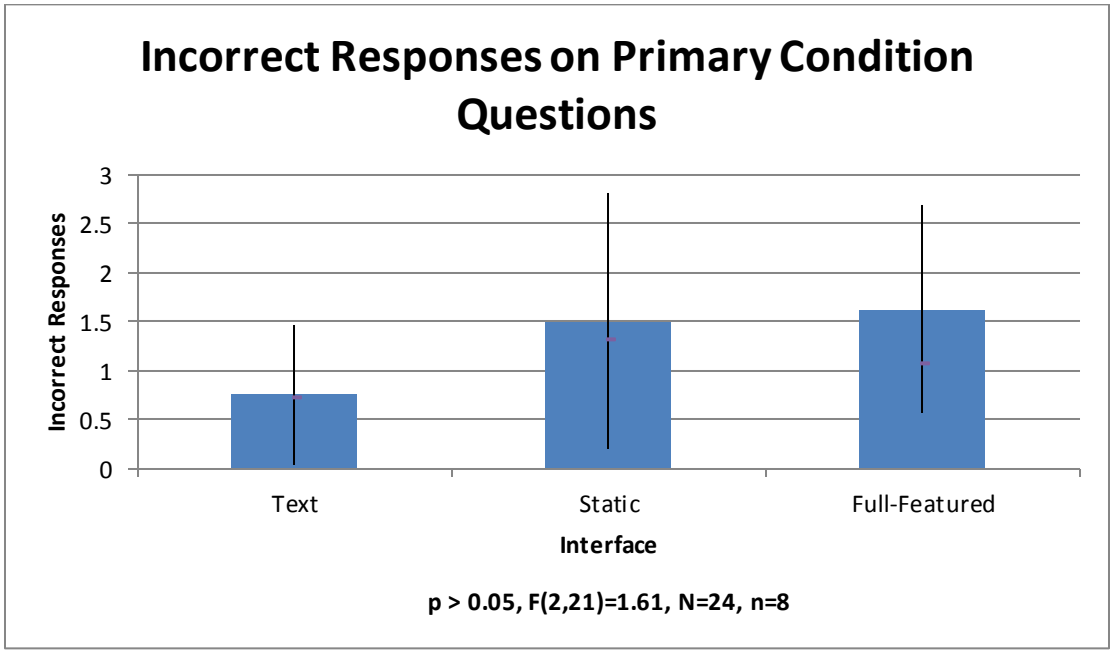


Figure 35. Means and Standard Deviation for Incorrect Responses on Primary Condition Questions

Finally, participants were asked for a subjective rating of the interface they used during the debriefing portion of their session. A one-way independent-measures ANOVA did not show significant differences in the subjective rating of the interfaces ($p > 0.05$, $F(2,21) = 0.46$, see table 9, figure 36).

Table 9. Means and Standard Deviation for Participants' Subjective Interface Rating

	Text-Only	Static Graphic	Full-Featured
M	6.75	7	6.25
SD	0.89	1.93	2.31

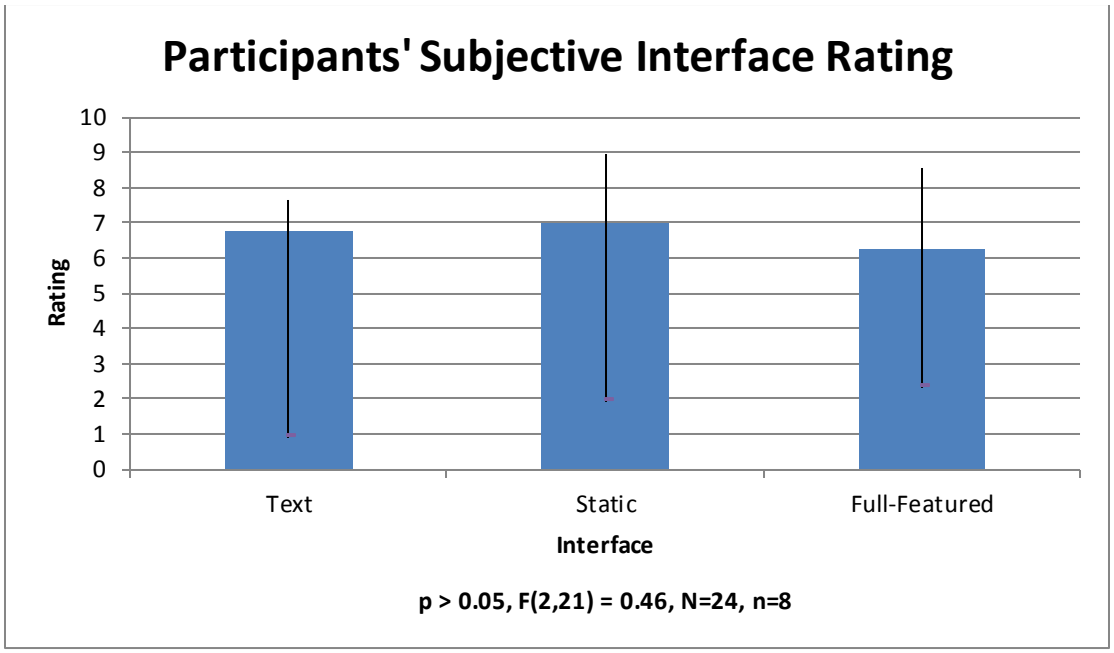


Figure 36. Means and Standard Deviations for Participants' Subjective Interface Rating

Discussion

While the small number of participants in this preliminary evaluation produced a high within-treatment variance, the results from the preliminary user evaluation suggest that refinements to the study procedure may produce significant results. Overall, participants using the full-featured interface appeared to view their decision aid fewer times on average than participants using the text only interface.

Because the null hypothesis cannot be rejected for research question 1, it can be said that the guided narration and novel visualization of the fully-featured TreatmentExplorer did not negatively impact participant completion times. That is, walking through a series of guided narration did not hinder participants when compared with participants using familiar formats like text. Most participants using the full-featured interface needed only two interactions with the decision aid to complete the questionnaire (average 2.75). Two participants in the group, however needed more interactions than the others and thus had a skewing effect on the group. Without these two scores the full-featured interface's average time to completion drops to 2.3. Similarly, most patients in the static graphic interface were also able to complete the questionnaire with two interactions with the interface (average 2.625). One participant, however, misunderstood how the timing of the study would be recorded and would only interact with the interface for about one minute at each opportunity before stating that they were done and ready for the questionnaire (other participants used the full 3-minute allotment). After removing this participant's relatively high number of trials, the average time to completion of the static graphic group drops to 2.28.

For this preliminary evaluation, the scoring of time was kept at a high level and decision to count only the number of times a decision aid was consulted was made to encourage participants to focus on exploring the content of each decision aid, rather than attempting to read as fast as possible or find ways to "fast forward" through animations. During the study however, it was noted that most participants using the full-featured interface had viewed all guided narration within the first

minute of the allotted 3, explored the treatments for the next minute, and then spent any remaining time rapidly clicking between treatments. It also appeared that participants using the text-only interface did not take the time to thoroughly read the content of the interface. Rather, text-only participants appeared to switch rapidly between treatment options and compare treatment data points individually. Improvements to the guided narration intended to reduce the amount of time it takes to view the narration have been made since the running of the evaluation. In addition, other interface improvements may make the full-featured interface more memorable to participants with just the guided narration alone. A future repetition of this study could lower the allotted time per interface interaction from the 3 minutes. With less time to re-read interface content, significant differences may appear revealing which interface is more memorable with the least amount of interaction.

The significant differences in participants' ability to understand and remember treatment side effects is in contrast with the lack of overall significant difference in accuracy between interfaces. This supports in-study observations that participants using the full-featured interface attempted to answer more questions each time they worked on their questionnaire than participants using the text-only interface. In the preliminary version of this evaluation, time to completion was a separate variable recorded at a high level independent of each question. Further, the number of unanswered questions was not tracked. Future iterations of this evaluation should track both the number of unanswered questions after each attempt at the questionnaire as well as how many interface-interactions were needed before each question was correctly answered.

Finally, because this preliminary evaluation was a between-subjects design, the phrasing of the question to elicit a subjective interface rating from participants likely needs to be changed. The procedure of this study dictated that participants return to the interface and continue working on the questionnaire to correct completion. However, the debriefing question was phrased "On a scale of 1 to 10, how well did this decision aid support you in answering the questionnaire?" Arguably, because all participants were able to complete the questionnaire, all participants could provide a high subjective rating. Participants were also not exposed to other interface options which could bias them towards a more (or less) favorable rating of their interface due to a lack of available options. In a future version of the study, alternate subjective questions could be asked such as "How easy was it to answer questions with this aid?" or "How easy was it to find the answers for these questions?" Additional subjective questions could target specific features of each decision aid. Because the null hypothesis cannot be rejected for research question 3, it can be said that the novelty of the full-featured interface did not adversely affect participants' subjective experience and in fact the full featured and static experiences provided the subjective experience as familiar formats such as text.

This preliminary evaluation attempted to simulate a realistic scenario where a recently-diagnosed patient must learn about several treatment options in a short amount of time. However, to protect participants from the remote possibility of thinking they were actually being given a real diagnosis a fictitious disease was invented for use with the evaluation. This fictitious disease also prevented any participants from having any experience with the condition. This decision is not

without its detractions though. It's possible that the obvious fiction caused participants to take the questionnaire less seriously thus affecting their performance in some way. It is also possible that the contrived nature of the fictitious disease prevented participants from understanding what information they were presented with. For example, the most common incorrect response across all participants in all groups regardless of which interface was used was mistaking the primary symptom of the fictitious disease for a side effect of one of its treatments. Many participants listed the primary symptom as a side effect in question 1 or claimed that it was the most common side effect in question 2. The use of a more realistic health condition might reduce the number of symptom vs. side effect errors.

While the preliminary evaluation did not show significance for all the posited research questions, the results are encouraging enough to warrant further investigation. Particularly, refinements to the procedure may produce greater differences between treatment conditions.

5.3 Evaluation Limitations

Participants for the pilot studies and preliminary user evaluation were both recruited from convenient populations of University of Maryland students, faculty, and staff which is somewhat limiting to the generalizability of results. However, medical risks and treatment decisions are universal experiences regardless of population. As such, all participants were equally qualified to provide perspective on the TreatmentExplorer prototype. A greater diversity of participants in future evaluations will improve the generalizability of findings related to TreatmentExplorer. Future

evaluations should draw from diverse populations of varying age groups, computer skills, and literacy and education levels.

5.4 Improvements Resulting from the Preliminary User Evaluation

The feedback from the participants formed consensus on several points of improvement which have been implemented and included in the latest version of TreatmentExplorer.

Separating Symptoms from Side Effects in the Legend

The high number of symptom vs. side effect errors in the full-featured and static graphic conditions were unexpected and participant feedback prompted the addition of labels in the legend of TreatmentExplorer. The legend was regrouped such that symptoms would be separate from side effects and both groups labeled.

In-Visualization Labeling

Many participants in the static graphic and full-featured interface groups remarked that while the visualization helped make the treatment differences memorable and made them easier to recall when working on the questionnaire, they relied on the treatment highlight list to inform them of the exact numeric risks associated with each treatment. Others remarked that looking between a separate legend and the visualization took effort. To address both these comments, additional labeling was added to the TreatmentExplorer visualization which provided the name of each colored region within that region of the visualization as well as the proportion of patients represented by that region.

5.5 Preliminary Evaluations Summary

Based on the feedback of domain experts, a preliminary evaluation was designed to begin assessing the efficacy of the TreatmentExplorer prototype. A small pilot study was used to determine the best procedure and the parameters for a larger evaluation. Feedback from the pilot study resulted in minor modifications to the treatment transition animations to better support comparison between treatments. Results from the preliminary evaluation were promising and indicated that participants assisted by the full-featured TreatmentExplorer prototype were better able to understand both primary conditions and treatment side effects better than participants using other versions of the prototype. Additional feedback from the full evaluation has resulted in interface improvements to the full featured TreatmentExplorer prototype. Particularly, the organization of the legend has been improved to distinguish condition symptoms from treatment side effects and additional labels have been included to improve the readability of the visualization. Improvements to the evaluation procedure have been made and a follow-up evaluation is planned to study the effects of the interface improvements.

Chapter 6: Conclusions and Future Work

6.1 Medical Risk Communication and Treatment Exploration Using TreatmentExplorer

TreatmentExplorer is a prototype decision aid designed to allow patients to explore their treatment options and educate themselves on the associated risks of their condition and its treatments. It is capable of showing the likely outcomes of treatment options based on the actual outcomes of similar patients and includes information that would otherwise be scattered amongst disparate decision aids. Though graphics, animation, and guided narration patients with no prior experience can quickly learn both the system and the treatment information it contains. Data personalized to the patient increases the relevancy of TreatmentExplorer's presentation. It can provide a starting point for patients to discuss treatment outcomes with their physicians and other medical professionals on their care team. Patients with greater understanding of their conditions will be more prepared to take greater responsibility for their care and to assert their preferences. Preliminary evaluations show promising results which suggest that TreatmentExplorer is an effective tool for medical risk communication and treatment exploration.

6.2 Guidelines for Interactive Medical Decision Aids

TreatmentExplorer has been evaluated with a number of experts as well as pilot tested with a small group of users and finally compared against other versions of a similar interface through a preliminary evaluation. The feedback and performance of these users has provided insight into best-practices or the design of interactive decision aids for medical risk communication and treatment exploration. The following guidelines have been distilled from the TreatmentExplorer design process and evaluation.

Use of Text

Use Text to Support Graphics

Text is useful for concisely communicating a specific data point. However, text is frequently not memorable and many patients will prefer to use a short bulleted list rather than read large paragraphs of text. Text should therefore be used to support the meaning of graphics and not be the only representation of a data point in a decision aid.

Use Simple Sentence Structure

Patients seeking information from a decision aid may be pressed for time or facing other stressors at the time. They will be interested in obtaining information as quickly as possible. In addition literacy is highly variable among patients who need information from a decision aid. Simple sentence structures should be used so that information can be easily processed by patients who may be rushing or under duress. Communicating only one data point per sentence may be the best option.

Use of Graphics

Label All Graphics

Patients using a decision aid will have varying levels of familiarity with graphics and numerical information. Graphics need to be placed into a context with labels so that patients can quickly understand what information they can expect to learn from a graphic.

Label Graph Pieces

Varying competencies with graphics also implies that patients will need assistance to understand the information conveyed in a graphic. In order to make graphics fully useful to patients, labels should be used to help patients understand how to interpret individual portions of a graphic. Graphic axes may not be enough to help patients understand numeric values conveyed by a graph.

Organize/Order Legends to Show Data Structure

Graphic legends can be used to set up patient expectations about graphics by suggesting what information is portrayed in a graphic. If the data follows an implied structure this structure should be reflected in the legend. For TreatmentExplorer, this means that symptoms of a condition and side effects of treatments should be separated from each other and labeled. There is value in communicating graphic information at a high level through the legend.

Use of Animation

Provide User Control for Animation Speed

Patients will have varying degrees of comfort with computer use and animated graphics. Some patients will be impatient with slow moving graphics while others intimidated by rapidly changing visualizations. The same patient may want to view the same animation at different speeds depending on how well they understand the meaning of the animation. Patients should have the ability to control the speed at which animations occur.

Animate One Object with One Movement per Step

Patients consulting a decision aid are motivated to learn the information it contains. Using simple animations will support the message of a decision aid without being a distraction. Animation should be minimal with one animation used to represent only one data point. In order to allow patients to fully absorb and understand the data point, only one should be expressed at a time suggesting simplistic animations with only one object in motion at a time are best.

Use of Guided Narration

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. This will help patients learn to separate narratives and understand that the data presented in one narration is independent from the data presented in another narration.

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 narration. Patients who are learning a narrative interface will also be less intimidated if their expectations about what will happen next are not defied. Thus, following a consistent order with all treatment narratives will make the narration more accessible.

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 from a narration and feel they only need to see it once. These patients 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 patients who feel they have missed information or want to view the detailed explanation again.

Explain One Data Point per Stage

Following from the recommendations to use simple sentence structures and simple animations, stages of narration should also be simple to allow patients to focus their attention on one data point at a time. Related data points should be grouped into a sequence of stages.

Give Users Control of 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. This may mean providing patients with options to control the speed of an auto-playing narration or pausing and waiting for patients to indicate that they are ready for the next step before the narration advances.

6.3 Limitations

Dataset Limitations

TreatmentExplorer is envisioned to operate as part of a clinical records system where it would have access to large numbers of electronic health records. These patient records would serve as the data source. TreatmentExplorer would aggregate data from patients similar to a given patient in order to provide relevant treatment outcomes. As a prototype, however, it has been built using synthetic data and not tested on actual patient records. This limits the prototype's initial readiness to handle actual patient records. While the data of the preliminary user evaluation was truly fictitious, other data sets have been created based on real-world statistics available from credible organizations. These realistic datasets have been used to guide the development of TreatmentExplorer so that it will be better prepared for the complexity of real-world electronic health records. Future work should expand

TreatmentExplorer to include capabilities to draw data from more realistic sources such as anonymized patient records.

Problem Scope Limitations

Medical risk communication is a vast field with many research questions left to be answered. Continued research is needed to identify the best practices for designing decision aids. TreatmentExplorer is a limited exploration into the efficacy of applying information visualization strategies such as animation, interaction, and guided narration to these diverse fields of research. With its design and early evaluations, TreatmentExplorer has been demonstrated to be a viable tool for continued exploration. Future work will continue to explore subtleties of animation, graphics and text use, and interaction.

6.4 Future Work

The TreatmentExplorer prototype has been established as a first step towards meeting the design goals established in Chapter 1 and supporting patients in making better healthcare decisions. Because of its early stage of development, there are several directions future work could take.

Technical Improvements

Several simplifying assumptions were made to speed the development of TreatmentExplorer. Firstly, it was assumed that patients would be investigating treatments for a single condition with a single primary symptom that must be

alleviated. Secondly, treatments were assumed to have one adverse side effect. Thirdly, the prototype was visually laid out to accommodate a small number of treatment options (though a maximum number of treatment options is not set or enforced anywhere in the code base). Each of these assumptions must be addressed by future work in order to improve the scalability of the prototype. Additionally, future work should attempt to make use of real clinical data instead of the synthetic datasets used during this phase of development. Future research related to the development of TreatmentExplorer is also needed to understand how to best reconcile large, broadly distributed EHR data sources with locally provided EHRs. Small clinics may need to draw from nationally established records systems but it is possibly that patients would be best served if records were weighted to lend more importance to the outcomes of local patients when possible.

Additional Features

A minimal set of desired features has been implemented in this prototype. Feedback from expert reviewers and study participants have provided a number of additional desired features ranging from the inclusion of more measurements (treatment costs, subjective ratings from other patients, ranking treatments by invasiveness, etc.) to additions to the guided narration. A frequently discussed idea is that of spoken narration. Combining text and graphics necessarily splits a patient's attention between reading text and visual processing of graphics. Spoken narration might be a solution to improving patient engagement with TreatmentExplorer (though in its

current implementation as an HTML based web site this will require extensive technical work).

User Experience Improvements

The preliminary work on TreatmentExplorer often sacrificed aesthetics for functionality. Several participants commented on the color choices in particular. While some felt the bold colors made the visualization very memorable, others remarked that they were too striking. The current color selections were chosen to support fully colorblind patients by default at the request of one of the expert reviewers. Future work should improve the balance between inviting, user-pleasing aesthetics and function.

Evaluations

The preliminary evaluations conducted as part of this thesis focused on determining whether or not the TreatmentExplorer prototype showed any indication of improvement over a baseline comparative decision aid. A text-only baseline was chosen to reduce development time needed before evaluations could take place. Other, more complex baselines should be used in future evaluations to address the utility of the TreatmentExplroer visualization specifically. It is possible that a set of coordinated and interactive basic graphs such as bar and line charts could produce similar knowledge gains as TreatmentExplorer. Future evaluations should also seek to better measure how quickly participants are able to learn about treatments as well as test for long-term knowledge gains (possibly by re-testing participants at another

time days or weeks after using TreatmentExplorer). Future evaluations must also seek to test the efficacy of TreatmentExplorer with more representative users. Other populations such as the elderly and patients currently undergoing treatment decisions should be consulted for design feedback as well as recruited for additional controlled studies.

Additional Use Scenarios

This thesis work has focused on addressing the needs of patients making a single healthcare decision. But the TreatmentExplorer prototype could be studied in other contexts in the future. One possibility is to improve informed consent processes in medical research. Potential participants could use TreatmentExplorer to obtain a more complete picture of what they can anticipate experimental procedures to entail. This would be an improvement over current practices which provide minimal information, typically in text format. Another possible use scenario is that of policy decision making. This would involve a different interpretation of "treatment" and represent different populations as they are affected by potential policy changes. Policy makers could use TreatmentExplorer to understand different variations of the same policy on one population or to understand how different populations fare under a single policy.

Conclusion

This thesis makes five contributions, restated below:

1) A multi-dimensional model for describing the content of decision aids based on an extensive literature review
2) A prototype interactive decision aid, TreatmentExplorer, designed to communicate treatment risks and benefits through a guided narration and animation
3) An evaluation of TreatmentExplorer with four experts in health communication
4) A preliminary usability evaluation comparing the performance of TreatmentExplorer against design alternatives
5) Guidelines for interactive decision aids based on the results of these preliminary user evaluations.

Appendix A: User Evaluation Materials

User Evaluation Scenario

Imagine that you have recently been to your doctor's office because you have been experiencing dry, red skin on your hands and feet as well as a mild tingling sensation in your fingers and toes. Your doctor examined you and determined that you have developed a condition known as "Crimson Blot Syndrome". Along with uncomfortable dry skin, this condition carries a serious risk of sudden paralysis and needs to be treated immediately. Fortunately, there are several treatment options but each carries a risk of some side effect. You must now make a decision about which treatment options is the best choice for you.

To help you make this decision, you will be given a decision aid which explains the success rate of each treatment and its side effects. Your doctor has also given you a short list of questions for you to answer about each treatment to help you learn about them. You will need to be able to answer these questions from memory and without the use of the decision aid later. You will be able to use the decision aid for 3 minutes and then be asked to answer as many questions from the list as you can in one minute. You will be able to return to the decision aid and repeat this process as many times as you need to complete the questionnaire.

User Evaluation Questionnaire

- What are the side effects associated with treatments for Crimson Blot Syndrome (CBS)?

- What is the most common side effect of treating CBS?

- What is my risk of getting headaches if I chose Treatment C?

- When can I expect headaches if I chose Treatment B?

- What is my risk of severe rash if I chose Treatment A?

- Which treatment has the highest risk of paralysis?

- What is my risk of paralysis if I chose Treatment C?

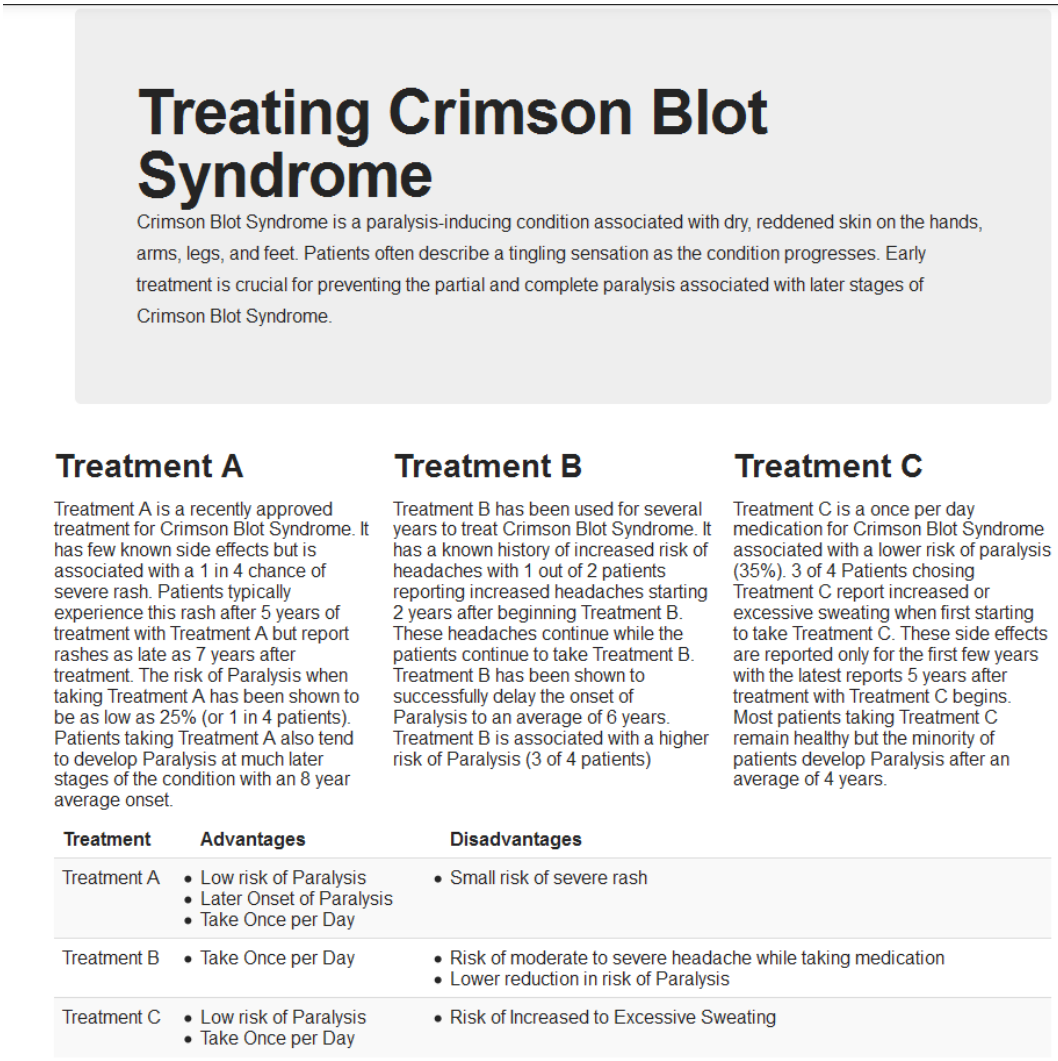
- When can I expect paralysis to occur if I chose Treatment A?

- Which treatment has the earliest average onset of paralysis?

- Which treatment has the most paralysis-free patients?

Appendix B: User Evaluation Interfaces

Text-Based Interface



Treating Crimson Blot Syndrome

Crimson Blot Syndrome is a paralysis-inducing condition associated with dry, reddened skin on the hands, arms, legs, and feet. Patients often describe a tingling sensation as the condition progresses. Early treatment is crucial for preventing the partial and complete paralysis associated with later stages of Crimson Blot Syndrome.

Treatment A

Treatment A is a recently approved treatment for Crimson Blot Syndrome. It has few known side effects but is associated with a 1 in 4 chance of severe rash. Patients typically experience this rash after 5 years of treatment with Treatment A but report rashes as late as 7 years after treatment. The risk of Paralysis when taking Treatment A has been shown to be as low as 25% (or 1 in 4 patients). Patients taking Treatment A also tend to develop Paralysis at much later stages of the condition with an 8 year average onset.

Treatment B

Treatment B has been used for several years to treat Crimson Blot Syndrome. It has a known history of increased risk of headaches with 1 out of 2 patients reporting increased headaches starting 2 years after beginning Treatment B. These headaches continue while the patients continue to take Treatment B. Treatment B has been shown to successfully delay the onset of Paralysis to an average of 6 years. Treatment B is associated with a higher risk of Paralysis (3 of 4 patients)

Treatment C

Treatment C is a once per day medication for Crimson Blot Syndrome associated with a lower risk of paralysis (35%). 3 of 4 Patients chosing Treatment C report increased or excessive sweating when first starting to take Treatment C. These side effects are reported only for the first few years with the latest reports 5 years after treatment with Treatment C begins. Most patients taking Treatment C remain healthy but the minority of patients develop Paralysis after an average of 4 years.

Treatment	Advantages	Disadvantages
Treatment A	• Low risk of Paralysis • Later Onset of Paralysis • Take Once per Day	• Small risk of severe rash
Treatment B	• Take Once per Day	• Risk of moderate to severe headache while taking medication • Lower reduction in risk of Paralysis
Treatment C	• Low risk of Paralysis • Take Once per Day	• Risk of Increased to Excessive Sweating

Figure 37. Screen capture of the text-only interface used in the preliminary user evaluation (See Chapter 5)

Static Graphic Interface

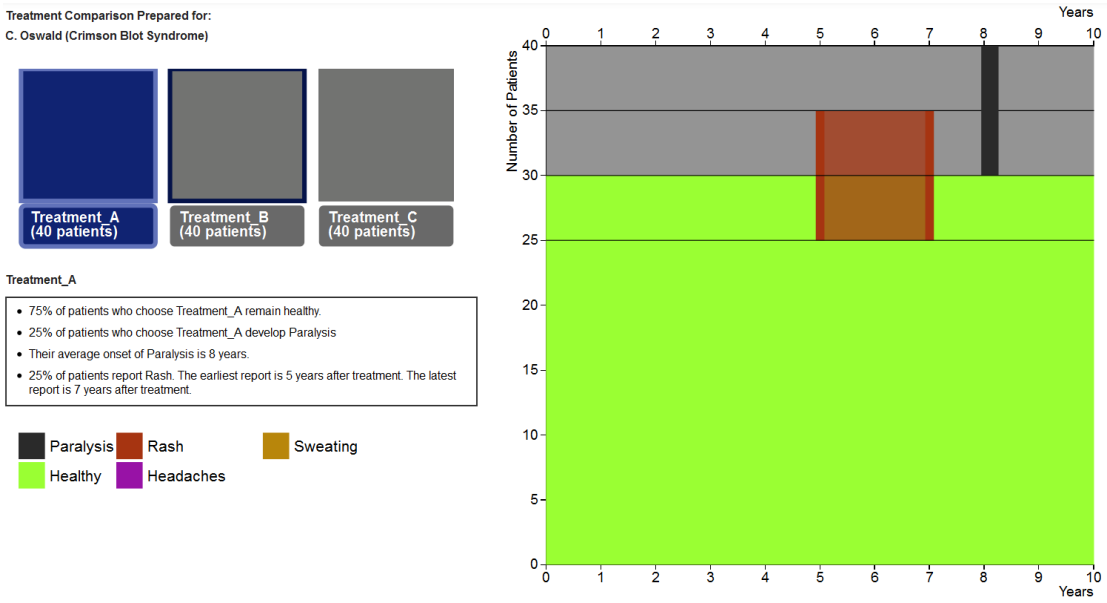


Figure 38. Screen capture of the static graphic interface used in the preliminary user evaluation (See Chapter 5).

Full-Featured Interface

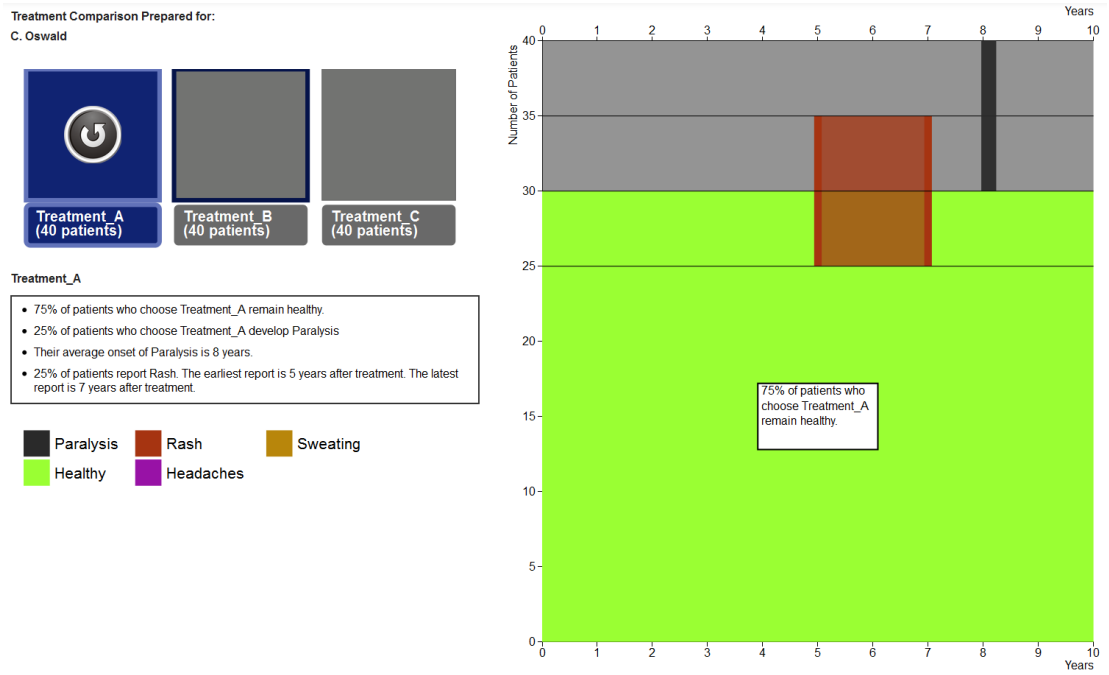


Figure 39. Screen capture of the full-featured interface used in the preliminary evaluation. This interface included animation, guided narration, and interactive explanatory texts.

Appendix C: TreatmentExplorer Code and Demo

Code

The code for the latest prototype version of TreatmentExplorer as completed for this thesis project can be downloaded from:

http://www.cs.umd.edu/~lyndsey/TreatmentExplorer.zip

This package contains the source code and five synthetic datasets which together can be used to run TreatmentExplorer. Included in the five synthetic datasets is the dataset used in the evaluations of TreatmentExplorer.

Demo

For a demonstration of this thesis project using the dataset built for the TreatmentExplorer user evaluations, visit

http://www.cs.umd.edu/hcil/treatmentexplorer/

To change to a different dataset from the included five synthetic datasets, edit the _Scenario variable located in the script section of demo.html.

Using New Datasets for New Applications

The code provided in the TreatmentExplorer.zip has a JavaScript file called data.js which is located in the js directory. This script loads a dataset based on a global variable, _Scenario. _Scenario is created and set in the script section of demo.html at the top level of the TreatmentExplorer directory structure.

To use a new dataset with TreatmentExplorer:

1. **Create Two Files for Your Dataset**: a meta file and a records file. The meta file describes how the events in the records file are to be interpreted by TreatmentExplorer. The records file contains the raw events to be represented.
2. **Edit data.js**: data.js already has a file-loading section of script which reads the contents of a meta- and records-file. You may add to the switch statement which selects from the five synthetic demo datasets or remove the switch leaving only your loading code.
3. **(Optional) Edit demo.html**: demo.html is intended to demonstrate all versions of TreatmentExplorer and to provide an easy mechanism. You may remove versions of the prototype from display by editing demo.html and removing the unwanted tags.

Works Cited

[1] D. Feldman-Stewart and M. D. Brundage, "Challenges for designing and implementing decision aids," *Patient Education and Counseling,* vol. 54, no. 3, pp. 265-273, 2004.

[2] A. J. Dillar, L. Scherer, P. A. Ubel, D. M. Smith, B. J. Zikmund-Fisher, J. B. McClure, S. Greene, A. Stark and A. Fegerlin, "Breast Cancer Anxiety's Assisations with Responses to a Chemoprevention Decision Aid," *Social Science & Medicine,* vol. 77, pp. 13-19, 2013.

[3] A. Edwards, J. Gray, A. Clarke, J. Dundon, G. Elwyn, C. Gaff, K. Hood, R. Iredale, S. Sivell, C. Shaw and H. Thornton, "Interventions to improve risk communication in clinical genetics: Systematic review," *Patient Education and Counseling,* vol. 71, no. 1, pp. 2-25, 2008.

[4] S. M. Brown, J. O. Culver, K. E. Osann, D. J. MacDonald, S. Sand, A. A. Thornton, M. Grant, D. J. Bowen, K. A. Metcalfe, H. B. Burke and others, "Health literacy, numeracy, and interpretation of graphical breast cancer risk estimates," *Patient Education and Counseling,* vol. 83, no. 1, pp. 92-98, 2011.

[5] D. R. Timmermans, "Prenatal screening and the communication and perception of risks," in *Gynaecology, Obstetrics, and Reproductive Medicine in Daily Practice: Proceedings of the 15th Congress of Gynaecology, Obstetrics and Reproductive Medicine*, 2005.

[6] A. E. Waters, W. H. Sullivan, W. Nelson and W. B. Hesse, "What Is My Cancer

Risk? How Internet-Based Cancer Risk Assessment Tools Communicate Individualized Risk Estimates to the Public: Content Analysis," *J Med Internet Res,* vol. 11, no. 3, p. e33, 2009.

[7] J. S. Ancker, Y. Senathirajah, R. Kukafka and J. B. Starren, "Design Features of Graphs in Health Risk Communication: A Systematic Review," *Journal of the American Medical Informatics Association,* vol. 13, no. 6, pp. 608-618, 2006.

[8] A. Fagerlin, P. A. Ubel and D. M. Smith, "Making Numbers Matter: Present and Future Research in Risk Communication," *American Journal of Health Behavior,* vol. 31, no. Suppl 1, pp. s47-256, 2007.

[9] R. L. Sudore and D. Schillinger, "Interventions to improve care for patients with limited health literacy," *Journal of clinical outcomes management: JCOM,* vol. 16, no. 1, p. 20, 2009.

[10] J. K. Belkora, S. Volz, A. E. Teng, D. H. Moore, M. K. Loth and K. R. Sepucha, "Impact of decision aids in a sustained implementation at a breast care center," *Patient Education and Counseling,* vol. 86, no. 2, pp. 195-204, 2012.

[11] A. K. Ghosh and K. Ghosh, "Translating evidence-based information into effective risk communication:," *Journal of Laboratory and Clinical Medicine,* vol. 145, no. 4, pp. 171-180, 2005.

[12] A. M. O'Connor, F. Legare and D. Stacey, "Risk communication in practice: the contribution of decision aids," *BMJ,* vol. 327, no. 7417, pp. 736-740, 2003.

[13] J. Jones, N. R. Shah, C. A. Bruce and W. F. Stewart, "Meaningful use in practice: using patient-specific risk in an electronic health record for shared

decision making," *American Journal of Preventive Medicine,* vol. 40, no. 5, pp. s179-s186, 2011.

[14] H. L. Bekker, J. Hewison and J. G. Thornton, "Understanding why decision aids work: linking process with outcome," *Patient Education and Counseling,* vol. 50, no. 3, pp. 323-329, 2003.

[15] J. K. Belkora, M. K. Loth, S. Volz and H. S. Rugo, "Implementing decision and communication aids to facilitate patient-centered care in breast cancer: A case study," *Patient Education and Counseling,* vol. 77, no. 3, pp. 360-368, 2009.

[16] PatientsLikeMe, "Patients Like Me," PatientsLikeMe, 15 August 2013. [Online]. Available: http://www.patientslikeme.com/. [Accessed 15 August 2013].

[17] B. Dervin, "Sense-Making Theory and Practice: an Overview of User Interests in Knowledge Seeking and Use," *Journal of Knowledge Management,* vol. 2, no. 2, pp. 36-47, 1998.

[18] B. Dervin, L. Foreman-Wernet and E. Lauterback, Sense-Making Methodology Reaser: Selected Writings of Brenda Dervin., Cresskill, NJ: Hampton Press, Inc., 2003.

[19] B. Tversky, "Visualizing Thought," *Topics in Cognitive Science,* vol. 3, pp. 499-535, 2011.

[20] M. Bunge, I. Muhlhauser, A. Steckelberg and others, "What constitutes evidence-based patient information? Overview of discussed criteria," *Patient education and counseling,* vol. 78, no. 3, p. 316, 2010.

[21] A. G. Levy, S. S. Sonnad, J. E. Kurichi, M. Sherman and K. Armstrong,

"Making Sense of Cancer Risk Calculators on the Web," *Journal of General Internal Medicine,* vol. 23, no. 3, pp. 229-235, 2008.

[22] A. Fagerlin, B. J. Zikmund-Fisher and P. A. Ubel, "Helping Patients Decide: Ten Steps to Better Risk Communication," *Journal of the National Cancer Institute,* vol. 103, no. 19, pp. 1436-1443, 2011.

[23] B. J. Zikmund-Fisher, A. Fagerlin and P. A. Ubel, "Improving Understanding of Adjuvant Therapy Options by Using Simpler Risk Graphics," *Cancer,* vol. 113, no. 12, pp. 3382-3390, 2008.

[24] P. K. J. Han, W. M. P. Klein and N. K. Arora, "Varieties of Uncertainty in Health Care: A Conceptual Taxonomy," *Medical Decision Making,* vol. 31, no. 6, pp. 828-838, 2011.

[25] P. K. Han, W. M. Klein, B. Killam, T. Lehman, H. Massett and A. N. Freedman, "Representing randomness in the communication of individualized cancer risk estimates: effects on cancer risk perceptions, worry, and subjective uncertainty about risk," *Patient Education and Counseling,* vol. 86, no. 1, January 2012.

[26] S. T. Hawley, B. Zikmund-Fisher, P. Ubel, A. Jancovic, T. Lucas, A. Fagerlin and others, "The impact of the format of graphical presentation on health-related knowledge and treatment choices," *Patient education and counseling,* vol. 73, no. 3, p. 448, 2008.

[27] Z. Hildon, D. Allwood and N. Black, "Making data more meaningful: Patients' views of the format and content of quality indicators comparing health care providers," *Patient Education and Counseling,* vol. 88, no. 2, pp. 298-304, 2012.

[28] G. J. Hollands and T. M. Marteau, "The impact of using visual images of the body within a personalized health risk assessment: An experimental study," *British Journal of Health Psychology,* 2012.

[29] B. M. Chaudry, K. H. Connelly, K. A. Siek and J. L. Welch, "Mobile interface design for low-literacy populations," in *Proceedings of the 2nd ACM SIGHIT International Health Informatics Symposium*, Miami, FL, USA, 2012.

[30] N.-J. Lee, S. Bakken and others, "Development of a prototype personal digital assistant-decision support system for the management of adult obesity," *International journal of medical informatics,* vol. 76, p. s281, 2007.

[31] J. Finkelstein and J. Wood, "Information technology for continuous patient health education," in *Proceedings of the 2011 conference on Information technology education*, West Point, NY, USA, 2011.

[32] A. V. Dhukaram and C. Baber, "An Approach to Designing Interactive Decision Aid for Cardiac Patients," in *25th British Computer Society Conference on Human-Computer Interaction*, Swinton, UK, 2011.

[33] G. Leroy, "Improving Consumer Health Literacy with Information Technology," *Encyclopedia of Healthcare Information Systems,* 2008.

[34] M. Agrawala, W. Li and F. Berthouzoz, "Design Pricinples for Visual Communication," *Communications of the ACM,* vol. 54, no. 4, pp. 60-69, 2011.

[35] R. Kosara and J. Mackinlay, "Storytelling: the Next Step for Visualization," *Computer,* vol. 46, no. 5, pp. 44-50, 2013.

[36] E. Segel and J. Heer, "Narrative Visualization: Telling Stories with Data," *IEEE*

Transactions in Visualization & Computer Graphics (Proc. InfoVis.), 2010.

[37] J. Hullman, N. Diakopoulos and E. Adar, "Contextifier: Automatic Generation of Annotated Stock Visualizations," in *Conference on Human Factors in Computing Systems (CHI)*, 2013.

[38] A. Perer and D. Gotz, "Data-Driven Exploration of Care Plans for Patients," in *CHI 2013 Extended Abstracts*, Paris, France, 2013.

[39] T. Longman, R. M. Turner, M. King and K. J. McCaffery, "The effects of communicating uncertainty in quantitative health risk estimates," *Patient Education and Counseling,* vol. 89, no. 2, pp. 252 - 259, 2012.

[40] M. Price, R. Cameron and P. Butow, "Communicating risk information: The influence of graphical display format on quantitative information perception—Accuracy, comprehension and preferences," *Patient Education and Counseling,* vol. 69, pp. 121-128, 2007.

[41] B. Xie, I. Watkins and M. Huang, "Making web-based multimedia health tutorials senior-friendly: design and training guidelines," in *Proceedings of the 2011 iConference*, Seattle, WA, USA, 2011.

[42] A. Fagerlin, B. J. Zikmund-Fisher and P. A. Ubel, ""If I'm better than average, then I'm ok?": Comparative information influences beliefs about risk and benefits," *Patient education and counseling,* vol. 69, no. 1, pp. 140-144, 2007.

[43] P. A. Ubel, D. M. Smith, B. J. Zikmund-Fisher, H. A. Derry, J. McClure, A. Stark, C. Wiese, S. Greene, A. Jankovic and A. Fagerlin, "Testing whether decision aids introduce cognitive biases: Results of a randomized trial," *Patient*

Education and Counseling, vol. 80, no. 2, pp. 158 - 163, 2010.

[44] A. Edwards, R. Thomas, R. Williams, A. L. Ellner, P. Brown and G. Elwyn, "Presenting risk information to people with diabetes: evaluating effects and preferences for different formats by a web-based randomised controlled trial," *Patient education and counseling,* vol. 63, no. 3, pp. 336-349, 2006.

[45] J. S. Ancker, C. Chan and R. Kukafka, "Interactive Graphics for Expressing Health Risks: Development and Qualitative Evaluation," *Journal of Health Communication,* vol. 14, no. 5, pp. 461-475, 2009.

[46] C.-L. Chi, W. N. Street, J. G. Robinson and M. A. Crawford, "Individualized patient-centered lifestyle recommendations: An expert system for communicating patient specific cardiovascular risk information and prioritizing lifestyle options," *Journal of Biomedical Informatics,* 2012.

[47] H. Kharrazi, "Improving healthy behaviors in type 1 diabetic patients by interactive frameworks," in *AMIA Annual Symposium Proceedings*, 2009.

[48] C.-A. Waldron, T. van der Weijden, S. Ludt, J. Gallacher and G. Elwyn, "What are effective strategies to communicate cardiovascular risk information to patients? A systematic review," *Patient education and counseling,* vol. 82, no. 2, pp. 169-181, 2011.

[49] National Center for Chronic Disease Prevention and Health Promotion, *National Diabetes Fact Sheet,* 2011.

[50] B. J. Zikmund-Fisher, P. A. Ubel, D. M. Smith, H. A. Derry, J. B. McClure, A. Stark, R. K. Pitsch and A. Fagerlin, "Communicating side effect risks in a

tamoxifen prophylaxis decision aid: the debiasing influence of pictographs," *Patient education and counseling,* vol. 73, no. 2, pp. 209-214, 2008.

[51] "National Cancer Institute," National Institutes of Health, 2013. [Online]. Available: www.cancer.gov. [Accessed 1 May 2013].

[52] P. Denig, M. Dun, J. Schuling, F. Haaijer-Ruskamp and J. Voorham, "The effect of a patient-oriented treatment decision aid for risk factor management in patients with diabetes (PORTDA-diab): study protocol for a randomised controlled trial," *Trials,* vol. 13, no. 1, p. 219, 2012.

[53] J. S. Ancker, E. U. Weber and R. Kukafka, "Effects of game-like interactive graphics on risk perceptions and decisions," *Medical Decision Making,* vol. 31, no. 1, pp. 130-142, 2011.

[54] Hughes Risk Apps, "User Guide 2.6.x," Hughes Risk Apps, 2009.

[55] B. Drohan, E. M. Ozanne and K. Hughes, "Electronic Health Records and the Management of Women at High Risk of," *The Breast Journal,* vol. 15, pp. s46-s55, 2009.

[56] M. W. Kreuter and R. J. Wray, "Tailored and Targeted Health Communication: Strategies for," *American Journal of Health Behavior,* vol. 27, no. s3, pp. s227-s232, 2003.

[57] B. J. Zikmund-Fisher, M. Dickson and H. O. Witteman, "Cool but counterproductive: interactive, Web-based risk communications can backfire," *Journal of medical Internet research,* vol. 13, no. 3, 2011.

[58] M. Monroe, R. Lan, K. Wongsuphasawat, C. Plaisant, B. Shneiderman, J.

Millstein and S. Gold, "EventFlow: Exploring Point and Interval Event Temporal Patterns," Human Computer Interaction Lab, University of Maryland College Park, 23 May 2013. [Online]. Available: http://www.cs.umd.edu/hcil/eventflow/. [Accessed 8 July 2013].

[59] National Cancer Institute, "What You Need to Know About Breast Cancer," 26 September 2012. [Online]. Available: http://www.cancer.gov/cancertopics/wyntk/breast/page7. [Accessed 30 June 2013].

[60] M. Bostock, "Data Driven Documents," 2012. [Online]. Available: http://d3js.org/. [Accessed 28 August 2013].

[61] "Susan G. Komen," Suan G. Komen Breast Cancer Foundation, 8 April 2013. [Online]. Available: http://ww5.komen.org/Default.aspx. [Accessed 1 May 2013].

[62] Consumer Reports Health, "Treating Type 2 Diabetes: The Oral Diabetes Drugs, Comparing Effectiveness, Safety, and Price," December 2012. [Online]. Available: http://www.consumerreports.org/health/best-buy-drugs/type2diabetes.htm. [Accessed 30 June 2013].

[63] B. L. St. Jean, *Information Behavior of People Diagnosed with a Chronic Serious Health Condition: A Longitudinal Study,* University of Michigan, 2012, p. 336.

[64] I. M. Lipkus, "Numeric, Verbal, and Visual Formats of Conveying Health Risks: Suggested Best Practices and Future Recommendations," *Medical Decision*

Making, vol. 27, no. 5, pp. 696-713, 2007.

[65] N. Kong and M. Agrawala, "Graphical Overlays: Using Layered Elements to Aid Chart Reading," *IEEE Transactions on Visualization and Computer Graphics,* vol. 18, no. 12, pp. 2631-2638, 2012.

[66] C. D. Hundhausen, S. A. Douglas and J. T. Stasko, "A Meta-Study of Algorithm Visualization Effectiveness," *Journal of Visual Languages and Computer,* vol. 13, pp. 259-290, 2002.