

- [Contents](#) |
 - [Author index](#) |
 - [Subject index](#) |
 - [Search](#) |
 - [Home](#)
-

Context-based interactive health information searching

[Tesfahun Melese Yilma, Anushia Inthiran, Daniel D Reidpath and Sylvester Olubolu Orimaye](#)

Introduction. This paper deals with the impact of contextual features, such as sex, age, mother tongue, health status, health literacy, Internet use experience, and frequency of health information seeking on health information searching.

Method. An interactive information retrieval approach was used to study users' searching behaviour. An online survey and experiment using simulated situation technique were used as data collection methods. The online survey gathered data about user features, such as sex, age, mother tongue, health status, health literacy, Internet use experience, and health information seeking. An experiment was then carried out using four simulated tasks to collect information about health information searching.

Analysis. The multiple linear regression analysis method was used to identify contextual factors affecting query length and number of queries. In addition, binary logistic regression analysis method was used to identify contextual factors affecting result clicking.

Results. Frequent health information seeking leads to more queries and long query length, and English as a mother tongue and being healthy contribute to long query length. Queries with spelling errors and those formulated outside task descriptions are found to be ineffective.

Conclusion. Contextual features such as frequency of health information seeking, mother tongue, and health status influence query formulation. In addition, spelling errors and source of query affect the effectiveness of queries. The findings could be useful for health information retrieval systems to learn and predict users' information needs to aid effective retrieval.

Introduction

In the past, researchers have focused on the evaluation of information retrieval systems to improve precision and recall of search results. Nowadays, however, researchers are paying attention to the role of humans to design and improve information retrieval systems ([Kelly, 2009](#)). This is because users are central in information retrieval as they are the ultimate consumers of information retrieval systems. Understanding users and their context could help information retrieval systems to deliver the right information to the users in the right way.

The focus of our research is on the health domain due to the popularity of health information searching on the Web ([European Commission, 2014](#); [Fox and Duggan, 2013](#); [Medlock, et al., 2015](#)). Besides its popularity, health information searching is crucial to increase the ability to comfortably discuss health issues with health professionals in order to have enhanced shared decision making ([Feinberg, Greenberg and Frijters, 2015](#); [Jung, 2014](#); [Lambert and Loiselle, 2007](#)), to improve the ability to cope with stresses, to increase self-care management skills and commitment to treatments ([Jung, 2014](#); [Lambert and Loiselle, 2007](#)), and to increase medical treatment satisfaction ([Jung, 2014](#)). As a result of its importance, health information searching has become an essential research area that calls for a systematic investigation of searching behaviour.

Existing research studies focus on demographic characteristics, health information search experience, medical topic familiarity, task clarity, task easiness ([Inthiran, Alhashmi and Ahmed, 2011a](#)), and relevant assessment of search results to contextualize users' information need and searching behaviour ([Lopes and Ribeiro, 2010](#)). However, the context of information searching is often ignored ([Bierig and Göker, 2006](#); [Ingwersen and Jarvelin, 2005](#)) while it can have the potential to improve information retrieval process. Context remains an indefinable concept in information retrieval ([Jones, 1981](#)) but is commonly referred to. Kelly ([2006](#)) states that context in information retrieval comprises individual and situational variables. The individual variables consist of personal characteristics, knowledge structures, and cognitive and learning styles, whereas that of situational variables include search task characteristics.

Context could help information retrieval systems to learn and predict information needs of a user, relate one piece of information to another, display required information in a suitable manner to users, relate tasks performed by a searcher and enable other searchers to use it. Therefore, in this study, the impact of contextual features on health information searching is investigated. Specifically, the influence of user features (such as sex, age, mother tongue, health status, health literacy, Internet use experience, and frequency of health information seeking) on the formulation of queries are examined. These features are selected because they have the potential to influence query formulation. For example, health literacy ([Ellis, Mullan, Worsley and Pai, 2012](#); [Gutierrez, Kindratt, Pagels, Foster and Gimpel 2014](#)), health status ([Kim, 2015](#); [Wang, Viswanath, Lam, Wang and Chan, 2013](#)), and Internet use experience ([Escoffery et al., 2005](#); [Janeice, Eileen and Trauth, 2013](#)) are found to influence health information seeking.

In addition, the effects of query features (such as the type of query, the source of the query, and spelling error) on the result accessing behaviour are analysed. The type of query could be word-based, question-based, or based on word-strings. The classification could help to point out users' level of knowledge on query language and searching skills. Users are likely to use queries based on word strings with Boolean operators (complex query) if their information retrieval knowledge is high, with the reverse being true, i.e., using simple word-based queries, if their information retrieval knowledge is low ([Sutcliffe and Ennis, 1998](#)). A question-based query could also help to suggest an automated question-answering technique in the design of an information retrieval system ([Ramprasath and Hariharan, 2014](#)). Understanding the effects of user contextual features on query formulation and query features on result clicking could help to design effective health information retrieval systems that can allow users to retrieve needed information effectively and efficiently.

Related work

In this section, we provide information about query behaviour, contextual factors of query formulation and result clicking from previous research. The information could help us to understand the research coverage and gaps in research on contextual health information searching behaviour.

Query behaviour

Users are central in information retrieval and they need to correctly phrase their queries to successfully achieve their goals ([Yeganova, Comeau, Kim and Wilbur, 2009](#)). However, laypeople sometimes use plain English sentences ([Inthiran, Alhashmi and Ahmed, 2011b](#); [Luo, Tang, Yang and Wei, 2008](#)) or questions ([Zhang, 2013](#)) as a query to search for health information. In addition, users from non-health domains may not be able to identify medical or health terms for their query, or their query may not match with medical vocabularies. For example, Yi ([2015](#)) identifies difficult health terminology as one of the barriers to finding health information. These make existing Web search engines to less likely to give targeted query suggestions. When plain English sentences are used as queries, the hits returned from the system are unlikely to be accurate or even relevant. Consequently, users are commonly disappointed by the search results and unlikely to click and view the results, resulting in unsatisfactory searching outcome ([Inthiran et al., 2011b](#)). In addition, the use of abbreviations or acronyms, use of slang expressions, or misspellings during query formulation are common barriers to retrieving health information ([Boden, 2014](#)). Hence, this paper attempts to understand how laypeople formulate health queries.

Users are commonly observed to use simple and short queries which might lead them to unsatisfactory results. A research study on searching behaviour of students when looking for specific health-related Information in MedlinePlus finds that the average query length issued by users ranged from 1.79 to 4 terms per task with an average number of queries ranging from 1 to 2.63 per task ([Zhang, 2013](#)). However, most of the terms that form the queries are found to be stop words which do not help to characterize users' information needs. From 10,257 queries extracted from [www.alltheWeb.com](#), Spink *et al.* ([2004](#)) find 2.3 average terms per medical or health query with a mean of 2.2 queries per medical or health session, which is short. This study also finds that users show less query reformulation efforts on their medical or health queries which show succinct query issuance. In the point of views of these research studies, the use of too-short queries and stop words to formulate queries could be obstacles to searching for the required information. This is because, when a query is too short, searching becomes challenging due to a lack of information regarding user intent.

One study indicates that longer queries are significantly associated with increased users' satisfaction with search results ([Belkin et al., 2003](#)). As users issue longer queries that characterize their information need, their searching effectiveness increases, resulting in a better interactive information retrieval performance. However, Inthiran, Alhashmi and Ahmed ([2015](#)) observed medical students while issuing long queries when searching on a difficult task. The students were found to be slow to complete search tasks and experienced ineffective search sessions. These two findings conflict with each other in terms of the effectiveness of using long queries even though medical students are perceived to have adequate knowledge of medical terms which should enable them to formulate queries better. These contradictory findings are worth further research.

Contextual factors in query formulation

Information retrieval can be affected by contextual factors, such as searchers and search task characteristics ([Kelly, 2006](#)). Query formulation which is an essential part of the information retrieval system can also be affected by context.

Existing research studies that explore contextual factors of query formulation are limited. These few studies focus mostly on task features and often ignore user features that possibly affect information searching on the Web. Lopes and Ribeiro ([2010](#)) studied the effects of task and user features on query formulation in health information searching. The authors found that sex, years of experience in Web search, frequency of health information

seeking, previous experience on search topics and ease of tasks are associated with query length. That is, being female, less experienced in Web searching, frequently seeking health information, having previous search experience on a topic, and difficult search tasks, are associated with longer query length. In the study, the effect of user features on a number of queries is ignored which is a key feature in the query formulation process. Kelly (2009) points out that including users in information retrieval system evaluation and investigating users' information searching behaviour are essential for improving information retrieval systems. Hence, in our study, we explore the effect of user features on the query length as well as the number of queries.

Another research study has attempted to identify factors affecting query formulation in Web information searching (Aula, 2003). This study identifies experience in using computers and the Web as factors affecting the query formulation process, indicating that experienced users issue longer and more queries than users with less experience in using computers and in searching the Web. The study by Aula was conducted in the general information domain. In our study, we explore the contextual factors in the health-specific domain. The health domain is rich in context. In addition, its search process is based on well-defined scenarios (e.g., prevention, treatment, or medication) whose contexts could help to improve health information retrieval systems.

Result clicking

A challenge for users to search for health information is to filter and select relevant results. A study (Broussard and Zhang, 2013) finds that users were uncertain which Websites are relevant after search results were returned by the search engine. Users either selected results through trial and error or selected results from the top hits, believing that the more relevant results are located at the top of a result page. They also selected results based on Website familiarity simply because they feel comfortable with the familiar Website. Another study found that users selected results based on Website reputability and advertisement (Fiksdal *et al.*, 2014). Users were likely to select a Website if it was reputable and did not have a lot of advertisements and popups. Users considered a Website as reputable if its organization name was well known, if it gained credibility through academics and research and if it was a local Website. In our study, we explore how users select search results after a search engine returned the results.

Users may be frustrated while searching for health information on the Internet. The frustration could be during searching or at the result viewing stage. For example, Janeice *et al.*, (2013) find that participants are frustrated because of their limited searching skills. Users with better searching skills are more likely to specify their queries and commit fewer errors in getting a satisfactory search result (Hu, Lu and Joo, 2013). In contrast, when users are inexperienced with the Internet, the search process will take more time and at the end they may be frustrated. The above research studies by Janeice *et al.*, (2013) and Hu *et al.*, (2013) show the impact of users' searching skills on search satisfaction. However, in our study, we will investigate the effect of query features, such as spelling errors, the source of the query and query type on result clicking.

Conceptual framework

The process of information searching consists of problem identification, need articulation, query formulation and results evaluation (Sutcliffe and Ennis, 1998) and these are mainly affected by factors such as environment (e.g., search engine and search topic), searcher, search process (e.g., command used) and search outcome (Fenichel, 1981). Since the current study is focused on investigating the contextual factors affecting query formulation during health information searching on the Web, it is further explored using earlier studies. Based on Lopes and Ribeiro (2010) who have studied the context effect of query formulation in health information searching, two main factors are expected to have a significant impact on query formulation. These are user and task features. The user features included in study by Lopes and Ribeiro are age, sex, language, Web search

experience, health search experience and familiarity with the search topic. The task features included are a medical specialty, the task clarity and task easiness.

In the current study, the following contextual features are included: demographic characteristics (sex, age, mother tongue and Internet use experience) and health-related variables (health literacy, health status, frequency of health information seeking, and task search experience). Health literacy and health status are included because of their influence on health information seeking. Adequate health literacy level ([Ellis et al., 2012](#); [Eriksson-Backa, Ek, Niemelä and Huotari, 2012](#); [Gutierrez et al., 2014](#)) and poor health status ([Kim, 2015](#); [Wang et al., 2013](#); [Weaver et al., 2010](#)) are associated with frequent health information seeking which could increase knowledge and skills required to articulate search queries. Considering the user and task features, the conceptual framework shown in Figure 1 is developed.

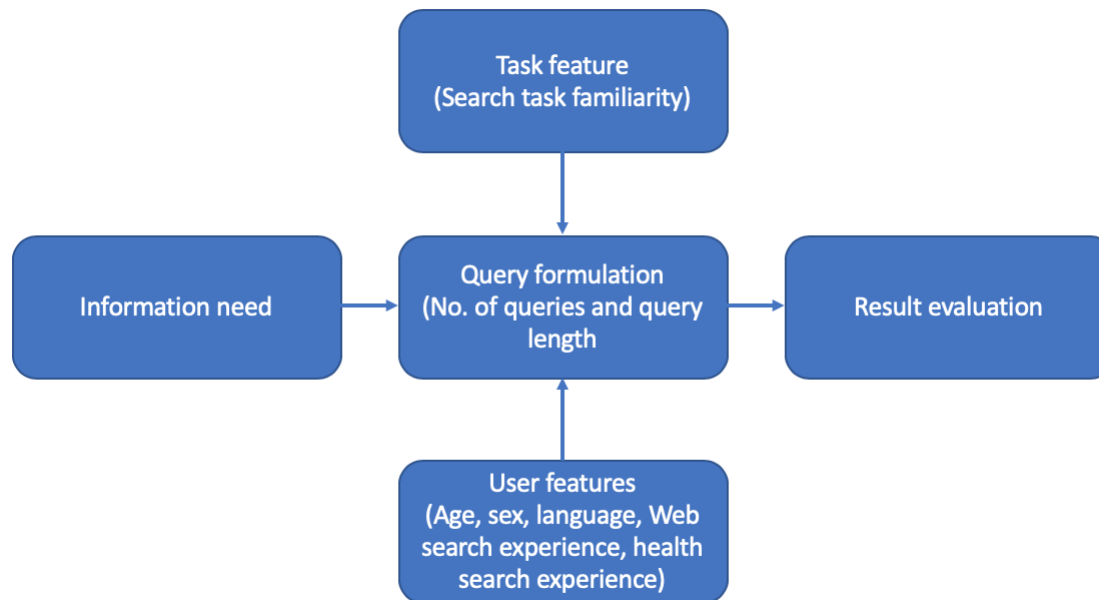


Figure 1: Conceptual framework of contextual factors affecting query formulation

Objective

The objective of this research study is to assess the search behaviour of consumers and identify contextual factors affecting health information searching on the Web. The research questions associated with this research objective are:

- 1. What are the contextual factors associated with query formulation during health information searching on the Web?
- 2. What behaviour do consumers exhibit when they search for health information on the Web?

Method

Participants

Undergraduate students from a university in Malaysia are included as study participants. However, students from health-related discipline are excluded as they are more likely to be well-versed in health knowledge, which may potentially affect the result of the study. A total of fifty-eight students participated in the study. The students were invited to participate through personal communication at the end of their class.

Search tasks

Simulated situations were employed to study the search behaviour of participants. A simulated situation consists of a short story that elaborates a situation that motivates an individual to search for information on the Web ([Borlund, 2000](#)). In our study, participants were provided with four simulated tasks to perform in any Website they chose. Four search tasks were used because real search behaviour might not be obtained with only one or two tasks. Kelly ([2009](#)) suggests that simulated tasks should not last over an hour to reduce participants' fatigue which could affect the result of a study. In our study, the average time it took to complete the four tasks was forty-five minutes, with a range of thirty to sixty minutes. Previous research studies used three ([Zhang, Wang, Heaton and Winkler, 2012](#)) to four ([Puspitasari, Moriyama, Fukui and Numao, 2015](#)) simulated tasks to study users' health information searching behaviour. To reflect real information needs, the tasks were developed based on the health information needs identified from a pilot study.

The four tasks selected related to dengue fever, smoking, obesity and physical activity. These four health issues are a burden to young populations in developing countries like Malaysia. For example, dengue fever is a major health problem causing an economic burden in developing countries, particularly in Southeast Asian countries ([Guzman and Istúriz, 2010](#)). The disease is among the most serious health problem in Malaysia affecting social function, well-being and quality of life, predominantly amongst children and young adults ([Shepard *et al.*, 2013](#)). The prevalence of the disease is higher among people aged 20-29 years in Malaysia ([Chew, Rahman and Salleh, 2012](#)). As of October 2015, over 98,500 dengue cases with 267 deaths had been reported in Malaysia ([World Health Organization, 2015b](#)).

Smoking is also the major risk factor for non-communicable diseases which contributed to 80% of premature deaths in developing countries ([World Health Organization, 2015c](#)). In Malaysia, non-communicable diseases are the leading cause of premature deaths ([Malaysia. Ministry of Health, 2011](#)) contributing to 73% of total deaths ([World Health Organization, 2014](#)). According to the 2012 Global Adult Tobacco Survey Report, over 10,000 Malaysian die as a result of smoking-related illnesses per annum. The daily use of tobacco among people aged 15-24 years is 32% in Malaysia ([Malaysia, Institute... 2012](#)). Obesity is another major risk factor contributing to cardiovascular diseases which are responsible for three-quarters of deaths in developing countries ([World Health Organization, 2015a](#)).

Therefore, using these four health problems, four simulated situations containing a short cover story that can motivate the study participant to search an information retrieval system were developed. Each simulated situation contained goals for participants to perform a search. For example, the simulated situation for physical activity was described as follows:

Simulated situation: Imagine last night you listened to a radio programme about physical inactivity. You have heard that physical inactivity is a key risk factor for cancer, diabetes, heart disease, stroke and chronic respiratory diseases. The radio programme also pointed out that more than 80% of the world's adolescent population is insufficiently physically active. In comparison, you also heard that

sufficient physical activity gives benefits of reduced stress, improved sleep and leads to a better quality of life. You have noted the benefit of physical activity, and, at the same time, you are concerned that you are physically inactive. You want to find out proper guidelines or recommendations on physical activity.

Indicative request: Find information about guidelines or recommendations to help you to be sufficiently physically active.

Data collection tools

Two data collection methods are used: an online survey and an experimental simulation using simulated tasks. Through the online survey, we gathered data about user features, such as sex, age, mother tongue, health status, health literacy, Internet use experience and health information seeking. Health literacy is measured using a tool from the European Health Literacy Survey ([Sorensen et al., 2013](#)). The Asian Health Literacy Association adopted the tool to be used by researchers in Asian countries ([Duong et al., 2017](#)). Furthermore, a study was undertaken to assess the validity and internal reliability of the tool in the Malaysian context ([Mohamad, Su, Majid and Chinna, 2014](#)). This study found high reliability and concluded that the tool can be utilized in the Malaysian context. Therefore, the forty-seven item European Health Literacy Survey questionnaire was used to determine the health literacy of the participants under study. The forty-seven items are assessed using a four-point self-report scale (very easy, easy, difficult and very difficult). Following the survey, the experimental procedure was carried out to collect information about health information searching behaviour of the participants using the [Morae keylogger software](#), which is useful to record users' interactions with information retrieval systems. The Morae software captures audio, video, on-screen activity and keyboard and mouse inputs during a search session. However, in this study, only on-screen activity and keyboard and mouse inputs are captured.

A file name with a similar unique identification number to that given in the survey was used to save the screen recording. This enabled us to link the data from the screen recording with the data from the survey. An observation checklist was used to extract data from computer screen recordings using the Morae software.

Experimental procedure

Students were approached and invited to participate in the study through personal communication made at the end of their class. First, a short introduction to the study and its purpose was given. After they agreed to participate in the study, an appointment was made. Upon arrival, the participants were provided with the ethics approval letter and an explanatory statement of the study. Then, we requested them to review and sign the informed consent form. After consent, they were requested to complete the survey questions. Following the survey, four tasks in random order (to balance task order effect) were given to the participants. We told the participants to perform each task in any Website or information retrieval system they prefer. They were also told to take as much time as they want to complete the tasks.

Data analysis

The data from the online survey were downloaded from Google Forms as an Excel spreadsheet and then transferred to SPSS v.20 for coding and analysis. We transcribed key logging data from the experiment for analysis. Transcribed data were entered using a pre-designed observation checklist in Google Forms. The observational checklist was designed based on the search activities (querying and result clicking) of the participants.

Descriptive statistics were used to describe demographic characteristics and contextual features of the respondents. In addition, the multiple linear regression method was used to identify contextual factors affecting query length and the number of queries issued by the participants. Binary logistic regression was also used to identify contextual factors affecting the result clicking of the participants. Multiple linear regression analysis was used because the dependent variables (query length and number of queries) are continuous. Logistic regression was also an appropriate analysis method as the dependent variable is dichotomous or binary ([Harrell, 2015](#)). In this research study, the dependent variable (result clicking) had two responses (clicked/not clicked). We selected the stepwise logistic regression method as it controls the confounding effect, which occurs when a variable correlates with both dependent and independent variables thereby affecting the statistical outcome ([Sterne and Kirkwood, 2003](#)). A p-value of 0.05 was used as a cutoff value for statistically significant findings.

Results

Demographic information

A total of fifty-eight computer screen recordings from fifty-eight participants were transcribed. Thirty-nine males and nineteen females took part in the experiment. The age of the participants ranged from 18 to 24 years, mean = 22, standard deviation = 1.1. From the total, 74.1% of the participants use the Internet to look for health information whereas 25.9% never use it. The average years of experience of the participants in online health information searching was 3 ± 3.3 SD. Participants were asked about their experience in searching the experimental tasks. From the total, about 71 %, 52%, 72% and 81% of the participants responded that they had searched for dengue, smoking, obesity and physical activity information before, respectively. Table 1 provides detail information about the demographic characteristics of the participants.

Table 1: Demographic information on the participants

Variables	Category	Frequency	Percentage
Sex	Male	39	67.2
	Female	19	32.8
Age	18-21	25	43.1
	22-24	33	56.9
Mother tongue	English	11	19.0
	Others*	47	81.0
Internet health information seeking (HIS)	No	15	25.9
	Yes	43	74.1
Frequency of HIS in the last one month	Median = 1, SD = 1.73		
Internet use experience in hours per day	Median = 6, SD = 4.67		

*Other mother tongues include Arabic, Bahasa Indonesia, Bengali, Cantonese, Creole, Dhivehi, Hakka, Memon, Punjabi and Tamil

Based on the analysis of computer screen recordings, the following aspects of health information searching behaviour were examined: query characteristics (type of query, source of query, spelling errors, number of queries and query length) and result clicking actions (result clicked and

source opened).

Query characteristics

Type of query

In this research study, three types of query were identified: question-based, word-based and word-string-based. Question-based queries contain question words, such as what, how, when, where, which and why ([Jadhav, et al., 2014](#); [Jadhav, Sheth and Pathak, 2014](#)). A word-based query contains a single word, while a word-string query contains more than one word. About 63% of the queries issued were word-strings, while question-based queries constituted 24.5%.

Source of query

While 85.9% of the queries were formulated using keywords from the tasks description, 14.1% used keywords outside the task description. For example, a participant formulates a query of guidelines or recommendations on physical activity. In this query, the keywords (guidelines, recommendations and physical activity) were taken from the task description of physical activity. Another participant formulates a query of moderate intensity aerobic exercise. In this case, the keywords (moderate, intensity, aerobic and exercise) were not taken from the physical activity task description. The task description for physical activity was described under the search tasks of the method section.

Spelling errors

Issuing a query with correct spelling plays a major role in retrieving effective results. In this research study, the presence of spelling errors during query formulation was checked through observation of the computer screen recordings. It was found that only 3% of the queries issued had spelling errors. However, participants used the auto-correction method suggested by search engines for the 30.2% of the queries.

Number of queries and query length

The fifty-eight participants, each of whom completed all four tasks, issued a total of 441 queries to perform the four tasks. The number of queries issued ranged from 1 to 10 across the tasks whereas the average number of queries was 1.9 for the tasks. The query length ranged from 1 to 48 with an average query length of 4 terms for the tasks. Table 2 shows the basic query features stratified by tasks.

Table 2: Basic query features

	Query feature	Task 1 (Dengue)	Task 2 (Smoking)	Task 3 (Obesity)	Task 4 (Physical activity)	All tasks
Number of queries	Mean	1.8	1.7	2.4	1.8	1.9
	Median	1	1	2	1	1
	SD	1.51	1.03	1.77	1.13	1.4

	Min	1	1	1	1	1
	Max	10	5	8	5	10
	Sum	106	94	140	101	441
	Mean	3.66	4.38	4.35	5.55	4.46
	Median	3	4	4	5	4
Query length	SD	2.7	4.5	4.4	5.4	4.38
	Min	1	1	1	1	1
	Max	21	30	25	48	48
	Sum	388	412	609	560	1969
Query type	Word-based	10	14	25	5	54
	Question-based	17	33	25	33	108
	Word-string based	79	47	90	63	279
Spelling error	Yes	7	1	4	1	13
	No	99	93	136	100	428
Source of query	Task description	104	87	115	73	379
	Outside task description	2	7	25	28	62

Multiple linear regression analysis was conducted to examine the effect of contextual features, such as sex, age, mother tongue, health status, health literacy, Internet use experience, task search experience and frequency of health information seeking on the average number of queries and average query length. To fulfil the assumption of normal distribution, square root transformations were carried out for both the number of queries and query length. In addition, outliers for query length were removed (query length >7). Possible multicollinearity is also checked using tolerance and variance inflation factor for each independent variable. A tolerance closer to 0 and/or a variance inflation factor greater than 2 show high multicollinearity ([IBM Corp, 2011](#)). In this study, the tolerances and variance inflation factors for each independent variable were above 0.8 and below 1.36, respectively, which show low multicollinearity. Levene's test of equality of error variances was performed to test for homoscedasticity of errors for all observations ([Levene, 1960](#)). P-values of above 0.123 were found which are not statistically significant, indicating the errors were homoscedastic (i.e., having the same finite variance).

The regression models with all the possible predictors produced $R^2 = 0.071$, $F(1, 56) = 4.29$, $p = 0.043$ for the number of queries and $R^2 = 0.341$, $F(3, 47) = 8.1$, $p < 0.0001$ for query length. Table 3 summarizes the descriptive statistics and analysis results. As can be seen in Table 3, frequency of health information seeking had significant positive regression weights for both the number of queries and query length, indicating that participants with frequent health information seeking are found to issue more queries and longer query length after controlling for other variables in the model. However, mother tongue and health status had a significant negative weight on the query length, indicating that participants whose mother tongue was not English and those who perceived their health status as not healthy were more likely to issue shorter query length. Other contextual features, such as age, sex, health literacy, Internet use experience and task search experience did not contribute to the multiple regression models.

Table 3: Predicting the number of queries and query length using multiple linear regression analysis

Predictors	Category	Frequency (%)	Number of queries		Query length	
			Coefficient	P-value	Coefficient	P-value
Sex	Male	39 (67.2)	0.048	0.712	0.004	0.976
	Female	19 (32.8)				
Mother tongue	English	11 (19.0)	-0.172	0.185	-0.384	<0.0001
	Other	47 (81.0)				
Health literacy	Adequate	46 (79.3)	0.01	0.940	0.012	0.921
	Limited	12 (20.7)				
Health status	Healthy	45 (77.6)	-0.077	0.560	-0.220	-0.025
	Not healthy	13 (22.4)				
Task search experience	Have search experience on tasks	18 (31.0)	0.005	0.970	0.213	0.076
	Have no search experience on tasks	40 (69.0)				
Age			-0.002	0.988	-0.141	0.296
Internet use experience (hours/day)			-0.041	0.755	-0.080	0.513
Frequency of health information seeking			0.049	0.043	0.065	0.006

Result clicking

Of the 441 queries issued to search engines, 85% of them were effective. In this study, a query is said to be effective if and only if one or more results were clicked from the result page. Most of the results clicked were those ranked high on the result list. None of the participants clicked results beyond the first page.

In order to analyse the effects of spelling errors, type of query and source of query on the effectiveness of queries, a logistic regression model was computed. The result showed that spelling errors and source of a query were found to have a significant effect on the effectiveness of queries. Queries with spelling errors were 77% less likely to be effective when compared to queries issued with correct spellings [Adjusted odds ratio (AOR) = 0.23; 95% Confidence interval (CI) (0.07, 0.013)]. Regarding the source of the query, if the whole or part of the words of a query was taken from task description, it was 2.22 times more likely to be effective than a query formulated from words outside the task description [Adjusted odds ratio = 2.22; 95% Confidence interval (1.15, 4.29)]. In this study, the type of query was not found to affect the effectiveness of a query. The analysis output of the logistic regression result is presented in Table 4.

Table 4: Logistic regression analysis showing the effects of spelling errors, type of query and source of a query on the effectiveness of queries

Variables	Category	Result clicked		AOR (95% CI)	P-value
		Yes (%)	No (%)		
Spelling error	Yes	8 (2.1)	5 (7.6)	0.23 (0.07, 0.73)	0.013

	No	367 (97.9)	61 (92.4)	1*	
	Keyword	46 (12.3)	8 (12.1)	1.21 (0.53, 2.78)	0.651
Type of query	Question based	96 (25.6)	12 (18.2)	1.69 (0.85, 3.36)	0.138
	Word strings	233 (62.1)	46 (69.7)	1*	
Source of query	Task description	328 (87.5)	51 (77.3)	2.22 (1.15, 4.29)	0.018
	Outside task description	47 (12.5)	15 (22.7)	1*	

*1 = reference category (a category selected to compare with other category during data analysis)

Source opened

The most frequent search engine used was [Google](#) (98%). The other search engine used were [YouTube](#), [Reddit](#), [Bing](#), [MedlinePlus](#) and [A Calorie Counter](#). Participants used the above search engines to jump to specific Websites. In total, participants visited 289 Websites when performing all the tasks. Among these sites, health-specific sites, such as [smokefree.gov](#), [National Heart, Lung and Blood Institute](#), [Patient](#), [National Health Service \(UK\)](#), [Centers for Disease Control and Prevention](#), [health.gov](#), [American Heart Association](#), [KidsHealth](#), [World Health Organization](#) and [WebMD](#) were found to be the most frequently visited sites. [Wikipedia](#), a general-purpose site, also appeared to be a popular site visited by the participants.

Discussion

This paper attempts to describe the basic query features, identify the contextual features affecting query formulation and explore the result clicking behaviour in the domain of health. In order to address these objectives, we first discuss findings on the basic query features (type of query, number of queries and query length) followed by a discussion on the contextual features affecting query formulation. Then, findings of the result clicking behaviour are discussed.

Our research study identified three types of queries, word-based (12.5%), question-based (24.5%) and word-string-based (63%). Most of the participants used word-string-based queries (63%), implying that the participants had attempted to find specific answers for their information needs. Analyses of query reformulation pattern in previous research studies showed that users issue a single-word query when they want to have a deeper understanding of their information need (Broussard and Zhang, 2013; Rieh and Xie, 2006). On the other hand, users added terms to their original query to find specific information to fulfil their information need (Broussard and Zhang, 2013; Rieh and Xie, 2006). The use of word-string-based queries could indicate a lack of defining search strategies and understanding the level of specificity of the search tasks. Hence, providing a query refining feature to health information retrieval systems could help users to refine their search and retrieve specific answers for their information needs. Ways to provide query refining features could be refining by keywords and health topics. Refining by keyword could help to access results related to the query whereas refining by health topics could help to access results based on specific health topics. Moreover, query suggestions could be provided for users to perform related searches that allows them either to formulate a new query or to add terms to an existing query.

In addition, the question-based query was found to be a popular type of query identified in this study. Out of all queries submitted to search engines, about 25% of them were found to be question-based queries. This indicates participants prefer to express their query or information need in a more natural way using natural language questions. Wildemuth, De Bliet, Friedman and Miya (1994) state that the most natural way for people to seek information is to ask questions. The popularity of question-based queries could indicate the participants' preference to find specific answers to their

information needs. Including a question-answering technique in health information retrieval systems could help users to express their information need in natural language questions and retrieve specific answers for their questions. A question and answering system can automatically analyse documents and return answers in response to users' questions ([Dodiya and Jain, 2013](#)). Information extraction and natural language processing techniques could be employed to provide relevant answers to questions posed by users.

In our study, the average number of queries issued by the participants was 1.9 for the search tasks which indicates fewer reformulation efforts, implying succinct query issuance by the participants. Spink *et al.*, ([2004](#)) point out that health-related searches lack query reformulation which could be attributed to succinct query issuance. On the other hand, the average query length issued to search engines is four terms per query which is longer than the average query length (2.3 terms per query) identified in a research study conducted by Zhang ([2013](#)) on the searching behaviour of students when looking for specific health-related information in MedlinePlus. The use of longer query length could mean that participants express their information need well enough to search engines to retrieve sufficient information to satisfy their need. Longer query length is associated with increased user satisfaction ([Belkin *et al.*, 2003](#); [Lopes and Ribeiro, 2010](#)). In this study, the level of satisfaction with search results was 98%.

Among all possible contextual factors fitted to the multiple linear regression model, only frequency of health information seeking was found to affect the number of queries. Participants who frequently seek health information issue more queries. Their previous experience seems to motivate frequent health information seekers to be exploratory searchers. In exploratory searching, a user issues the first query representing his/her initial understanding of the problem. As the search continues, the user learns new keywords and is triggered to modify the initial query. The searching continues until the user is satisfied with the information found ([Pang, Chang, Pearce and Verspoor, 2014](#)). Hence, frequent health information seeking enables users to explore information in depth. Perhaps, providing related queries in information retrieval systems could help users to select suitable terms to explore and gain a deeper understanding of their information needs.

On the other hand, mother tongue, frequency of health information seeking and health status were found to affect query length. Participants whose mother tongue is English were more likely to issue longer query length. This could be because they can express their ideas better than those whose mother tongue is not English. Frequent health information seekers were also more likely to issue longer query length. This finding is in line with other research studies which indicate that frequent Web searchers have the tendency to formulate longer queries ([Aula, 2003](#); [Lopes and Ribeiro, 2010](#)). Previous research studies indicate that longer queries are associated with increased user satisfaction ([Belkin *et al.*, 2003](#); [Lopes and Ribeiro, 2010](#)). Hence, experienced users, whose search skills are cultivated by their frequent information searching, seem to use longer queries to satisfy their information need. Similarly, participants who perceive their health status as healthier are more likely to issue longer queries. Healthier people are more likely to actively and frequently search for health information to increase their health knowledge and improve their health status ([Pálsdóttir, 2008](#)), implying increased experience and skills in health information searching ([Belkin *et al.*, 2003](#); [Lopes and Ribeiro, 2010](#)). These people could issue longer queries to increase their search effectiveness.

Other contextual features, such as age, sex, health literacy and Internet use experience were not found to influence the number of queries and query length. The possible reason for sex and age not to influence query formulation could be because there is no sex and age difference in health information seeking as identified in a previous research study (Yilma, Inthiran, Reidpath and Orimaye, 2017). Health literacy is presumed to influence query formulation because previous studies point out that people with adequate health literacy level can express their health problems, interact better with their physicians and likely to look for health information ([Ellis *et al.*, 2012](#); [Gutierrez *et al.*, 2014](#)). However, in this study, health literacy is not found to influence query length and number of queries. The reason could be due to the higher health literacy level identified in this study (79.3%). Internet use experience is not also found to influence query length and number of queries in this study. The reason could be because university students presumably have access to the Internet in campuses and hence may have some experience in using the Internet.

Regarding result clicking behaviour, we looked at the effectiveness of queries and the sources opened. Out of all queries issued to search engines, 15% of them were ineffective, that is, none of the results or links were clicked from the result page. Contextual features, such as spelling error and source of query were found to influence the effectiveness of queries. Queries with spelling errors and queries formulated from outside task description contributed to the ineffectiveness of the queries. Issuing a query with correct spellings plays a major role to retrieve effective results. When a user issues queries with spelling errors, the hits returned from search engines could be irrelevant and the user may be disappointed with the search results. As a result, the user is unlikely to click and view results. Hence, it is crucial to provide a function to detect spelling errors and make suggestions during the development of health information retrieval systems. Similarly, a user may not express his/her information need correctly to a search engine, resulting in irrelevant results. Consequently, the user may not click the results. Providing related terms by search engine could help users to express their information need and retrieve relevant results.

In this study, Google, which is a general-purpose search engine, was found to be the most frequently used search engine by participants. General purpose search engines and Websites are useful to support searchers engaged in basic search tasks, but they may not be helpful for searchers who need a deeper understanding of a health topic with multifaceted search tasks that need multiple search sessions and continuous interactions ([Pang et al., 2014](#)). Hence, it is necessary to promote domain-specific search engines or health information retrieval systems, such as [WebMed](#), [MedlinePlus](#) and national health information portals (e.g., [Myhealth](#)) which were rarely used by participants in this study. Results of other research studies in Iran ([Gavvani, Qeisari and Jafarabadi, 2013](#)), Southeast Asia ([Inthiran and Soyiri, 2015](#)), China ([Yuli, Zhang and Xia, 2012](#)) have supported our finding that Google is the most preferred search engine to search for health information.

A total of 289 Websites were viewed by participants when performing the four tasks. Even though health-specific Websites were frequently viewed, general purpose Websites, e.g., Wikipedia, appeared to be popular sites visited by participants. The reason for participants to use Wikipedia frequently could be because it is ranked high in search engines. Moreover, the contents could be easier for participants to understand. In the future, it would be worthwhile to study the role of general-purpose Websites, such as Wikipedia. Another aspect of the Websites visited was that the national health portal, [Myhealth](#), was rarely used. Promoting such Websites among university population could help to get national-specific health information.

Conclusion and recommendation

We conducted a user-based interactive information retrieval study to understand the effect of contextual features on health information searching behaviour. We found interesting results that could have the potential to improve health information retrieval systems.

Query length and number of queries are associated with frequency of health information seeking. Frequent health information seeking leads to more queries enabling users to explore information deeply and increase health knowledge. However, since most of the participants use a general search engine which supports basic searches ([Pang et al., 2014](#)), it is important to promote health-specific search engines that can support exploratory searching.

Mother tongue and health status are the other contextual features found to affect query length. English as mother tongue and being healthy contribute to long query length, indicating better health information retrieval as long query length is associated with increased user satisfaction with search results ([Belkin et al., 2003](#); [Lopes and Ribeiro, 2010](#)).

Regarding the effectiveness of queries, queries with spelling errors and queries formulated outside task descriptions are found to be ineffective. Providing automatic error detection and alternative terms in health information retrieval system could help users to retrieve required information

effectively.

Limitations

This research study has a few limitations. First, the sample size is relatively small which could make it difficult to generalise the study. Moreover, the small sample size may have less power to conduct regression analysis. In future observational studies, we will increase the number of participants so that we have sufficient data to undergo regression analysis. Additionally, the homogeneity with respect to the ‘age of the participants’ can be considered a limitation because it is difficult to conclude whether age is a predictor of query formulation or not. Future research studies should consider a study population with a different age group.

Acknowledgements

The authors would like to thank the participants of this research study.

About the author

Tesfahun Melese Yilma is a doctoral candidate at the School of Information Technology, Monash University, Malaysia. He has a master’s degree in public health informatics from the University of Gondar, Ethiopia and bachelor’s degree in industrial engineering from Mekelle University, Ethiopia. His research interest includes health informatics, health information retrieval, health information system analysis, design and evaluation, evidence-based medicine and data management and analysis. He may be reached at tesfahun.melese@monash.edu

Anushia Inthiran is a Senior Lecturer at the Department of Accounting and Information Systems in the University of Canterbury, New Zealand. She has a PhD in Information Technology from Monash University, a master’s degree in technology management from Staffordshire University, a graduate certificate in higher education from Monash University and a bachelor’s degree in computing from Monash University. Her research interest includes online health search, interactive information retrieval, computer human interaction and library and information science. She may be reached at anushia.inthiran@canterbury.ac.nz.

Daniel D Reidpath is a Professor of Population Health and Public Health at Jeffrey Cheah School of Medicine and Health Sciences, Monash University, Malaysia. He holds a PhD from the University of Western Australia, a diploma of educational psychology from Monash University and a Bachelor of Arts from Swinburne Institute of Technology. His research focus is mainly on the measurement of population health, social stigma (particularly HIV related stigma) and health equity. He may be reached at daniel.reidpath@monash.edu.

Sylvester Olubolu Orimaye works at the College of Public Health in East Tennessee State University, United States of America. He has a PhD in information technology from Monash University, a Master of Philosophy in information technology from Monash University and a BSC in computing from Staffordshire University. His research interest includes Natural Language Processing and Machine Learning. He can be contacted at orimaye@mail.etsu.edu.

References

- Aula, A. (2003). Query formulation in Web information search. In Isaías, P. and Karmakar, N. (Eds.), *Proceedings of the IADIS International Conference WWW/Internet 2003, vol. 1*, (pp. 403-410).

- Belkin, N. J., Kelly, D., Kim, G., Kim, J.-Y., Lee, H.-J., Muresan, G., . . . Cool, C. (2003). [Query length in interactive information retrieval](#). In *Proceedings of the 26th annual international ACM SIGIR conference on Research and development in informaion retrieval, Toronto, Canada, July 28 to August 1, 2003*, (pp. 205-212). New York, NY: ACM. Retrieved from <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.96.3106&rep=rep1&type=pdf>. (Archived by WebCite® at <http://www.webcitation.org/77YdYoiMV>)
- Bierig, R. & Göker, A. (2006). [Time, location and interest: an empirical and user-centred study](#). In *Proceedings of the 1st international conference on Information interaction in context, Copenhagen, Denmark, October 18 - 20, 2006*, (pp. 79-87). New York, NY: ACM. Retrieved from <http://bierig.net/wp-content/papercite-data/pdf/bierig2006timelocationinterest.pdf>. (Archived by WebCite® at <http://www.webcitation.org/77Ydt0uY9>)
- Boden, C. (2014). Overcoming the linguistic divide: a barrier to consumer health information. *Journal of the Canadian Health Libraries Association/Journal De L'Association Des Bibliothèques De La Santé Du Canada*, 30(3), 75-80.
- Borlund, P. (2000). Experimental components for the evaluation of interactive information retrieval systems. *Journal of documentation*, 56(1), 71-90.
- Broussard, R. & Zhang, Y. (2013). Seeking treatment options: consumers' search behaviors and cognitive activities. *Proceedings of the ASIS&T Annual meeting*, 50(1), 1-10
- Chew, M. H., Rahman, M. & Salleh, S. A. (2012). Dengue in Malaysia: an epidemiological perspective study. *Pakistan Journal of Medical Sciences*, 28(3), 643-647.
- Dodiya, T. & Jain, S. (2013). Comparison of question answering systems. In A. Abraham & S. M. Thampi (eds.), *Intelligent informatics: proceedings of the International Symposium on Intelligent Informatics ISI'12 held at August 4-5 2012, Chennai, India* (pp. 99-107). Berlin: Springer.
- Duong, T. V., Aringazina, A., Baisunova, G., Nurjanah, Pham, T. V., Pham, K. M., . . . Chang, P. W. (2017). Measuring health literacy in Asia: validation of the HLS-EU-Q47 survey tool in six Asian countries. *Journal of Epidemiology*, 27(2), 80-86.
- Ellis, J., Mullan, J., Worsley, A. & Pai, N. (2012). [The role of health literacy and social networks in arthritis patients' health information-seeking behavior: a qualitative study](#). *International journal of family medicine*, 2012, Article ID 397039. Retrieved from <https://www.hindawi.com/journals/ijfm/2012/397039/>. (Archived by WebCite® at <http://www.Webcitation.org/76EJEq2x6>)
- Eriksson-Backa, K., Ek, S., Niemelä, R. & Huotari, M. L. (2012). Health information literacy in everyday life: a study of Finns aged 65–79 years. *Health Informatics Journal*, 18(2), 83-94.
- Escoffery, C., Miner, K. R., Adame, D. D., Butler, S., McCormick, L. & Mendell, E. (2005). Internet use for health information among college students. *Journal of American College Health*, 53(4), 183-188.
- European Commission. (2014, November 28). [Europeans becoming enthusiastic users of online health information](#). *Digital Single Market. News*. Retrieved from <https://ec.europa.eu/digital-agenda/en/news/europeans-becoming-enthusiastic-users-online-health-information>. (Archived by WebCite® at <http://www.Webcitation.org/76EJZTY41>)
- Feinberg, I., Greenberg, D. & Frijters, J. (2015). [Understanding health information seeking behaviors of adults with low literacy, numeracy and problem solving skills: results from the 2012 US PIAAC study](#). Washington, DC: US Program for the Assessment of Adult Competencies. Retrieved from https://piaac.squarespace.com/s/Feinberg_Greenberg_Frijters_PIAAC.pdf. (Archived by WebCite® at <http://www.Webcitation.org/76EK7t0fE>)
- Fenichel, C. H. (1981). Online searching: measures that discriminate among users with different types of experiences. *Journal of the American Society for Information Science*, 32(1), 23-32.
- Fiksdal, A. S., Kumbamu, A., Jadhav, A. S., Cocos, C., Nelsen, L. A., Pathak, J. & McCormick, J. B. (2014). Evaluating the process of online health information searching: a qualitative approach to exploring consumer perspectives. *Journal of Medical Internet Research*, 16(10), e224.
- Fox, S. & Duggan, M. (2013). [Health online 2013](#). Washington, DC: Pew Research Institute. Retrieved from <http://www.pewinternet.org/2013/01/15/health-online-2013/>. (Archived by WebCite® at <http://www.Webcitation.org/76EKIfT9T>)

- Gavvani, V. Z., Qeisari, E. & Jafarabadi, M.A. (2013). [Health information seeking behavior: a study of a developing country](#). *Library Philosophy and Practice (e-journal)*, Paper 902. Retrieved from <http://digitalcommons.unl.edu/libphilprac/902>. (Archived by WebCite® at <http://www.Webcitation.org/76EKVk1Np>)
- Gutierrez, N., Kindratt, T.B., Pagels, P., Foster, B. & Gimpel, N.E. (2014). Health literacy, health information seeking behaviors and internet use among patients attending a private and public clinic in the same geographic area. *Journal of Community Health*, 39(1), 83-89.
- Guzman, A. & Istúriz, R.E. (2010). Update on the global spread of dengue. *International Journal of Antimicrobial Agents*, 36(Supplement 1), S40-S42.
- Harrell, F.E.A. (2015). *Regression modeling strategies: with applications to linear models, logistic and ordinal regression and survival analysis*. (2nd ed.). Cham, NY : Springer.
- Hu, R., Lu, K. & Joo, S. (2013). [Effects of topic familiarity and search skills on query reformulation behavior](#). *Proceedings of the ASIS&T annual meeting*, 50(1), 1-9. Retrieved from <https://onlinelibrary.wiley.com/doi/full/10.1002/meet.14505001062>. (Archived by WebCite® at <http://www.webcitation.org/77jKHap2D>)
- IBM Corp. (2011). *IBM SPSS Statistics for Windows (Version 20.0)*. Armonk, NY: IBM Corporation.
- Ingwersen, P. & Järvelin, K. (2005). Information retrieval in context – IRiX. *ACM SIGIR Forum* 39(2), 31-39.
- Inthiran, A., Alhashmi, S., & Ahmed, P. (2011a). Investigating interactive search behaviour of medical students: an exploratory survey. In *Proceedings of the 23rd Australian Computer-Human Interaction Conference, Canberra, Australia — November 28 - December 02, 2011* (pp. 156-165). New York, NY: ACM.
- Inthiran, A., Alhashmi, S. & Ahmed, P. (2011b). Profiling a non-medical professional searcher on a medical domain: what do search patterns and demographic details reveal? In M. Salem, K. Shaalan, F. Oroumchian, A. Shakery, and H. Khelalfa (Eds.), *Information retrieval technology* (pp. 157-168). Berlin: Springer. (Lecture Notes in Computer Science, volume 7097).
- Inthiran, A., Alhashmi, S.M. & Ahmed, P.K. (2015). A user study on the information search behaviour of medical students. *Malaysian Journal of Library and Information Science*, 20(1), 61-77.
- Inthiran, A. & Soyiri, I. (2015). Searching for health information online for my child: a perspective from South East Asia. In B. R. Allen, J. Hunter and L. M. Zeng (Eds.), *Digital Libraries: Providing Quality Information: 17th International Conference on Asia-Pacific Digital Libraries, ICADL 2015, Seoul, Korea, December 9-12, 2015. Proceedings* (pp. 76-81). Cham, Switzerland: Springer International Publishing.
- Jadhav, A., Andrews, D., Fiksdal, A., Kumbamu, A., McCormick, J. B., Misitano, A., ... & Pathak, J. (2014). [Comparative analysis of online health queries originating from personal computers and smart devices on a consumer health information portal](#). *Journal of Medical Internet Research*, 16(7). Retrieved from <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4115262/>. (Archived by WebCite® at <http://www.Webcitation.org/76EKsHXek>)
- Jadhav, A., Sheth, A. & Pathak, J. (2014). Analysis of online information searching for cardiovascular diseases on a consumer health information portal. In *AMIA Annual Symposium Proceedings, 2014*, (pp. 739-748). Bethesda, MD: American Medical Informatics Association.
- Janeice, A., Eileen, M. & Trauth, M. (2013). Socio-economic influences on health information searching in the USA: the case of diabetes. *Information Technology and People*, 26(4), 324-346.
- Jones, K.S. (1981). *Information retrieval experiment*. London: Butterworths.
- Jung, M. (2014). Determinants of health information-seeking behavior: implications for post-treatment cancer patients. *Asian Pacific Journal of Cancer Prevention*, 15(16), 6499-6504.
- Kelly, D. (2006). Measuring online information seeking context, part 1: background and method. *Journal of the American Society for Information Science and Technology*, 57(13), 1729-1739.
- Kelly, D. (2009). [Methods for evaluating interactive information retrieval systems with users](#). *Foundations and Trends in Information Retrieval*, 3(1-2), 1-224. Retrieved from https://ils.unc.edu/courses/2019_spring/inls509_001/papers/FnTIR-Press-Kelly.pdf. (Archived by WebCite® at <http://www.webcitation.org/77jMpl9ug>)

- Kim, S. (2015). An exploratory study of inactive health information seekers. *International Journal of Medical Informatics*, 84(2), 119-133.
- Lambert, S.D. & Loiselle, C.G. (2007). Health information seeking behavior. *Qualitative Health Research*, 17(8), 1006-1019.
- Levene, H. (1960). Robust testes for equality of variances. In I.E. Olkin (Ed.), *Contributions to probability and statistics: essays in honor of Harold Hotelling* (pp. 278–292). Stanford, CA: Stanford University Press.
- Lopes, C. T. & Ribeiro, C. (2010). Context effect on query formulation and subjective relevance in health searches. In *Proceedings of the third symposium on Information interaction in context, New Brunswick, New Jersey, USA — August 18 - 21, 2010* (pp. 205-214). New York, NY: ACM.
- Luo, G., Tang, C., Yang, H. & Wei, X. (2008). MedSearch: a specialized search engine for medical information retrieval. In *Proceedings of the 17th ACM conference on Information and knowledge management, Napa Valley, California, USA — October 26 - 30, 2008* (pp. 143-152). New York, NY: ACM.
- Malaysia. *Institute for Public Health* (2012). [Global adult tobacco survey. Malaysia 2011](https://www.who.int/tobacco/surveillance/survey/gats/malaysia_country_report_2011.pdf). Kuala Lumpur: Ministry of Health, Institute for Public Health. Retrieved from https://www.who.int/tobacco/surveillance/survey/gats/malaysia_country_report_2011.pdf. (Archived by WebCite® at <http://www.webcitation.org/77jOPhDDn>)
- Malaysia. *Ministry of Health*. (2011). [NCD prevention and control: annual report](http://www.moh.gov.my/images/gallery/nspncd/AnnualReport2010NCD.pdf). Putrajaya, Malaysia: Ministry of Health. Retrieved from <http://www.moh.gov.my/images/gallery/nspncd/AnnualReport2010NCD.pdf>. (Archived by WebCite® at <http://www.Webcitation.org/76EL4Pzw5>)
- Medlock, S., Eslami, S., Askari, M., Arts, D. L., Sent, D., de Rooij, S. E. & Abu-Hanna, A. (2015). Health information-seeking behavior of seniors who use the internet: a survey. *Journal of Medical Internet Research*, 17(1), e10.
- Mohamad, E., Su, T. T., Majid, H. A. & Chinna, K. (2014). *Health literacy in Malaysia: HLS-Asia questionnaire validation*. Paper presented at the The Second International Conference on Health Literacy and Health Promotion, Taipei, Taiwan.
- Pálsdóttir, Á. (2008). [Information behaviour, health self-efficacy beliefs and health behaviour in Icelanders' everyday life](http://www.informationr.net/ir/13-1/paper334.html). *Information Research*, 13(1), Paper 334. Retrieved from <http://www.informationr.net/ir/13-1/paper334.html>. (Archived by WebCite® at <http://www.Webcitation.org/76ENpGSDz>)
- Pang, P. C. I., Chang, S., Pearce, J. M. & Verspoor, K. (2014). Online health information seeking behaviour: understanding different search approaches. In *Proceedings of the Pacific Asia Conference on Information Systems (PACIS), 2014, Chengdu, China* (paper 229). Atlanta, GA: Association for Information Systems.
- Puspitasari, I., Moriyama, K., Fukui, K. I. & Numao, M. (2015). [Effects of individual health topic familiarity on activity patterns during health information searches](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4381811/). *Journal of Medical Internet Research*, 3(1), e16. Retrieved from <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4381811/>. (Archived by WebCite® at <http://www.webcitation.org/77jNElwnJ>)
- Ramprasath, M. & Hariharan, S. (2014). Query refinement based question answering system using pattern analysis. *Advances in Natural and Applied Sciences*, 8(17), 8-15.
- Rieh, S. Y. & Xie, H. (2006). Analysis of multiple query reformulations on the Web: the interactive information retrieval context. *Information Processing and Management*, 42(3), 751-768.
- Shepard, D. S., Lees, R., Ng, C. W., Undurraga, E. A., Halasa, Y. & Lum, L. (2013). [Burden of dengue in Malaysia: report from a collaboration between universities and the Ministry of Health of Malaysia](http://people.brandeis.edu/~shepard/Report_dengue_in_Malaysia_v50.pdf). Waltham, MA: Brandeis University. Retrieved from http://people.brandeis.edu/~shepard/Report_dengue_in_Malaysia_v50.pdf. (Archived by WebCite® at <http://www.Webcitation.org/76ELJhU1D>)
- Sørensen, K., Van den Broucke, S., Pelikan, J. M., Fullam, J., Doyle, G., Slonska, Z., ... and Brand, H. (2013). Measuring health literacy in populations: illuminating the design and development process of the European Health Literacy Survey Questionnaire (HLS-EU-Q). *BMC Public Health*, 13(1), 948.

- Spink, A., Yang, Y., Jansen, J., Nykanen, P., Lorence, D. P., Ozmutlu, S. & Ozmutlu, H. C. (2004). A study of medical and health queries to Web search engines. *Health Information and Libraries Journal*, 21(1), 44-51.
- Sterne, J. & Kirkwood, B. (2003). *Essential medical statistics*. Hoboken, NJ: John Wiley and Sons, Ltd.
- Sutcliffe, A. & Ennis, M. (1998). Towards a cognitive theory of information retrieval. *Interacting with Computers*, 10(3), 321-351.
- Wang, M. P., Viswanath, K., Lam, T. H., Wang, X. & Chan, S. S. (2013). [Social determinants of health information seeking among Chinese adults in Hong Kong](https://doi.org/10.1371/journal.pone.0073049). *PloS one*, 8(8), e73049. Retrieved from <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0073049>. (Archived by WebCite® at <http://www.webcitation.org/77jNbseHn>)
- Weaver, J. B., Mays, D., Weaver, S. S., Hopkins, G. L., Eroğlu, D. & Bernhardt, J. M. (2010). Health information-seeking behaviors, health indicators and health risks. *American Journal of Public Health*, 100(8), 1520-1525.
- Wildemuth, B. M., De Bliek, R., Friedman, C. P. & Miya, T. S. (1994). Information-seeking behaviors of medical students: a classification of questions asked of librarians and physicians. *Bulletin of the Medical Library Association*, 82(3), 295-304.
- World Health Organization. (2014). Noncommunicable diseases (NCD) country profiles: Malaysia. In Malaysia (pp. 1). Geneva, Switzerland: WHO.
- World Health Organization. (2015a). [Obesity and overweight](http://www.who.int/mediacentre/factsheets/fs311/en/). Geneva, Switzerland: World Health Organization. Retrieved from <http://www.who.int/mediacentre/factsheets/fs311/en/>. (Archived by WebCite® at <http://www.Webcitation.org/76ELV28sa>)
- World Health Organization. (2015b). [Update on the Dengue situation in the Western Pacific Region](http://www.wpro.who.int/emerging_diseases/dengue_biweekly_20151103.pdf?ua=1). Manila, Philippines: World Health Organization, Western Pacific Region. (Dengue Situation Update 477) Retrieved from http://www.wpro.who.int/emerging_diseases/dengue_biweekly_20151103.pdf?ua=1. (Archived by WebCite® at <http://www.Webcitation.org/76ELdXqe6>)
- World Health Organization. (2015c). [WHO report on the global tobacco epidemic, 2015. Raising taxes on tobacco](http://apps.who.int/iris/bitstream/10665/178574/1/9789240694606_eng.pdf?ua=1&ua=1). Geneva, Switzerland: World Health Organization. Retrieved from http://apps.who.int/iris/bitstream/10665/178574/1/9789240694606_eng.pdf?ua=1&ua=1. (Archived by WebCite® at <http://www.Webcitation.org/76ELoKY6R>)
- Yeganova, L., Comeau, D. C., Kim, W. & Wilbur, W. J. (2009). How to interpret PubMed queries and why it matters. *Journal of the American Society for Information Science and Technology*, 60(2), 264-274.
- Yi, Y. J. (2015). Health literacy and health information behavior of Florida public library users: a mixed methods study. *Journal of Librarianship and Information Science*, 47(1), 17-29.
- Yilma, T. M., Inthiran, A., Reidpath, D. & Orimaye, S. O. (2017). Health Information seeking and its associated factors among university students: a case in a middle-income setting. In *Proceedings of the Pacific Asia Conference on Information Systems (PACIS), 2017, Langkawi, Malaysia*, (paper 229). Atlanta, GA: Association for Information Systems.
- Yuli, Y., Zhang, S. & Xia, L. (2012). Preventive health information seeking behavior among general population in China. In *Proceedings of the 2012 International Symposium on Information Technology in Medicine and Education* (Vol. 2, pp. 608-612). New York, NY: IEEE.
- Zhang, Y. (2013). Searching for specific health-related information in MedlinePlus: behavioral patterns and user experience. *Journal of the Association for Information Science and Technology*, 65(1), 53-68.
- Zhang, Y., Wang, P., Heaton, A. & Winkler, H. (2012). [Health information searching behavior in MedlinePlus and the impact of tasks](https://www.ischool.utexas.edu/~yanz/fp088_Zhang.pdf). In *Proceedings of the 2nd ACM SIGHIT international health informatics symposium, January 28-30, 2012, Miami, Florida, USA*. (pp. 641-650). New York, NY: ACM. Retrieved from https://www.ischool.utexas.edu/~yanz/fp088_Zhang.pdf. (Archived by WebCite® at <http://www.webcitation.org/77jOfiRj6>)

Yilma, T.M., Inthiran, A., Reidpath, D.D. & Orimaye, S.O. (2019). Context-based interactive health information searching *Information Research*, 24(2), paper 815. Retrieved from <http://InformationR.net/ir/24-2/paper815.html> (Archived by WebCite® at <http://www.webcitation.org/78mjB8y17>)

Find other papers on this subject

Scholar Search

Google Search

Bing

Check for citations, [using Google Scholar](#)

Facebook

Twitter

LinkedIn

More

© the authors, 2019.

14 Last updated: 10 April, 2019

-
- [Contents](#) |
 - [Author index](#) |
 - [Subject index](#) |
 - [Search](#) |
 - [Home](#)
-