Diagnoses, Decisions, and Outcomes: Web Search as Decision Support for Cancer

Michael J. Paul*
Johns Hopkins University
3400 N. Charles Street
Baltimore, MD 21218
mpaul@cs.jhu.edu

Ryen W. White
Microsoft Research
One Microsoft Way
Redmond, WA 98052
ryenw@microsoft.com

Eric Horvitz
Microsoft Research
One Microsoft Way
Redmond, WA 98052
horvitz@microsoft.com

ABSTRACT

People diagnosed with a serious illness often turn to the Web for their rising information needs, especially when decisions are required. We analyze the search and browsing behavior of searchers who show a surge of interest in prostate cancer. Prostate cancer is the most common serious cancer in men and is a leading cause of cancer-related death. Diagnoses of prostate cancer typically involve reflection and decision making about treatment based on assessments of preferences and outcomes. We annotated timelines of treatment-related queries from nearly 300 searchers with tags indicating different phases of treatment, including decision making, preparation, and recovery. Using this corpus, we present a variety of analyses toward the goal of understanding search and decision making about treatments. We characterize search queries and the content of accessed pages for different treatment phases, model search behavior during the decision-making phase, and create an aggregate alignment of treatment timelines illustrated with a variety of visualizations. The experiments provide insights about how people who are engaged in intensive searches about prostate cancer over an extended period of time pursue and access information from the Web.

Categories and Subject Descriptors: H.2.8 [Database management]: Database applications—*data mining*

Keywords: medical search, decision making, cancer

1. INTRODUCTION

Upon diagnosis of a major illness, people frequently turn to the Web for information about the course and prognosis of the disease, and to better understand treatments and outcomes [6, 19, 27, 36]. We seek to understand the use of Web search as a medical decision support system by patients who have been diagnosed with a significant disease. In particular, we study the use of Web search to support the decisions of searchers who show salient signs of having been recently diagnosed with prostate cancer. We aim to characterize and enhance the ability of Web search engines to provide decision support for different phases of the illness.

We focus on prostate cancer for several reasons. Prostate cancer is typically slow-growing, leaving patients and physicians with time to reflect on the best course of action. There is no medical consensus about the best treatment option, as the primary options all have similar mortality rates, each with their own tradeoffs and side effects [24]. The choice of treatments for prostate cancer is therefore particularly sensitive to consideration of patients' preferences about outcomes and to the assessed likelihoods of achieving different outcomes. For these reasons, prostate cancer is an "archetypal condition" for the use of decision aids [25].

Guidance on decisions is provided primarily via consultation with a patient's physician and larger care team. Formal decision-making materials may be available through patients' physicians. However, Web search is becoming a common supplement to traditional decision aids [29, 4]. In a 2011 survey of prostate cancer patients, the Internet was found to be the second most common information source for making treatment decisions, after "doctor's recommendation" [37]. An earlier survey found that more than half of the respondents who had searched the Web prior to deciding on a treatment reported that information reviewed online had influenced their decision [32]. Search for information on major illnesses includes information gathering about treatments and outcomes, including characterization of uncertainties about outcomes for different treatments. Beyond providing information about illness, search and retrieval sessions serve as opportunities for interactive assessment of preferences about outcomes and risk. Thus, we are particularly interested in how people learn about different therapeutic options, including how they progress over time in navigating trees of possibilities.

To pursue insights about search and retrieval as medical decision support, we adapt methods developed in a prior study on the use of Web search for pursuing information on breast cancer [31]. In that work, classifiers were constructed from annotated logs to infer that a searcher had likely received a diagnosis of breast cancer. An ontology of different kinds of information needs was introduced to characterize the dynamics of information-seeking for a large set of searchers aligned by an inferred date of diagnosis. We harness similar methods to seek insights about prostate cancer, but focus on decision making and other search activities aligned with different phases of treatment. Specifically, we

*Work conducted during a Microsoft Research internship.

WWW 2015, May 18–22, 2015, Florence, Italy.
ACM 978-1-4503-3469-3/15/05.
http://dx.doi.org/10.1145/2736277.2741662.

perform studies that extend previous work in a number of ways and make the following contributions:

- We create a hierarchy of treatments and associated search terms, as well as an annotated corpus of 272 timelines of treatment search queries.
- We present a characterization of different phases of treatment search, exemplified by n-grams from queries and the content of visited webpages associated with each phase in the annotated corpus. We further characterize the phases and their progression over time by creating a multiple sequence alignment of timelines. We create a series of visualizations illustrating how the phases and queries evolve over time, based on the alignment.
- Focusing specifically on queries tagged as pursuit of decision support, we analyze the number and specificity of treatments that are searched over time, the treatments that co-occur in comparative searches, and transitions among the treatments at the focus of attention in successive queries. We identify and visualize typical sequences of treatment queries traversed by searchers by applying a spanning-tree algorithm to a graph of query transitions.

After presenting analyses and findings on the use of the Web for decision support, we discuss implications and directions for the design of search and retrieval systems that help people to better understand diagnoses and treatment decisions moving forward.

2. RELATED WORK

The Web is an important source of health-related information for many people. According to a 2013 survey, 59% of American adults had used the Web to find health information in the year preceding the survey, 35% of those adults engaged in self-diagnosis, and over half of these self-diagnosing searchers then discussed the matter with a clinician [14]. Despite the potential benefits, concerns have been raised about the quality of online health information [9] including cancer information [17]. A survey of oncologists noted that Web use can "simultaneously make patients more hopeful, confused, anxious, and knowledgeable." [19] In a large-scale survey of the use of search for self-diagnosis, White and Horvitz [42] found that almost 40% of participants experienced increased anxiety from searching health information online.

Such challenges highlight the criticality of understanding how patients use the Web, including the nature and dynamics of queries, and the content delivered in response to queries. To better understand how people pursue health information, studies have examined online health search using a variety of methods, including interviews [33], surveys [38], and analyses of large-scale search log data [3, 1, 20, 5, 43]. Search logs analyses can provide insights about how people use search engines [41], predict future search actions and interests [23, 10, 11], and detect real-world events and activities [34]. Applications of search log data in the health and medical domains include the detection of influenza [16] and the discovery of side effects of medications [44]. Studies of online information-seeking for cancer [6, 19] have characterized how cancer patients use Web resources and have proposed patient taxonomies describing how people employ retrieved health information. In the realm of search and retrieval on cancer, Bader et al [2] categorized cancer-related search queries from three months in 2001. More recently, Ofran et al. [27] used search log data to identify five phases of cancer search activity and showed that the phases mirror those associated with coping and grief that had been previously documented in the literature.

Information access about treatment options has been found to be important for cancer patients facing difficult therapeutic decisions. Cancer patients typically seek access to all relevant information [35, 15]. Valuable information about treatments and outcomes can come via reviewing testimonials from those afflicted with similar conditions [28]. Beyond providing information about treatment options, sharing health experiences online can help people to feel supported and to better engage with health services [46]. Formal decision aids [26] have been used to help patients decide among treatment alternatives, by providing information that helps to resolve or to clarify their uncertainties [25]. A study found that working with decision aids for prostate cancer could influence decisions about treatment strategy [39].

3. TREATMENT TIMELINES

We focus on the analysis of **treatment timelines** extracted from search query logs. A treatment timeline for prostate cancer is a time-stamped sequence of search queries that contain terms pertaining to prostate cancer treatment, where sequences of queries are associated with unique, anonymized user identifiers. The set of treatment timelines was created by first identifying relevant search histories via a series of filters that are described in detail in Section 4. We tag the queries in the treatment timelines with labels indicating the assessed phase of treatment, including whether the searcher appears to be seeking information for an initial or a follow-on, secondary treatment, and whether the queries appear to be aimed at seeking information decisions about a treatment, preparation for a chosen treatment, or recovery from a treatment, as described in Section 4.4.

In Section 5, we present a series of experimental analyses of the treatment timelines. We characterize the content associated with different phases of information pursuit and show how these phases evolve over time for a set of searchers who are temporally aligned by inferred date of diagnosis.

3.1 Treatment Hierarchy

In order to extract treatment timelines from search histories, we identify a set of search terms that searchers tend to use to refer to prostate cancer treatment options. As queries range from very general (e.g. "cancer treatment") to very specific (e.g. "low dose radiation seed implants"), we organized the terms into a hierarchical ontology of known treatments, moving from broad categories down to detailed therapies. Such a treatment hierarchy enables us to analyze treatment timelines in terms of categories of treatments as well as the raw text used to describe options. We can characterize the different types of treatments that are searched and the degree of specificity of queries, based on the depth of the query terms in the hierarchy.

Table 1 shows the treatment hierarchy and the terms associated with each category. The treatment hierarchy was constructed by an extensive review of the literature on the management of prostate cancer. Categories in the treatment hierarchy reflect current standard treatment options for prostate cancer. "Observation" is a treatment option that refers to a decision to forgo treatment for the time being; the two common methods of observation are typically

Level 0	Level 1	Level 2	Level 3	Search terms
Treatment	–	–	–	treatment(s)
Treatment	Surgery	–	–	surgery, prostatectomy, prostate removal, remove prostate
Treatment	Surgery	Open	–	open [Surgery]
Treatment	Surgery	Laparoscopic	–	laparoscopic, minimally invasive
Treatment	Surgery	Laparoscopic	Robotic	robot, robotic, da( )vinci
Treatment	Radiation	–	–	radiation
Treatment	Radiation	Brachytherapy	–	brachytherapy, brachy, seed(s)
Treatment	Radiation	Brachytherapy	LDR	low dose [Brachytherapy], ldr
Treatment	Radiation	Brachytherapy	HDR	high dose [Brachytherapy], hdr
Treatment	Radiation	External	–	external [Radiation], external beam, ebrt
Treatment	Radiation	External	3DRT	3drt, 3dcrt, conformal
Treatment	Radiation	External	IMRT	imrt, intensity-modulated, igrt, calypso
Treatment	Radiation	External	SBRT	sbrt, stereotactic body, cyber( )knife, gamma( )knife, x-knife
Treatment	Radiation	External	Proton	proton, pencil beam
Treatment	Radiation	Drugs	Radium 223	radium 223, radium dichloride, xofigo
Treatment	Hormone therapy	–	–	hormone/hormonal therapy, hormone/hormonal treatment
Treatment	Hormone therapy	LHRH	...	*various hormone-therapeutic drugs are categorized*
Treatment	Hormone therapy	Anti-Androgen	...	*various hormone-therapeutic drugs are categorized*
Treatment	Chemotherapy	–	–	chemotherapy, chemo
Treatment	Chemotherapy	Drugs	...	*various chemotherapeutic drugs are categorized*
Treatment	HIFU	–	–	hifu, high-intensity
Treatment	Cryotherapy	–	–	cryotherapy, cryosurgery, cryoablation, cryo
Treatment	Observation	None	–	no treatment, without treatment
Treatment	Observation	Waiting	–	waiting [Treatment]
Treatment	Observation	Surveillance	–	active surveillance

Table 1: Four-level treatment hierarchy. For space, we do not display several specific drugs at level 3. Brackets indicate co-occurrence constraints with the term. For example "open [Surgery]" indicates that the term "open" is considered as an entry only if it occurs in the same query as a term in the Surgery category.

referred to as "watchful waiting" and "active surveillance" by clinicians. These options may be recommended when the cancer is low grade and the risks of treatment are assessed as outweighing the risks of the disease [21]. HIFU (high-intensity focused ultrasound) and cryotherapy are newer experimental treatments that are less common, though still found to be frequently searched. Another type of treatment, immunotherapy (affecting the patient's immune response), is rarely found in queries in our dataset, and is not included in our analysis.

4. DATASET CREATION

We now review the extraction and tagging of data for the study, including the formulation of relevant search terms, labeling of searchers as likely facing prostate cancer decisions, and annotating phases of long-term timelines.

4.1 Ontology of Relevant Terms

To identify relevant search queries for the study, we relied on a manually-curated ontology of terms of interest. Terms are organized in a four-level hierarchy and include terms related to screening methods, diagnosis (e.g. "biopsy"), cancer staging and grading information (e.g. "stage II", "low grade"), and various treatment options. The formulation of relevant query terms is similar to efforts to construct an ontology for search and retrieval for breast cancer [31], and is distinct from the treatment ontology in Table 1.

4.2 Search and Browsing Logs

The data for this study comes from a proprietary set of anonymized logs from consenting users of the Internet Explorer Web browser. The data includes time-stamped search queries issued through the browser (primarily via interactions with the Microsoft Bing search engine) and time-stamped webpage visits. Each log entry includes a unique anonymized user identifier. The data spans an 18-month period from March 2013 to August 2014.

The initial dataset consists of logs collected from users whose queries include the bigram "prostate cancer" at least three times during this time period. This policy was employed as an initial high-recall filter to identify search histories that would likely be relevant to the study. Given our focus on treatment-related search, we filtered the set for search histories including queries containing treatment-related terms listed in Table 1 (excluding the most general term, "treatment"). This extraction procedure yielded a set of 3,066 search histories.

4.3 Experiential vs. Exploratory Searchers

A key initial task with the extracted data is to identify a high-precision set of search histories from the 3,066 candidate histories. Ideally, the resulting focused set would contain only histories of those who actually experienced a diagnosis of prostate cancer—either personally, or via the intensive searching performed with regard to the diagnosis of a close family member or friend. However, we cannot conclude with certainty whether a searcher is in this situation based only on information in the logs. We consider searchers as *experiential* versus less-involved *exploratory* searchers based on an assessment of *sustained* and *focused* interest in prostate cancer. We exclude search histories that are inconsistent with a cancer diagnosis. Following an annotation of logs as being experiential versus exploratory, we train a classifier (described below) on a small set of labeled histories to identify experiential searchers among the set of candidates in an automated manner.

We annotated a sample of 100 histories with binary relevance tags using the criteria outlined above. One of the authors (EH), with formal medical and decision analysis train-

Phase	# histories	# queries
Initial decision	174	2008
Initial preparation	126	882
Initial post-treatment	140	1232
Secondary decision	84	769
Secondary preparation	24	114
Secondary post-treatment	18	85
Total	272	5090

Table 2: Number of search histories and search queries labeled with each phase in the dataset.

ing (including decision analyses for the treatment of prostate cancer), provided insights on the coding criteria. Beyond leveraging knowledge of prostate cancer and its treatment, we considered distinctions that capture sustained and intensive focus of attention. As examples, a search history with a brief burst of interest that fades away is considered negative because the searcher does not show sustained interest. Histories that include searches for many different diseases are also labeled negative, as people grappling with a new diagnosis are likely to focus on a single illness and its challenges and trajectory. We also exclude searchers who appear to be medical professionals, as their queries include such searches as billing codes or instructions for administering a treatment. The goal was to rule out histories that were not plausibly experiential, in order to reduce noise in the analyzed data, even if we cannot know the ground truth.

Each history was independently coded by two annotators (the first author and one of the two other authors). In cases of disagreement, the final labels were resolved after discussion among the annotators. In some cases, annotators did not commit to an initial label, and instead marked a history as ambiguous to be resolved upon discussion. Annotators were shown all anonymized search sessions containing medical terms contained in the ontology (described in Section 4.1), including terms related to diagnosis, staging, and treatment, as well as other terms such as a set of symptoms and names of diseases. In the final set of annotations, 63% of histories were tagged as experiential.

These annotated histories (combined with similar annotations for breast cancer) were used to train a classifier to identify experiential cancer search histories [31]. The classifier uses a variety of lexical, distributional, and temporal pattern features, and is estimated to have 96% precision and 78% recall (AUC .891) from 5-fold cross validation. After applying the classifier to the 3,066 candidate histories, we identified 1,413 experiential histories.

We then applied an additional filter condition to the positively classified histories: Given our focus on decision support and the process of decision making about treatments, we only include in the final dataset those search histories that contain terms indicative of information-gathering and deliberation about treatment: "vs", "better", "best", "pros", "cons", "which", "option(s)", "should i". This final step yielded a corpus of 272 histories, which we then annotated with additional information, described in the next subsection.

4.4 Annotation of Treatment Timelines

Last, we annotated the treatment timelines—the experiential search histories projected down to only those queries containing treatment terms—with richer tags to allow for finer-grained analysis. The queries are tagged as belonging to different phases of the treatment process that we had observed as common patterns in the timelines. We were particularly interested in tagging queries that appeared to indicate decision making, but we also tagged other phases of treatment-related queries for queries that appeared to come before and after a treatment had occurred.

Each query was annotated with two labels. The first is one of three states of the treatment deliberation process:

- **Decision:** Queries that appear to be used to help a searcher decide between or learn more about different treatment options. These queries would sometimes contain explicit indicators that the searcher is considering different options (e.g. "best treatment options", "which is better", "pros and cons"). Queries for many different treatments in the same timeframe are considered decision queries.
- **Preparation:** Queries about a treatment that appears to be scheduled but before the treatment has taken place (e.g. "what to expect"). If a search history focuses on a single treatment for many days (as opposed to exploring multiple options), we consider these queries preparatory.
- **Post-treatment:** Queries that appear to take place after a treatment has commenced or completed. These may seek information about recovery, or queries regarding side effects that are experienced. Some queries include specific timing references (e.g., "3 weeks after surgery") which can help with determining this label.

The second label captures whether the context of a search session is an initial or follow-up treatment:

- **Initial:** The first round treatment that the searcher is considering, typically surgery or radiation.
- **Secondary:** Any treatment that follows an initial treatment, typically adjuvant radiation, hormone therapy, or chemotherapy for more advanced cancer. Secondary status is often clear from queries with explicit indicators like the term "adjuvant" or including such terms as "after surgery".

The cross-product of these tags defines a total of six different phases of treatment-related search.

The queries were categorized assuming that they were issued by patients experiencing cancer and in reference to treatment for a specific patient, who may be the searcher himself or a family member deeply involved in decisions about the illness. The goal was to group the search activity based on common characteristics that are observable in the data and consistent with a typical patient timeline.

Ambiguous queries were tagged with multiple phases. If more than one phase was included, the phases were ranked based on which phase the annotator believed was most likely. Queries that did not fit these phase labels were not annotated. The tagging of the phases was done by the first author and a second professional annotator formally trained in linguistics. The first annotator reviewed the secondary annotations to ensure consistency of tagging procedures.

Table 2 provides the number of queries labeled with each phase as well as the number of search histories with at least one query labeled with the phase. In the case of ambiguous annotations, only the most likely label was counted in this table. Additionally, the 272 histories contained 33,945 queries that were not annotated with these phase labels.

5. ANALYZING TREATMENT PATTERNS

We now discuss methods and results on characterizing the different phases of treatment as well as the dynamics of the progression of searchers through the phases over time.

5.1 Phase Characterization

We characterize different annotated phases of searcher timelines by identifying n-grams—from search queries and webpage bodies—and domain names that are most associated with retrieval in each phase. We wish to identify features that are *salient*—both probable and representative of the phase [8]. We achieve this with a two-component mixture model that mixes phase-specific feature distributions with a phase-independent background distribution which accounts for common features that are not representative of any particular phase [45].

With this model, the probability of a feature i (an n-gram or a domain name) in the text associated with phase k is a mixture of two parameters θ:

$$P(\text{feature} = i|\text{phase} = k) = \lambda\theta_i^B + (1-\lambda)\theta_i^k \qquad (1)$$

Each θ^k is a distribution over features specific to the phase k, while θ^B is a background distribution over features independent of phase, and λ is the mixing weight. Then, we examine the θ^k distributions to find salient feature associations with phase k, because these distributions will put the most mass on n-grams that are probable within the phase but are not better explained by the background distribution.

Experiment Details.

We created a model for each class of features (search and page n-grams, and page domain names) from the annotated queries. We modeled bigrams and trigrams. Features derived from webpage content include the pages visited during the same search session following an annotated query (i.e., on the post-query navigational trail), which is possible since we used browser logs for our analysis. To extract page content, we used the method described in [40], which extracts lines of HTML such that the ratio of text tokens to tags tokens is at least one standard deviation above the mean ratio. This is a simple heuristic for identifying the core content in the page, rather than supplementary text such as navigation menus and page footers.

The parameter posteriors for the mixture model are inferred using Gibbs sampling. We averaged the parameter values from every 100 sampling iterations, collected from 4000 iterations after a 2000-iteration burn-in. The θ parameters were given Dirichlet(0.01) priors, and λ was given a Beta(9000, 1000) prior, so that high values of λ (favoring the background distribution) are *a priori* more likely, resulting in stronger feature associations.[1]

When encoding feature values for the mixture model (i.e., the number of times each feature is observed within each class label), we used fractional values when annotators included multiple labels for a query. Since annotators provide labels in order of likelihood, we set the fraction to be twice as high for the more probable label. For example, two labels results in counts of $\frac{2}{3}$ and $\frac{1}{3}$ for features in that query.

[1]We used existing software for cross-collection Latent Dirichlet Allocation (ccLDA) [30], a topic model that learns topics for multiple collections of text as well as collection-independent background topics. This two-component model is a special case of ccLDA with only one topic.

Results.

Table 3 shows the top features for each phase, displaying the highest probability n-grams and domain names under each θ^k. For space and simplicity, the table only displays bigrams and not trigrams.

We see that the general query "treatment options" is associated with both the initial and secondary decision phases. The initial decision phase is also associated with queries containing the trigram "pros and cons", as well as explicit comparative n-grams like "surgery vs radiation". The decision phase also has a high probability of queries for "active surveillance" and "watchful waiting"—options for non-treatment that do not apply to later phases, once treatment has started. The page n-grams are similar, with general terms regarding treatment options. The bigrams "clinical trial(s)" are highly probable in the secondary decision pages.

A top query trigram for the initial preparation phase is "what to expect", while many of the top query n-grams for initial post-treatments are variants of "after surgery" or "after prostatectomy". Searchers look for general recovery information, as well as information about performing various activities after treatment (e.g. "sex after") and treatment side effects (e.g. "incontinence after"). The n-grams from retrieved pages for these two phases both include a number of treatment side effects ("erectile dysfunction", "urinary incontinence"), as well as n-grams containing the word "catheter".

The top search n-grams for all secondary phases include medications used in hormonal therapies (e.g. "lupron", "zytiga"), and "adjuvant radiation", referring to a type of radiation that is given after the initial surgery. The top page n-grams for the secondary phases have terms related to drugs and their side effects.

We see that `youtube.com` is the top domain name for initial preparation. In general, videos are associated with the initial preparation and decision phases. 26.2% and 28.2% of users visited pages with "video(s)" in the title or URL during initial preparation and decision sessions, while only 8.3% of users visited such pages during the initial post-treatment sessions. Almost no users visited video pages during the secondary phases.

We also observe that the initial recovery phase contains many n-grams containing first person pronouns in the top page content n-grams, and `cancerforums.net` is the top domain name. There is an association of forums with the initial post-treatment phase. 48.8% of users visited pages with "forum(s)", "discussion(s)", or "community" in the title or URL during this phase. This is substantially higher than the percentage during the initial decision (35.9%) or secondary decision (37.1%) phases, which had the next highest percentages of such visits.

The top domain names associated with the background distribution in the mixture model (the distribution independent of phase) are `webmd.com`, `cancer.org`, `cancer.gov`, `ask.com`, and `ehow.com`. These are the most probable domains visited, even though many of these are not associated with particular phases, and so do not appear in Table 3.

5.2 Evolution of Queries on Treatments

We now focus specifically on the "initial decision" phase, with the goal of seeking an understanding of the sequential patterns of information gathering about treatments and outcomes during decision making.

Initial Decision	Initial Preparation	Initial Post-treatment	Secondary Decision	Secondary Prep.	Secondary Post.
Search queries					
prostate cancer	after prostate	after prostate	after a	adjuvant radiation	seed implants
cancer treatment	surgery for	after prostatectomy	a radical	how much	hdr treatment
proton therapy	robotic prostatectomy	prostate surgery	psa of	taking lupron	pain in
best treatment	after prostatectomy	after surgery	radical prostatectomy	seed implants	cause pain
for prostate	on the	after radical	what are	lupron injections	radiation burns
cancer treatments	da vinci	radical prostatectomy	after radical	i stop	treatment cause
treatment options	what to	psa after	are the	stop taking	lupron treatment
pros and	home on	incontinence after	radiation after	i do	after seed
and cons	the same	how to	radiation therapy	can i	not effective
surgery for	vinci prostate	sex after	cancer treatment	treatment after	psa after
active surveillance	same day	after a	adjuvant radiation	will i	flomax after
da vinci	go home	after robotic	the side	to avoid	protectomy and
surgery vs	davinci prostate	after prostectomy	radiation be	prostate seed	after lupron
watchful waiting	for radical	radical prostectomy	treatment after	with catheter	radical protectomy
vs radiation	to expect	do i	whats next	zytiga cost	enlarged abdomen
treatment for	day of	what to	treatment options	catheter in	for high
cyberknife prostate	life after	on lupron	radiation what	lupron treatment	after medications
cons of	is surgery	levels after	if radiation	radiation after	medications not
prostate treatment	cryotherapy surgery	blood in	be next	taking casodex	long will
the best	kegel exercises	long does	post psa	on lupron	i take
Page content					
early stage	da vinci	prostate surgery	side effects	side effects	of radiation
stage prostate	prostate surgery	after surgery	advanced prostate	the radiation	how long
surgery and	surgery and	the catheter	radiation therapy	medical advice	side effects
cancer treatment	robotic surgery	surgery and	prostate cancer	hormone therapy	therapy is
cancer treatments	the catheter	the penis	early stage	call your	able to
da vinci	robotic prostatectomy	after prostate	hormone therapy	should not	may also
for prostate	surgery the	surgery the	of radiation	that you	cause the
radiation therapy	cancer surgery	surgery is	treatment of	radiation therapy	effects and
the tumor	an erection	the urethra	clinical trials	dose of	the radiation
robotic surgery	able to	an erection	the drug	do not	used to
side effects	erectile dysfunction	able to	treatment with	radiation oncology	weeks after
of urology	after prostate	the surgery	the treatment	used in	dose of
a treatment	of urology	after a	therapy for	your doctor	doctor if
diagnosed with	robotic prostate	most men	has spread	weeks after	talk with
minimally invasive	is removed	da vinci	cancer that	is given	your treatment
prostate surgery	the penis	catheter is	of treatment	list of	after treatment
the latest	minimally invasive	after radical	therapy is	to treat	do not
cancer surgery	the da	is removed	clinical trial	does not	change in
with prostate	most men	i had	of cancer	diagnosis or	have been
robotic prostatectomy	the bladder	incontinence after	psa level	radiation oncologist	the effects
Domain names					
cancer.gov	youtube.com	cancerforums.net	provenge.com	rxlist.com	chemocare.com
cancerfightingstrategies.com	aicr.org	urology.jhu.edu	pcf.org	crmc.org	cancerresearchuk.org
cancercenter.com	mayoclinic.org	healingwell.com	healthyliving.msn.com	uptodate.com	rxlist.com
seattlecca.org	surgery.about.com	medhelp.org	mciverclinic.com	medicinenet.com	radonc.ucla.edu
davinciprostatectomy.com	prostatecancer.org.au	livestrong.com	zytiga.com	cpmc.org	lvhn.org
internationalhifu.com	seer.cancer.gov	hisprostatecancer.com	cancer.net	link.springer.com	philaurology.com
davincisurgery.com	prostate-cancer-institute.org	uptodate.com	search.ask.com	zytiga.com	oncolink.org
provenge.com	ucomparehealthcare.com	healthcentral.com	lifescript.com	nature.com	healthline.com
iuhealth.net	prostatecancercare.com	simonfoundation.org	hisprostatecancer.com	patient.varian.com	ucsfhealth.org
seattlecancerwellness.com	urology.jhu.edu	myprostatedoc.blogspot.com	medscape.org	goodrx.com	hisprostatecancer.com

Table 3: Most probable bigrams associated with each of the six phases of treatment queries. Bigrams are estimated from the relevant search queries and text content and domain names of pages visited following those queries.

Number and Specificity of Treatments.

To understand the progression of treatment-related searches, we examine the average depth of the treatments searched (as defined in the treatment hierarchy displayed in Table 1) and the average cumulative number of treatments searched, as functions of the number of treatment queries conducted. These metrics can provide insights about the decision-making process, including changes over time in the specificity of treatment-related searches and the overall number of different treatments studied by the average user.

Figure 1 shows these values averaged across users for the first 11 initial decision queries, which is the average (10.9) number of queries in this phase (median 8).

We find that the first query is, on average, the broadest, with an average depth below 1, which means the bulk of initial queries contained general terms such as "treatment options". The depth increases nearly monotonically for the first 11 queries, showing that searchers ask for progressively more specific information. Beyond 11 queries, this trend levels off and becomes noisier, as the results are averaged among fewer users.

The average cumulative number of treatments searched in the initial query is 0.65, which means that only 65% of searchers specified a treatment initially, and the remaining 35% conducted a general query such as "best treatment options". After 11 queries, the average user has searched for 2.4 different treatments.

Treatment Comparisons.

We next consider treatments that are referenced together within the same query. Queries with multiple treatments are likely comparative, and we indeed observed a number of queries with explicit comparative language such as "vs", as highlighted in our n-gram characterization of the initial decision phase in the previous subsection.

We found that 9.6% of initial decision treatment queries contained more than one treatment in the same query, and 43.6% of users searched at least one query with multiple treatments. Examining the treatments that co-occurred, we found that 75% of such queries contained surgery and radiation, 7.3% contained different types of surgery, 7.3% con-

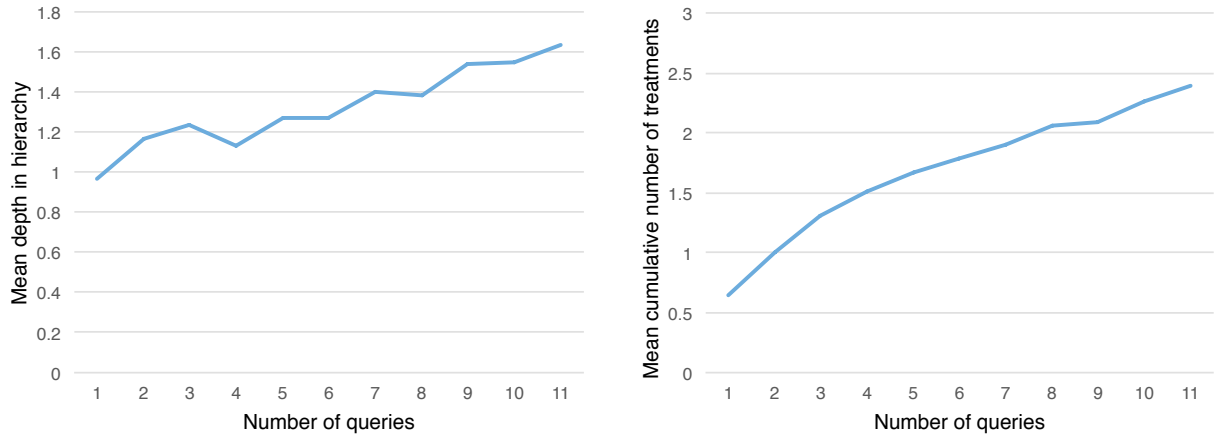


Figure 1: Left: Specificity of treatments searched over time by the average user during the initial decision phase, as given by each treatment query's depth in the treatment hierarchy in Table 1. Right: Cumulative number of different treatments searched over time by the average user. We use the mean number of queries in the decision phase (11 queries) as the range of the x axis.

tained surgery and observation, 6.3% contained radiation and hormone therapy, and 4.2% contained different types of radiation. Of the co-occurrences of surgery and radiation, 65.3% were for the most general terms (e.g., "surgery vs radiation"), while the others contained more specific types (e.g., "robotic surgery or seed implants").

Transitions among Treatments.

Finally, we examined the transition structure among different treatment types by analyzing the sequences of treatments that appear in consecutive queries within the same search session. These experiments were undertaken to pursue insights about how searchers refine their queries as they explore treatment options.

We examined how successive queries progress through the treatment hierarchy in Table 1. We found that 68.8% of the time, the same treatment is searched as in the previous query. In 12.7% of the cases, the subsequent query is more specific than the previous one, going deeper down the same branch in the hierarchy (e.g., searching "robotic surgery" after searching "surgery"), while the subsequent query is coarser (higher up along the branch) in 9.5% of cases. We found that 9.0% of subsequent queries are situated in an entirely different branch of the hierarchy (e.g. searching radiation after searching surgery).

We also explored *which* treatments are likely to be searched after a preceding treatment query. We construct a transition graph, where treatment categories are nodes, and directed edges are weighted by the number of times that one treatment was searched after the other. The graph includes a dummy START node, whose outgoing edges to each treatment node are weighted by the number of times that each treatment was in a user's *initial* query. To show a concise visualization of the typical transitions among treatments, we compute the maximum directed spanning tree of this dense graph, using the Chu-Liu-Edmonds algorithm [7, 12].

The induced tree is shown in Figure 2. The general treatment category (e.g., for non-specific queries such as "treatment options") and the coarsest surgery category follow from START. This may be taken as intuitive as our results above

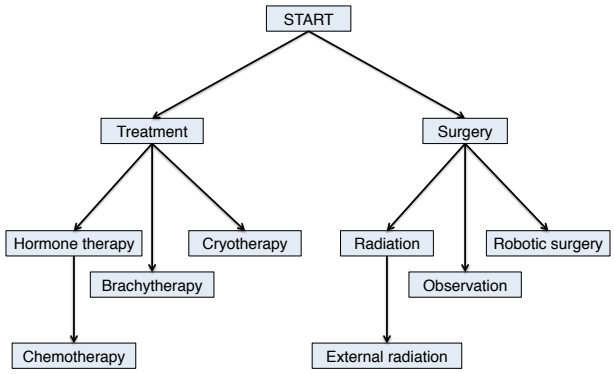
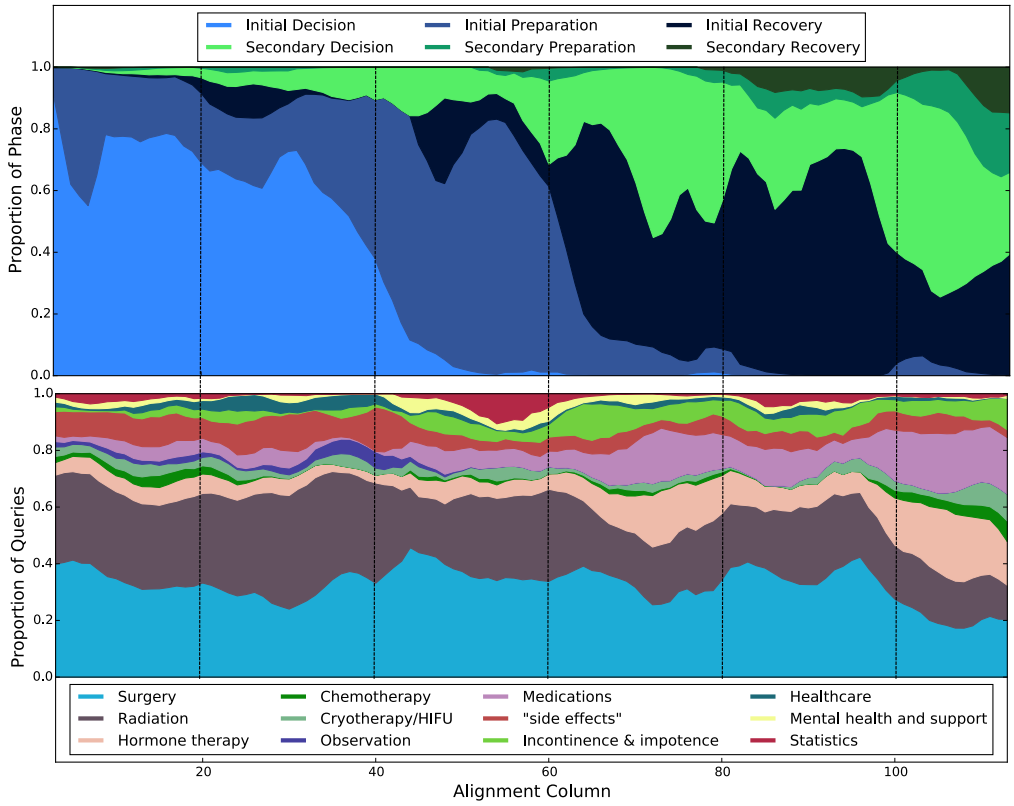


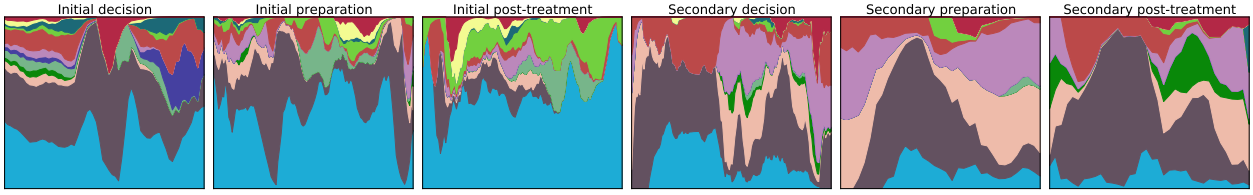
Figure 2: Maximum directed spanning tree induced from the treatment query transition graph.

showed that search histories most often begin with the general treatment category, and searches for surgery are the most common of all specific treatment types. Many of the edges are the same as those captured by the curated treatment hierarchy in Table 1: hormone therapy and cryotherapy following the general treatment category, robotic surgery following surgery, and external beam radiation following radiation. Other edges do not follow the natural hierarchy, but instead illustrate a typical order of treatments that searchers pursue via their queries. For example, searches for radiation and observation are most likely to follow surgery searches, which might be expected in light of the finding above that the most common comparative searches contain surgery and radiation, or surgery and observation.

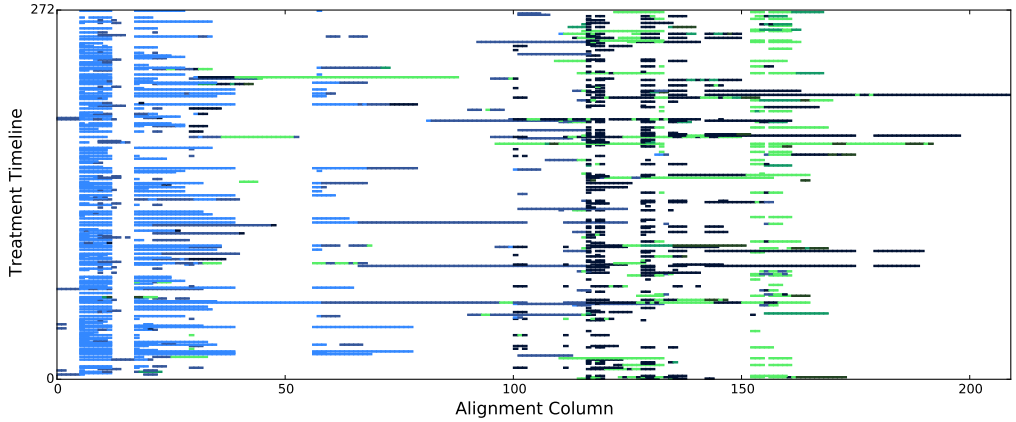
The lowest-weight edges (weight of 3) are excluded to keep the tree concise, and because very low weight edges are likely to introduce noise. (Recall that edge weights are the number of times each treatment category was searched after the other.) One edge was from "Treatment" to "HIFU", which fits with the hierarchy in Table 1. The other edge was from START to "Open surgery", which is not a sensible result, but open surgery happened to be searched in the first query (3 times) more than after other queries.



(a) Distribution over non-gap phases and content categories in each alignment column.



(b) The distribution over content categories in each alignment column, restricted to each particular phase.



(c) Multiple sequence alignment of 272 treatment timelines. Colored dots represent the label in each row/column, using the legend at the top of (a). White space represents gaps.

Figure 3: Different visualizations of treatment timeline alignments.

5.3 Progression of Phases and Search Content

We focused in the previous section on temporal patterns within the initial decision phase. We now seek to understand the temporal patterns across all phases. We wish to visualize how the phases progress over time, and how the content of treatment queries evolves over time within each phase and across entire timelines.

No individual tagged search history contained all six phases described in Section 4.4. However, we can align and stitch together the partial timelines to visualize an "average" complete timeline, aggregated across the hundreds of histories.

Toward this goal, we computed a *multiple sequence alignment* of the timelines. Multiple sequence alignment (MSA) methods were developed in computational biology and are typically used to build a molecular sequence via alignment of smaller sequence snippets. Alignments are scored based on how well symbols at each position align, penalizing gaps and mismatches. The optimization problem is then to solve for a single alignment that gives the highest score.

Solving for the best alignment between two sequences can be done efficiently with dynamic programming, using the same procedure that is used to compute string edit distance. The size of the dynamic programming table increases exponentially with the number of sequences, making this problem NP-hard for an arbitrary number of sequences [18], and impractical for more than a few. Many methods have been developed in computational biology to approximately solve for an MSA efficiently, such as the merging of pairwise alignments. For this experiment, we used ClustalW (from `clustal.org`), a software package for aligning protein sequences [22]. While protein alignments typically use domain-specific scoring functions, we created custom scores appropriate for our task.

Each treatment timeline was considered to be a sequence, and each phase label in the timeline was treated to be a symbol in the sequence. The most likely label was chosen in cases where annotators listed more than one possibility, using the annotators' highest-ranked choice. We created special symbols for the first query with each phase label in the timeline, to encourage the start of each phase to align, so that they are not aligned to arbitrary positions. This means that there are 12 total symbols and 272 sequences.

Alignments are scored such that each position in the alignment is given a score of 1 if the symbols match. A score of 0.1 is given if the phases match, but one of the symbols is the special 'first time' indicator and the other is not. No credit is given for aligning different phases. We did not apply strong penalties for gaps, preferring alignments where different phases do not overlap. We used a penalty of 0.1 for gap creation, to discourage gaps with all other options being equal, with no penalty for introducing successive gaps.

Figure 3(c) shows the resulting MSA of the timelines. We see that the left is dominated by initial phases (blue) while the right is dominated by secondary phases (green), though the phases do not progress monotonically. We note that the initial post-treatment and secondary decision phases are often interleaved, as searchers tend to search for recovery or side effects following the initial treatment at the same time as they search for the next steps.

To more clearly see the phase progression, Figure 3(a) shows the distribution of the phases after gaps are removed for each column. To reduce noise, we excluded columns with less than ten non-gap symbols. This roughly halves the number of columns included in the visualization. The values are smoothed by averaging the values from the preceding/following three columns.

Age	Sample	Filtered	Classified	Expected
20s	16.4%	7.3%	4.9%	0.0%
30s	17.0%	5.2%	2.8%	0.0%
40s	13.5%	9.0%	5.6%	1.4%
50s	18.8%	14.6%	12.7%	15.3%
60s	17.8%	39.1%	42.3%	43.1%
70s	8.1%	14.9%	23.9%	24.1%
80s	8.4%	9.8%	7.7%	16.1%

Table 4: The distribution of ages associated with searches from a 2-month sample of logs (left), the filtered set of searchers who searched "prostate cancer", and the set of positively classified users. The right column shows the expected distribution in the logs of real cancer diagnoses, based on ground truth incidence rates.

In addition to the distribution of phases, the figure also shows the distribution over various categories of search terms in each column. The categories are based on the term ontology (Section 4.1), including terms referencing treatments and side effects, as well as searches for healthcare providers, search terms referencing mental health or seeking social or emotional support, and searches that are aimed at retrieving statistics such as prognosis or success rates. We see some variation over time in the content. Searches for hormone therapy and prostate cancer medications (many of which are hormonal therapies, but categorized separately for this visualization) increase over time, while searches for observation only appear in the first half. The general term "side effects" is prominent initially and declines, and more specific terms for side effects (related to incontinence and impotence) begin to rise. This highlights a shift from a general interest in learning about side effects to more specific concerns.

Many of the differences in the category distributions over time are smoothed over due to overlapping phases at each point. Figure 3(b) shows the content distribution within each phase in isolation, clarifying the differences. For these images, we did not require a minimum number of non-gap values, other than excluding columns with only gaps. This reveals differences among the phases. For example, searches for healthcare (dark green) appear mostly in the initial decision phase, while searches describing mental health (yellow) appear mostly in the initial post-treatment phase. The treatment distribution differs between the initial and secondary phases, with fewer references to surgery and more references to hormone and chemotherapy in the latter.

6. AGE COMPOSITION OF SEARCHERS

We now present a final experiment focused on examining ages inferred for the experiential searchers in our dataset. We performed this experiment to understand the demographic composition of the dataset, as well as to provide an auxiliary form of validation, to determine whether the searchers in our dataset exhibit similar demographics as patients diagnosed with prostate cancer. Since rates of cancer are higher in older age groups than in the general population, we would expect to see a shift toward this demographic in our classified dataset, should our classifier be capturing a higher proportion of people experiencing cancer.

We associated age groups with searchers by looking for references to ages in search queries. Specifically, we matched queries against expressions of the form "at/age __", "__ year(s) old", and "in my/his/her __s" for different numeric values. A similar self-reporting methodology was used to estimate the stage in pregnancy of expectant mothers or the age of newborns, based only on search logs [13]. If a search history included multiple such expressions, the majority age group was chosen. We were able to associate ages with 142 user identifiers (out of the set of 1,413 classified) after applying this method.

We computed the distribution of age groups for the search histories identified by the classifier. For comparison, we computed the age distribution for the larger set of histories from our initial filter (those who searched "prostate cancer" at least three times), and the entire set of search logs from the most recent two months. From the two month sample, we estimated the expected distribution of ages in the search logs among those diagnosed with prostate cancer in the United States (US). We computed this estimate using Bayes' theorem: $\propto P(cancer|age)P(age)$, where $P(cancer|age)$ is defined by the age-specific prostate cancer incidence rates from the US National Cancer Institute[2] (NCI), and $P(age)$ is the distribution in the overall sample of logs (the first column of Table 4).

Table 4 shows the distribution of age groups from 20s to 80s (the NCI data does not list rates for specific age groups beyond 80s) for the three sets of log data as well as the expected distribution from incidence rates.

The age distribution among positively classified searchers is strikingly similar to the expected distribution, particularly for the ages of 60s and 70s, which are each within 1 percentage of the expected rate. The Pearson correlation between these two distributions is highly significant ($r = .959$, $p < .001$). The distribution among users passing through our initial "prostate cancer" filter (three queries for prostate cancer) skews toward the older distributions as might be expected. However, after applying the experiential classifier, the percentages further increase for the age groups with the highest incidence rates (people in their 60s and 70s) and decrease for the youngest age groups. We believe this provides additional evidence that many of the classified searchers included in our study were likely to be experiencing a diagnosis of prostate cancer.

7. DISCUSSION AND IMPLICATIONS

A limitation of using search log analysis to learn about Web-based decision support is that we lack the larger context that frames the information seeking activity. We do not have a means for knowing with certainty about the motivations behind searches. To address this shortcoming, we are developing methodologies to connect long-term log behaviors with self-reported data from consenting search participants. This approach provides a means of engaging directly with patients, who could provide details on the context behind the search and retrieval activity appearing in logs of online activity. This information would allow us to understand the influence of Web content on a patient's decision making and the details of the resources that were most helpful in deliberating about care decisions, whether the decision aligned with a doctor's recommendation, and the outcome following the decision. Many of these questions are difficult or impossible to answer from the observed activities in logs alone. Engaging directly with patients, and gaining their consent to align their time-stamped online activity with their evolving clinical situation and associated information needs, would enable us to perform a rich analysis grounded in detailed patient reports. While we hope to conduct such a study as future work, we also believe that there is complementary value in the use of the lighter-weight, larger-scale analyses presented in this paper.

We believe that there are opportunities to leverage the reported findings to inform the design of search and retrieval systems for supporting healthcare decision making. Given the focus of this paper, we are especially interested in enhancing the ability of Web search engines to serve as decision support systems, per the expectations that people appear to have when they turn to the Web for critical assistance with challenging treatment decisions under uncertainty. The different phases of treatment appear to be identifiable, with distinctive n-gram associations and timing characteristics. This suggests that search output can be tailored to the user's current phase. For example, a searcher during a decision-making phase may find it helpful to gain access to results that include comparisons of treatments. We found that comparative queries are common, with nearly half of users conducting an explicit comparison, and from our experiments, we know which comparisons are most common. These insights could be used to return results, or interface treatments such as direct answers using data pulled from external sources, that include relevant comparisons even if the query did not explicitly include a comparison. We also now have data on the progression of queries on treatments, including the depth over time and transitions between treatments, which can be used to model treatment search behavior. Such models could be used by search providers for important tasks such as appropriately suggesting or expanding queries.

Beyond tailoring results to a current phase of search aligned with a phase of care, it may be valuable to provide searchers with content that is typically viewed in later phases of an illness. For example, searchers making a decision may find it useful to read the content that is commonly viewed by those in a post-treatment phase, in order to understand the expected recovery process and side effects from particular treatments. Our analyses showed that many people may seek out discussions of the personal experiences of patients (e.g. through forums) during the post-treatment phase, and surfacing the concerns and issues of others could provide searchers facing decisions with more context than traditional decision-making materials. Such content might not be discovered in the course of normal searching in an early phase without designing a system for such proactive retrieval, as queries in the initial decision phase tend to be broad.

Promising future directions include adapting the annotations, classifiers, and overall methodology to understand the information needs and to guide decision support for treatment decisions for other illnesses. Beyond the pursuit of enhancing search for decisions about treatments, we can employ the methods to enhance search and retrieval for other healthcare needs, such as selecting a care provider.

[2] `http://seer.cancer.gov/csr/1975_2011/browse_csr.php?sectionSEL=23&pageSEL=sect_23_table.07.html`

8. REFERENCES

[1] S. L. Ayers and J. J. Kronenfeld. Chronic illness and health-seeking information on the internet. *Health*, 11(3), 2007.

[2] J. Bader and M. Theofanos. Searching for cancer information on the internet: Analyzing natural language search queries. *J. Med. Internet. Res.*, 5(4):e31, 2003.

[3] E. V. Bernstam, J. R. Herskovic, and W. R. Hersh. Query log analysis in biomedicine. 2009.

[4] P. Black and D. Penson. Prostate cancer on the internet – information or misinformation? *Journal of Urology*, 175(5):1836–1842, 2006.

[5] M.-A. Cartright, R. W. White, and E. Horvitz. Intentions and attention in exploratory health search. In *SIGIR*, 2011.

[6] K. Castleton, T. Fong, A. Wang-Gillam, M. Waqar, D. Jeffe, L. Kehlenbrink, F. Gao, and R. Govindan. A survey of internet utilization among patients with cancer. *Support Care Cancer*, 19(8), 2011.

[7] Y. Chu and T. Liu. On the shortest arborescence of a directed graph. *Science Sinica*, 14:1396–1400, 1965.

[8] J. Chuang, C. D. Manning, and J. Heer. Termite: Visualization techniques for assessing textual topic models. In *Advanced Visual Interfaces*, 2012.

[9] R. J. W. Cline and K. M. Haynes. Consumer health information seeking on the internet: the state of the art. *Health Education Research*, 16(6), 2001.

[10] D. Downey, S. Dumais, and E. Horvitz. Models of searching and browsing: languages, studies, and applications. In *IJCAI*, 2007.

[11] G. E. Dupret and B. Piwowarski. A user browsing model to predict search engine click data from past observations. In *SIGIR*, 2008.

[12] J. Edmonds. Optimum branchings. *J. Res. Nat. Bur. Standards*, 71(B):233–240, 1967.

[13] A. Fourney, R. W. White, and E. Horvitz. Exploring time-dependent concerns about pregnancy and childbirth from search logs. In *CHI*, 2015.

[14] S. Fox and M. Duggan. Health online 2013. Technical report, Pew Internet and American Life Project, 2013.

[15] C. M. Gaston and G. Mitchell. Information giving and decision-making in patients with advanced cancer: A systematic review. *Soc Sci Med*, 61(10), 2005.

[16] J. Ginsberg, M. Mohebbi, R. Patel, L. Brammer, M. Smolinski, and L. Brilliant. Detecting influenza epidemics using search engine query data. *Nature*, 457(7232), 2008.

[17] Y. Goto and T. Nagase. Oncology information on the Internet. *Japanese Journal of Clinical Oncology*, 42(5):368–374, 2012.

[18] D. Gusfield. *Algorithms on Strings, Trees, and Sequences: Computer Science and Computational Biology*. Cambridge University Press, 1997.

[19] P. R. Helft. Patients with cancer, internet information, and the clinical encounter: A taxonomy of patient users. In *American Society of Clinical Ontology*, 2012.

[20] R. Islamaj Dogan, G. C. Murray, A. Névéol, and Z. Lu. Understanding PubMed user search behavior through log analysis. *Database*, 2009:bap018, 2009.

[21] L. Klotz. Active surveillance for low-risk prostate cancer. *F1000 Med Reports*, 4(16), 2012.

[22] M. Larkin, G. Blackshields, N. Brown, C. R., P. McGettigan, and et al. Clustal W and Clustal X version 2.0. *Bioinformatics*, 23:2947–2948, 2007.

[23] T. Lau and E. Horvitz. Patterns of search: analyzing and modeling web query refinement. In *7th international conference on user modeling*, 1999.

[24] G. Lin, D. Aaronson, S. Knight, P. Carroll, and R. Dudley. Patient decision aids for prostate cancer treatment: A systematic review of the literature. *CA Cancer J Clin*, 59:379–390, 2009.

[25] A. M. O'Connor, V. Fiset, C. DeGrasse, I. D. Graham, W. Evans, D. Stacey, and et al. Decision aids for patients considering options affecting cancer outcomes: evidence of efficacy and policy implications. *JNCI Monographs*, 1999(25):67–80, 1999.

[26] A. M. O'Connor, A. Rostom, V. Fiset, J. Tetroe, V. Entwistle, H. Llewellyn-Thomas, and et al. Decision aids for patients facing health treatment or screening decisions: systematic review. *BMJ*, 319(7212):731–734, 1999.

[27] Y. Ofran, O. Paltiel, D. Pelleg, J. M. Rowe, and E. Yom-Tov. Patterns of information-seeking for cancer on the internet: An analysis of real world data. *PLOS One*, 7(9), 2012.

[28] R. Overberg. *Breast cancer stories on the internet: improving search facilities to help patients find stories of similar others*. PhD thesis, Leiden University, 2013.

[29] H. Patel, S. Mirsadraee, and M. Emberton. The patient's dilemma: Prostate cancer treatment choices. *Journal of Urology*, 169(3):828–833, 2003.

[30] M. Paul and R. Girju. Cross-cultural analysis of blogs and forums with mixed-collection topic models. In *EMNLP*, 2009.

[31] M. J. Paul, R. W. White, and E. Horvitz. Search and breast cancer: On disruptive shifts of attention over life histories of an illness. Technical Report MSR-TR-2014-144, November 2014.

[32] S. Pautler, J. Tan, G. Dugas, N. Pus, M. Ferri, W. Hardie, and J. Chin. Use of the Internet for self-education by patients with prostate cancer. *Urology*, 57(2):230–233, 2000.

[33] G. Peterson, P. Aslani, and K. A. Williams. How do consumers search for and appraise information on medicines on the internet? a qualitative study using focus groups. *J Med Internet Res*, 5(4), 2003.

[34] M. Richardson. Learning about the world from long-term query logs. *ACM Transactions on the Web*, 2(4), 2009.

[35] L. J. F. Rutten, N. K. Arora, A. D. Bakos, N. Aziz, and J. Rowland. Information needs and sources of information among cancer patients: a systematic review of research (1980–2003). *Patient Education and Counseling*, 57(3), 2005.

[36] M. J. Satterlund, K. D. McCaul, and A. K. Sandgren. Information gathering over time by breast cancer patients. *J Med Internet Res*, 5(3), 2003.

[37] A. Sidana, D. J. Hernandez, Z. Feng, A. W. Partin, B. J. Trock, S. Saha, and J. I. Epstein. Treatment decision-making for localized prostate cancer: What

younger men choose and why. *The Prostate*, 72(1):58–64, 2012.

[38] M. I. Trotter and D. W. Morgan. Patients' use of the internet for health related matters: a study of internet usage in 2000 and 2006. *Health Informatics*, 14(3), 2008.

[39] R. J. Volk, S. T. Hawley, S. Kneuper, E. W. Holden, L. A. Stroud, C. P. Cooper, and et al. Trials of decision aids for prostate cancer screening: a systematic review. *American Journal of Preventative Medicine*, 33(5):428–434, 2007.

[40] T. Weninger, W. H. Hsu, and J. Han. Cetr: Content extraction via tag ratios. In *WWW*, 2010.

[41] R. W. White and S. M. Drucker. Investigating behavioral variability in web search. In *WWW*, 2007.

[42] R. W. White and E. Horvitz. Cyberchondria: Studies of the escalation of medical concerns in web search. *ACM TOIS*, 27(4), 2009.

[43] R. W. White and E. Horvitz. From health search to healthcare: explorations of intention and utilization via query logs and user surveys. *J Am Med Inform Assoc*, epub, 2013.

[44] R. W. White, N. P. Tatonetti, N. H. Shah, R. B. Altman, and E. Horvitz. Web-scale pharmacovigilance: Listening to signals from the crowd. *J Am Med Informatics Assoc*, 20(3), 2013.

[45] C. Zhai and J. Lafferty. Two-stage language models for information retrieval. In *SIGIR*, pages 49–56, 2002.

[46] S. Ziebland, A. Chapple, C. Dumelow, J. Evans, S. Prinjha, and L. Rozmovits. How the internet affects patients' experience of cancer: a qualitative study. *BMJ*, 328(7439), 2004.