<u>Original Paper</u>

Analysis of PubMed User Sessions Using a Full-Day PubMed Query Log: A Comparison of Experienced and Nonexperienced PubMed Users

Illhoi Yoo[1,2*], PhD; Abu Saleh Mohammad Mosa[2,3*], M. S.

[1]Department of Health Management and Informatics, School of Medicine, University of Missouri, Columbia, MO, United States
[2]Informatics Institute, University of Missouri, Columbia, MO, United States
[3]Institute for Clinical and Translational Science, School of Medicine, University of Missouri, Columbia, MO, United States
[*]all authors contributed equally

Corresponding Author:
Illhoi Yoo, PhD
Department of Health Management and Informatics
School of Medicine
University of Missouri
Five Hospital Dr.
CE718 Clinical Support and Education Building (DC006.00)
Columbia, MO, 65212
United States
Phone: 1 573 882 7642
Fax: 1 573 882 6158
Email: yooil@health.missouri.edu

Abstract

Background: PubMed is the largest biomedical bibliographic information source on the Internet. PubMed has been considered one of the most important and reliable sources of up-to-date health care evidence. Previous studies examined the effects of domain expertise/knowledge on search performance using PubMed. However, very little is known about PubMed users' knowledge of information retrieval (IR) functions and their usage in query formulation.

Objective: The purpose of this study was to shed light on how experienced/nonexperienced PubMed users perform their search queries by analyzing a full-day query log. Our hypotheses were that (1) experienced PubMed users who use system functions quickly retrieve relevant documents and (2) nonexperienced PubMed users who do not use them have longer search sessions than experienced users.

Methods: To test these hypotheses, we analyzed PubMed query log data containing nearly 3 million queries. User sessions were divided into two categories: experienced and nonexperienced. We compared experienced and nonexperienced users per number of sessions, and experienced and nonexperienced user sessions per session length, with a focus on how fast they completed their sessions.

Results: To test our hypotheses, we measured how successful information retrieval was (at retrieving relevant documents), represented as the decrease rates of experienced and nonexperienced users from a session length of 1 to 2, 3, 4, and 5. The decrease rate (from a session length of 1 to 2) of the experienced users was significantly larger than that of the nonexperienced groups.

Conclusions: Experienced PubMed users retrieve relevant documents more quickly than nonexperienced PubMed users in terms of session length.

XSL•FO
RenderX

Introduction

Background

Methods of information seeking have become much easier, faster, and inexpensive since the 1990s with the advent of information technologies (ITs) including the Internet, digital libraries (eg, electronic full-text databases), and online search software/services such as Google Scholar and PubMed. [1-3]. Since then, immense change in scientific-information-seeking behavior has been observed, including among professionals, scholars, and scientists in the area of biomedical and health sciences [3-6]. There is unprecedented growth of biomedical information, which has been doubling every 5 years [7,8]. This large amount of scientific information from multiple sources (eg, journals) is currently integrated in electronic bibliographic databases and accessible through online search software [3,9]. For example, PubMed, which is maintained by the United States National Library of Medicine (NLM), is one of the largest and most authoritative online biomedical bibliographic databases in the world [10-12]. As of June 2015, PubMed contained more than 24 million citations and abstracts from approximately 5600 biomedicine and health-related journals. Health care professionals consider PubMed to be one of the most important and reliable sources of up-to-date health care evidence [13,14]. PubMed also plays a very important role in the process of literature-based discovery [15].

Recent years have seen a rising trend in biomedical information seeking from PubMed [16,17]. About two-thirds of PubMed users are domain experts (eg, health care professionals) and one-third are lay people [18]. Previous studies have examined the effects of domain expertise/knowledge on search performance using PubMed [6,19-21]. However, very little is known about PubMed users' knowledge of information retrieval (IR) functions and their usage in query formulation.

The goal of this study was to shed light on how PubMed users perform their search queries by analyzing a full-day query log. The hypotheses of this study were that (1) experienced PubMed users who use system functions such as Medical Subject Heading (MeSH) terms and search field tags quickly retrieve relevant documents and (2) nonexperienced PubMed users who do not use them have longer search sessions than experienced users, because they identify their information needs through subsequent queries by narrowing and/or broadening their queries. In order to test the hypotheses, we analyzed a full day of PubMed log data. We assumed that if a session was closed within a few queries, the session was successful (meaning that relevant documents were retrieved), even if a session close did not always mean successful IR.

In this study, experienced PubMed users were defined as users who used advanced PubMed IR functions for query formulation. The proper use of IR functions (described in the next section) is key for efficient and effective PubMed searches [6,8,22-27] because, unlike Google, PubMed does not sort search results by relevance. Studies have shown that experienced users are more likely to use IR functions than novice users. Xie and Joo (2012) [28] performed a study on factors affecting the selection of search tactics and demonstrated that expert participants were

more willing to use advanced IR functions. The study [28] used the definition of expert IR users from Holscher and Strube (2000) [29], in which expert users were defined as users having the "knowledge and skills" necessary to utilize information-seeking systems successfully. Holscher and Strube (2000) [29] also recognized that "expert users use advanced IR functions much more than average users." Earlier studies also demonstrated that experienced searchers are more knowledgeable of the content and structure of the IR system and more likely to interact with the system [30,31]. Penniman (1981) [32] defined experienced PubMed users based on the frequency of PubMed searches and concluded that experienced searchers use more search functions than nonexperienced searchers. In addition, many studies have demonstrated that experienced users use more advanced IR functions and show better IR performance than novices [33-37].

PubMed System Functions

PubMed system functions include search field tags, MeSH terms (used for indexing PubMed articles), truncation, and combining searches using search history. In PubMed, bibliographic information is stored in a structured database with 65 fields including title, abstract, author, journal or proceeding, publication type, and publication date. PubMed provides 48 search field tags in order to facilitate searching in its various database fields; a description for each search field is available at the NLM website [38] (last revised and updated November 2012). Thus, PubMed is a field-oriented search system in which search terms are tagged with search field tags and appended using Boolean operators (ie, AND, OR, and NOT). Using search field tags, PubMed users can limit the search to a specific field for each search keyword. A search field tag is attached to a search term by enclosing the search field name in square brackets (eg, "myocardial infarction" [Title]). The NLM indexes PubMed documents using the MeSH vocabulary after indexers read full papers (not just abstracts). Usually, 5-10 MeSH terms are assigned to a PubMed document. Truncation is used to search for the first 600 variations of a truncated word in PubMed. However, PubMed allows an asterisk (*) at the end of a word only; "?" is not used in PubMed. For example, the search term nutrition* will search for nutritional and nutritionists. Finally, the combining search function using search history enables PubMed users to readily use and combine previous search results using Boolean operators and search history indexes. For example, after a PubMed search for *diabetes mellitus*, the search result can be readily combined with one using a new search keyword *hypertension*: #1 AND *hypertension* (#1 indicates *diabetes mellitus*).

Related Studies

The study of information-seeking behavior is very important for the user-centric design of online IR systems including digital libraries. Individuals' knowledge and skills related to information seeking are the primary determinants of their online IR performance. According to Marchionini (1995) [39], there are four types of expertise that determine information-seeking performance: general cognitive abilities, domain knowledge, overall experience of online information seeking, and experience or knowledge of the functions of the IR system. Most intellectual

activities like the information-seeking process involve planning (eg, query term selection), progress monitoring (eg, the number of returned documents), decision making (eg, when to continue or stop the search), and reflecting on past activities (eg, refining the search query for a better search result). Marchionini (1995) [39] stated that people's perceptual and cognitive processes (known as cognitive abilities) are used in completing these tasks. As a common expectation, a person with higher cognitive abilities should perform better at information seeking than someone with lower cognitive abilities. However, few studies have investigated which cognitive abilities are linked to information seeking performance [1,29,39-42]. Hersh et al (2002) [24] assessed three cognitive factors (spatial visualization, logical reasoning, and verbal reasoning) that were found to affect IR performance, and found that PubMed/MEDLINE search experience and spatial visualization were the main factors in successful PubMed searches.

The second major area of expertise is the knowledge of information seekers in their area of interest (known as domain knowledge). The NLM reported that almost two-thirds of PubMed users are health care professionals and scientists (ie, domain experts), whereas the remainder are the general public [18]. Studies have demonstrated that methods of conducting information seeking tasks by domain experts are different from those of novice users [1,5]. In addition, overall IR performance of domain experts is better than that of novice users in various IR systems such as web and hypertext searches [29,42-46], and online bibliographic database searches [33,42,47]. A similar result has also been observed for PubMed searches [20,48]. PubMed search studies demonstrated that PubMed users with domain knowledge usually spent less time and retrieved more information than PubMed users with less domain knowledge. On the other hand, some studies measured user-searching performance (in terms of recall and precision) and concluded that domain knowledge did not significantly affect information-seeking performance. These studies were performed with the DIALOG database [49], an online library catalog [50], and the MEDLINE search system [19-21].

The other two determinants of search performance (ie, overall experience using online information seeking and experience or knowledge of the functions of the IR system) can be considered together as procedural knowledge for using the IR system [6]. Previous studies have demonstrated that such experience improves IR performance for various search systems such as web, hypertexts, file collections, and bibliographic DBs including PubMed [21,24,35,42,44,45,51]. Egan (1988) [52], Hölscher and Strube (2000) [29], and Jenkins et al (2003) [44] found that domain knowledge helped to improve search performance only if users had sufficient procedural knowledge including experience with online searching and search software/systems. In their literature review, Vibert et al (2009) [6] mainly compared the effects of domain knowledge on PubMed searches between expert and novice groups, and demonstrated that domain knowledge does not help to improve search performance if users do not have procedural knowledge. In addition, the study [6] suggested that knowledge in a broad scientific field can compensate for a lack of knowledge in a specific domain, and that the main determinant of bibliographic

search performance is individual cognitive abilities. Thus, people with basic domain knowledge in their area of interest, higher cognitive abilities, and sufficient procedural knowledge regarding the bibliographic search system should efficiently perform information-seeking tasks (eg, query selection and decisions about search discontinuation). Some recent studies found that most academic researchers and health care professionals including physicians do not use advanced IR functions but only natural language for PubMed searches [6,51,53-55]. Another very recent study of PubMed by Macedo-Route et al (2012) [56] concluded that the way researchers use PubMed is nearly the same as the way IR novices do ("mostly typing a few keywords and scanning the titles retrieved by the tool"). Several studies have shown that medical librarians (considered experienced users in the study) use more IR functions for PubMed searches and their IR performance is better than regular users [20,36,57,58].

In this study, our goal was to compare experienced versus nonexperienced users' searching behavior in terms of session length (ie, the number of queries per session). We used a full-day PubMed query log for that purpose. There are a number of approaches for studying user-searching behavior such as eye tracking, surveys, and search log analysis. Search log analysis has become a viable solution for many applications including search engines [16,17,59-63]. One major advantage of search log analysis over other methods is that actual searches by a large number of real users can be analyzed, while other methods usually examine searches from only tens up to hundreds of users. A search engine stores users' query text along with other information including user IP addresses in query log files.

Silverstein et al (1999) [59] and Jansen et al (2000) [60] analyzed a query log from the AltaVista and Excite web search engines, respectively. Silverstein et al (1999) [59] reported three important facts: (1) users rarely navigate beyond the first page of search results, (2) they rarely resubmit a refined query (similar to Jansen et al (2000)'s [60] finding), and (3) most queries are short in length. Herskovic et al (2007) [16] carried out a similar study with a PubMed log and reported statistical information on PubMed usage (including the number of users, queries per user, sessions per user, and frequently used search terms and search field tags). The PubMed log data were used for segmenting query sessions [64], evaluating the PubMed Automatic Term Mapping (ATM) [65], and annotating PubMed queries using the Unified Medical Language System (UMLS) [66]. NLM researchers used month-long PubMed log data for categorizing PubMed queries [17,66], creating a query suggestion database [67], and identifying related journals for user queries [68]. Both of the full-day-long and month-long datasets are publicly available. However, the month-long dataset does not contain actual user queries. For this reason, we used the full-day-long PubMed log data.

The focus of this study is different from that of the eight studies that used PubMed log data [16,17,63-68]. We focused on comparing experienced versus nonexperienced users' searching behavior in terms of session length (the number of queries in a session). To the best of our knowledge, there is no study with this focus.

Methods

Data Cleaning and Preprocessing

The dataset used in this study is a plain text file containing a full-day's query log of PubMed that was obtained from the NLM FTP site (Refer to [69] to access the data). There are nearly 3 million queries issued by 626,554 distinct users.

The data cleaning and preprocessing steps are presented in Figure 1. We found 1146 records with empty user IDs, 76 records with unusual user IDs (we believe they were errors), and 77,923 records with no user-query text. These records (79,145/2,996,301, 2.64%) were eliminated from the dataset.

Figure 1. Data cleaning and preprocessing.
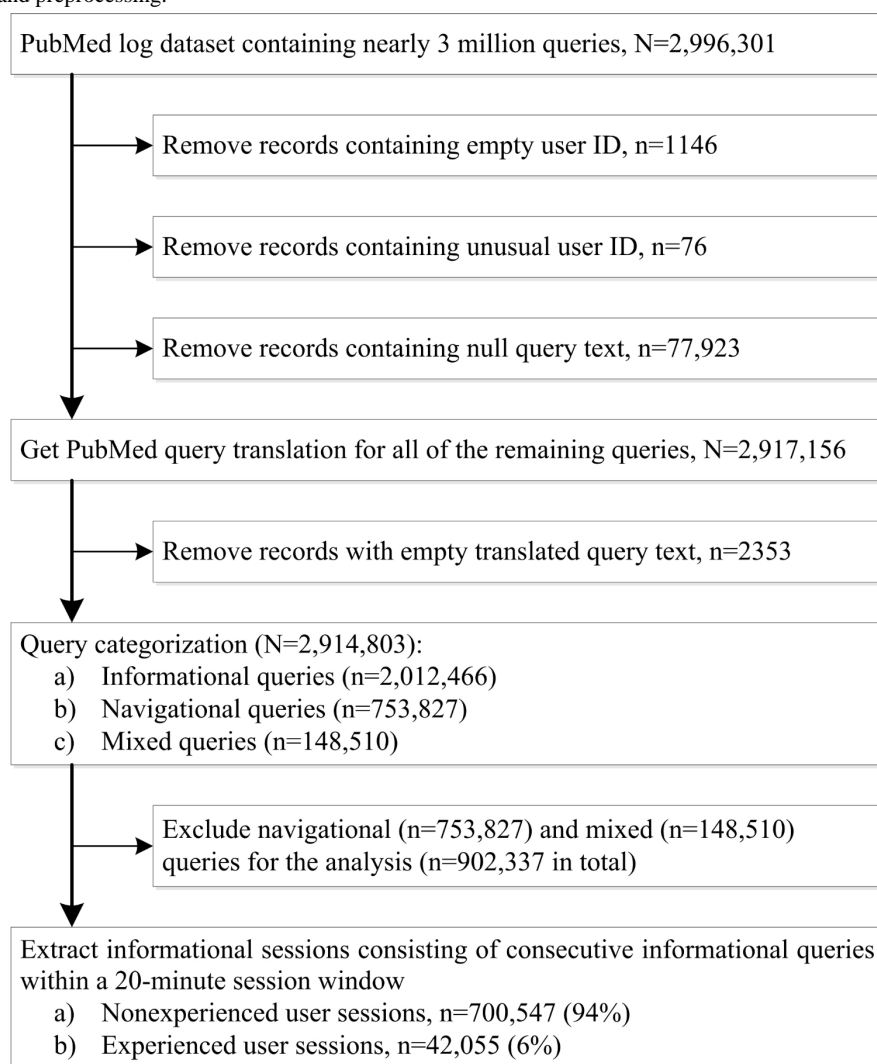
Query Categorization

The user queries in the PubMed log file are categorized as informational, navigational, or mixed according to the purpose of the search expressed in the query. Informational queries are intended to fulfill end users' information needs (eg, "diabetes mellitus" [MeSH]) and navigational queries are intended to retrieve specific documents (eg, Yoo [author] AND Mosa [author]). Mixed queries have both intentions (eg, searching for a specific topic within a specific journal). Refer to Broder (2002) [70] and Herskovic et al (2007) [16] for details of web search types and PubMed search types, respectively.

In order to identify the purpose of user queries for query categorization, we used PubMed's ATM. Every PubMed user query is automatically translated by ATM to improve overall IR performance and the translated query is actually used for the

PubMed search; if a query contains double quotation marks or search tags, those parts (words or terms) are not translated. The ATM translation identifies each term in a query and adds an appropriate search tag to the term. We categorized PubMed queries using ATM-added tags as well as user-added tags after ATM translations. PubMed provides 48 search tags (refer to the PubMed Help website [71] for details), which are classified into informational and navigational tags [69]. Queries containing only informational tags are identified as informational queries. Navigational queries are queries containing navigational or citation-related tags. Queries containing both informational and navigational tags are identified as mixed queries, unless the original query contains an indication of a navigational query. Figure 2 presents a flow diagram for query categorization. A total of 2353 queries resulted in empty query translation. These
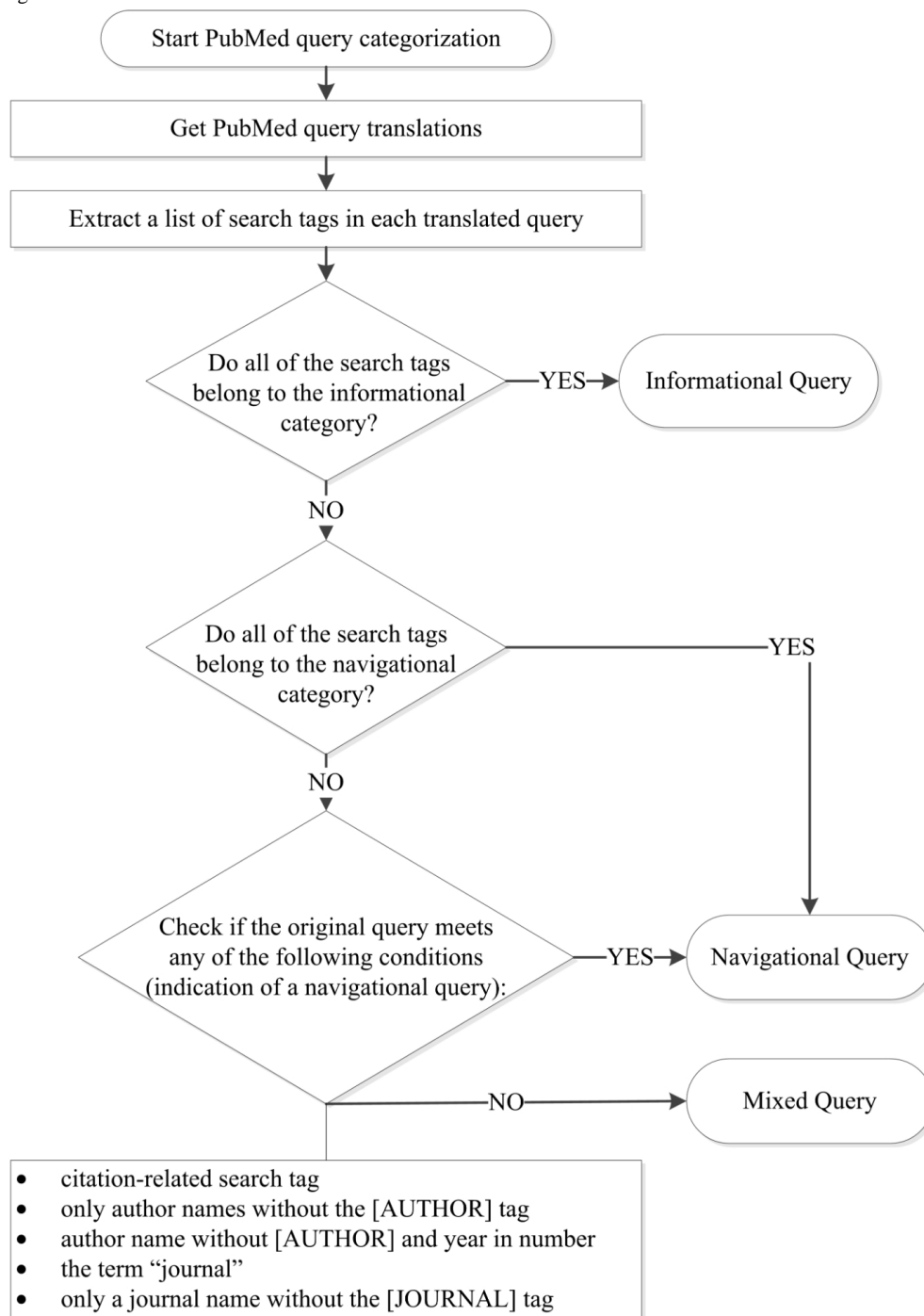
were removed from the analysis. The translated query texts were then parsed to extract the search tags.

The search tag extraction process involved a semiautomatic approach consisting of two steps: the semiautomatic construction of a list of search tags and their variations, and the automatic extraction of the search tags including their variations from the queries using the search tag list. A total of 963 unique substrings were extracted from the queries in the first step. The first step (a partial manual step) was required for two reasons: (1) for each search tag there are several variations that are not fully documented even though they are correctly recognized by the PubMed system; for example, [Author Name], [Author], [AU

Name], [Auth], and [AU] represent the same search tag header but only [Author Name] and [AU] are documented in the PubMed Help web page, and (2) incorrect search tags (eg, typos like [Atuhor]) used in PubMed queries are not recognized by the PubMed system but a domain expert could correctly recognize and read those intentions. The extracted search tags from the translated queries were then analyzed to identify query types. Since navigational search tags are mainly used to retrieve specific documents rather than to fulfill information needs, we excluded navigational and mixed queries from the analysis, assuming informational search tags are primarily used for information needs.

Figure 2. Query categorization.

- citation-related search tag
- only author names without the [AUTHOR] tag
- author name without [AUTHOR] and year in number
- the term "journal"
- only a journal name without the [JOURNAL] tag

Session Segmentation

Information seeking is defined as "the process of repeatedly searching over time in relation to a specific, but possibly an evolving information problem" [72]. Swanson et al (1977) [73] defined information seeking as a trial-and-error process, in which the initial search query is refined at every step, based on the search results in the previous queries. IR users often perform multiple queries in a row for the same information problem. The IR community has coined the term session in this regard. Silverstein et al (1999) [59] defined a session as "a series of queries by a single user made within a small range of time; a session is meant to capture a single user's attempt to fill a single information need." In order to segment queries by a user into sessions, most studies utilized temporal clues such as temporal threshold (ie, time cutoff) between two consecutive queries [59,74-78] or temporal constraint [79] (Refer to a recent survey article by Gayo-Avello (2009) [80] for details). This process (ie, session segmentation) provides valuable insights into users' search behavior and interactions with the IR system.

In this study, we employed both the session-shift and temporal-constraint-based sliding window for session segmentation. This is because several studies reported the average duration of user sessions for query log analysis (meaning that the maximum length of session window can be chosen based on those results for session segmentation) [81-83]. In our study, we set the maximum length of the sliding window to be 20 minutes. The choice of a 20-minute session window was based on two biomedical IR studies. The first was a qualitative study with human subjects that showed most PubMed users successfully completed their task within a 15-minute period, whereas many took more than 15 minutes [6]. The second was a randomized controlled trial on biomedical information retrieval demonstrating that the average time to solve a biomedical information problem ranges from 14 to 17 minutes [84]. In addition to temporal constraint, we used change of query types as session shift. As a result, a change from an informational query to a navigational query was considered a session boundary.

Using this method, we extracted 742,602 user sessions from more than 2 million informational queries. User sessions were divided into two categories: experienced and nonexperienced. Experienced sessions were those in which queries were formed using system functions such as MeSH terms and search field tags. Otherwise, a user session was considered nonexperienced. For example, while a query containing "hypertension [MeSH]" was considered experienced, a query with "high blood pressure" was considered nonexperienced, even though hypertension is a synonym of high blood pressure. This is because although for the query "high blood pressure," PubMed's ATM internally expands the query by adding the MeSH term *hypertension*, the MeSH term is ORed with the term high blood pressure (i.e., hypertension [MeSH] OR high blood pressure) and the lay term results in many irrelevant documents. Thus, the ATM is designed to increase recall at the cost of precision (refer to PubMed Help to understand how ATM works).

Results

First, we performed some basic statistical analysis on query and session data. The number of queries per user ranged from 1 to 8544 (an extreme outlier) with an average of 4.77 queries per user (SD 15.11, median 2). Figure 3 presents the proportion of users that submitted different numbers of queries and the proportion of queries submitted by the corresponding users. Many PubMed users submitted one query. About two-fifths (43%) of users submitted one query that represented around 9% of the total queries. The rest of the users (57%) performed multiple queries and those queries represented about 90% of the total queries. More than half of PubMed users performed one or at most two queries for their information needs. There was a gradual decrement in the proportion of users as the number of queries increased.

PubMed users may perform multiple IR sessions to fulfill their various information needs. In order to identify the purpose of each IR session, we categorized the queries in the log dataset as shown in Figure 2. Figure 4 presents the percentages of different query types. A total of 2,012,466 (69%) queries were identified as informational, 753,827 (26%) queries navigational, and 148,510 (5%) queries mixed. A total of 742,602 user sessions were identified from the informational queries. Because we compared experienced and nonexperienced search sessions, we further identified experienced and nonexperienced search sessions based on their system function usage from the user sessions (that are identified from the informational queries only, see Figure 4).

About 94% (=700,547/742,602) of the sessions were performed by nonexperienced-users and 6% (=42,055/742,602) of the sessions were performed by experienced users (see Figure 4). Some of the users (about 1.12%) performed both experienced and nonexperienced search sessions meaning that such sessions contain both experienced and nonexperienced queries. Since these users knew how to perform searches using advanced system functions, we considered them as experienced users. There are two possible explanations as to why they performed nonexperienced queries. First, they needed to express new concepts but there were no MeSH terms for the concepts. Thus, although they knew of advanced search functions such as MeSH terms, they could not avoid using natural language to describe concepts. Second, as Vibert et al (2009) [6] found, many PubMed users with search skills do not use search functions.

Figure 5 shows the histogram of the proportion of the experienced and nonexperienced users for the various session lengths (the number of queries in a session). Technically, the users in the figure indicate sessions. Because a user may have multiple sessions, a set of sessions that is performed by the same user cannot be matched with a specific (integer number of) session length, meaning that each session is independently treated in the analysis. For both of the groups, the proportion of users significantly decreased as the number of sessions increased. For experienced users, the session length ranged from 1 to 308 (an extreme outlier) with an average of 2.85 queries per session (SD 4.24, median 1). For nonexperienced users, session length ranged from 1 to 8522 (an extreme outlier) with

XSL•FO

RenderX

an average of 2.7 (SD 11.61, median 2). As the standard deviation values indicate, session length variation of nonexperienced sessions was higher than that of experienced sessions. Figure 5 clearly shows the difference between experienced users and nonexperienced users in terms of session length. While for users whose session length was 1 (ie, an ideal IR), the percentage of experienced users was higher than that of nonexperienced users (25,365/42,055, 60.31% vs 331,337/700,547, 47.30%), for users whose session length was 2 or 3, the percentage of the experienced group was lower than that of the nonexperienced group. This session length difference indicates that experienced users completed their searches earlier than nonexperienced users.

In addition, we measured user decrease rates of the experienced and nonexperienced users from the session length of 1 to 2, 3, 4, and 5. Because the ideal session length is 1 (meaning that a user fulfills his or her information need with only one query), the baseline session length should be 1 (the ideal session). Decrease rates from the baseline indicate the success of the IR session (at retrieving relevant documents). Figure 6 compares decrease rates from the baseline of the two user groups. The decrease rate of the experienced users at the session length of 2 was significantly higher than that of the nonexperienced group (the formula to calculate the rate of the experienced users at the session length of 2 is: $1 - \#$ of experienced sessions at the session length of 2/# of experienced sessions at the session length of 1, or $1 - 3969/25,365 = 84.30\%$). The decrease rates of the two groups indicated that most experienced PubMed user sessions were closed within only one query (note the median of the session lengths was 1) (in other words, the initial or first query satisfied the users' information needs) and nonexperienced user sessions (median of 2) were longer than those of the experienced group.

Figure 3. Percentage of users and queries per number of queries.
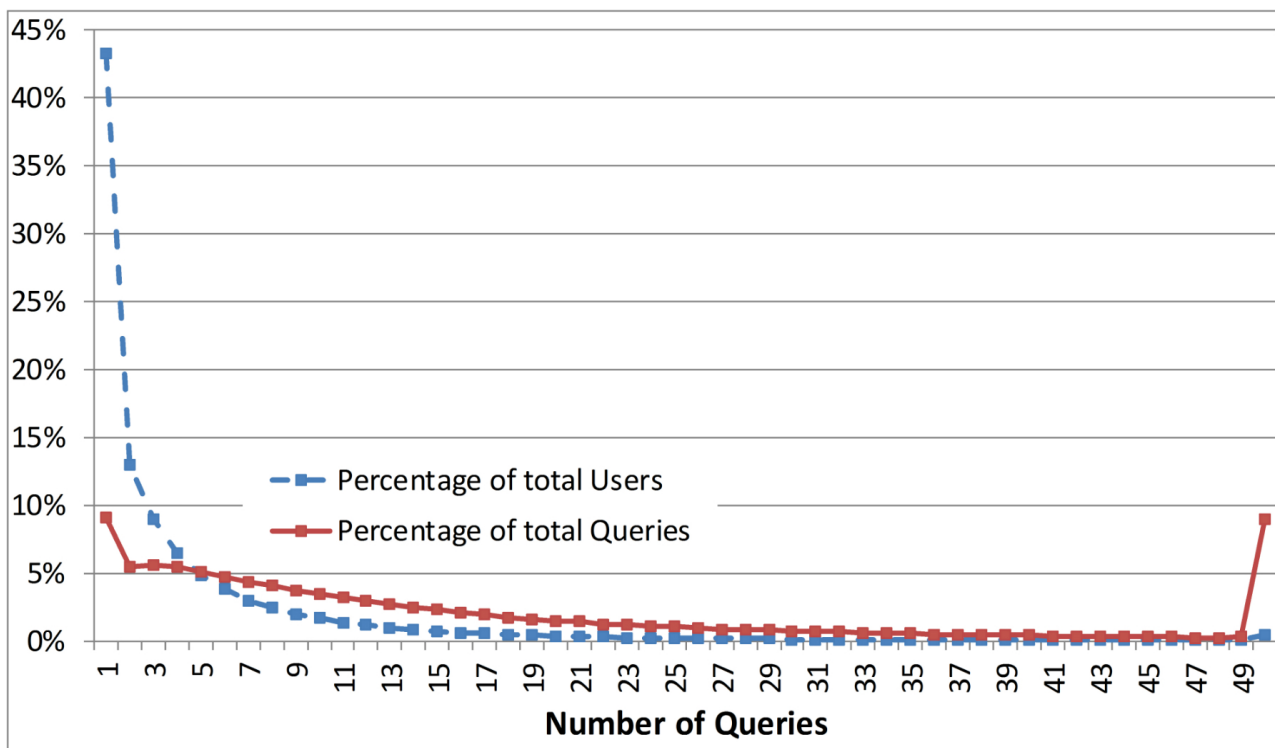
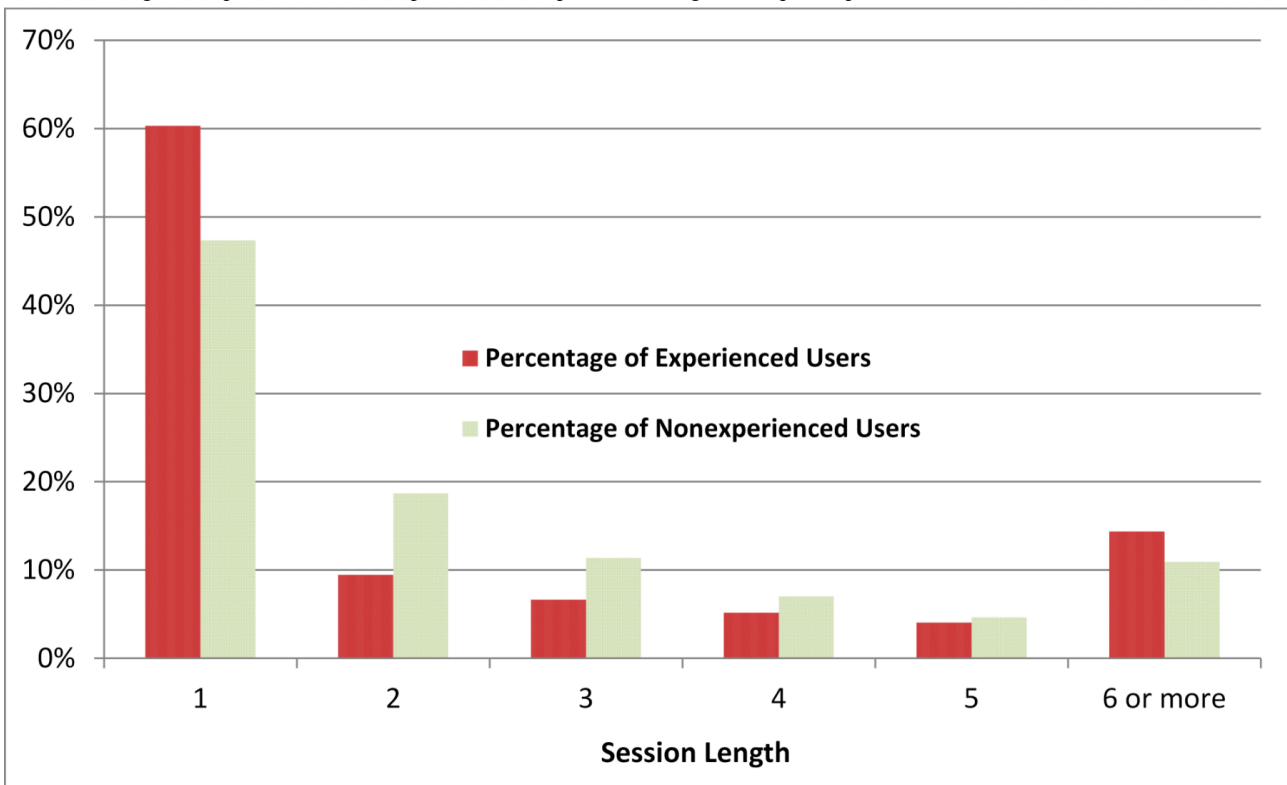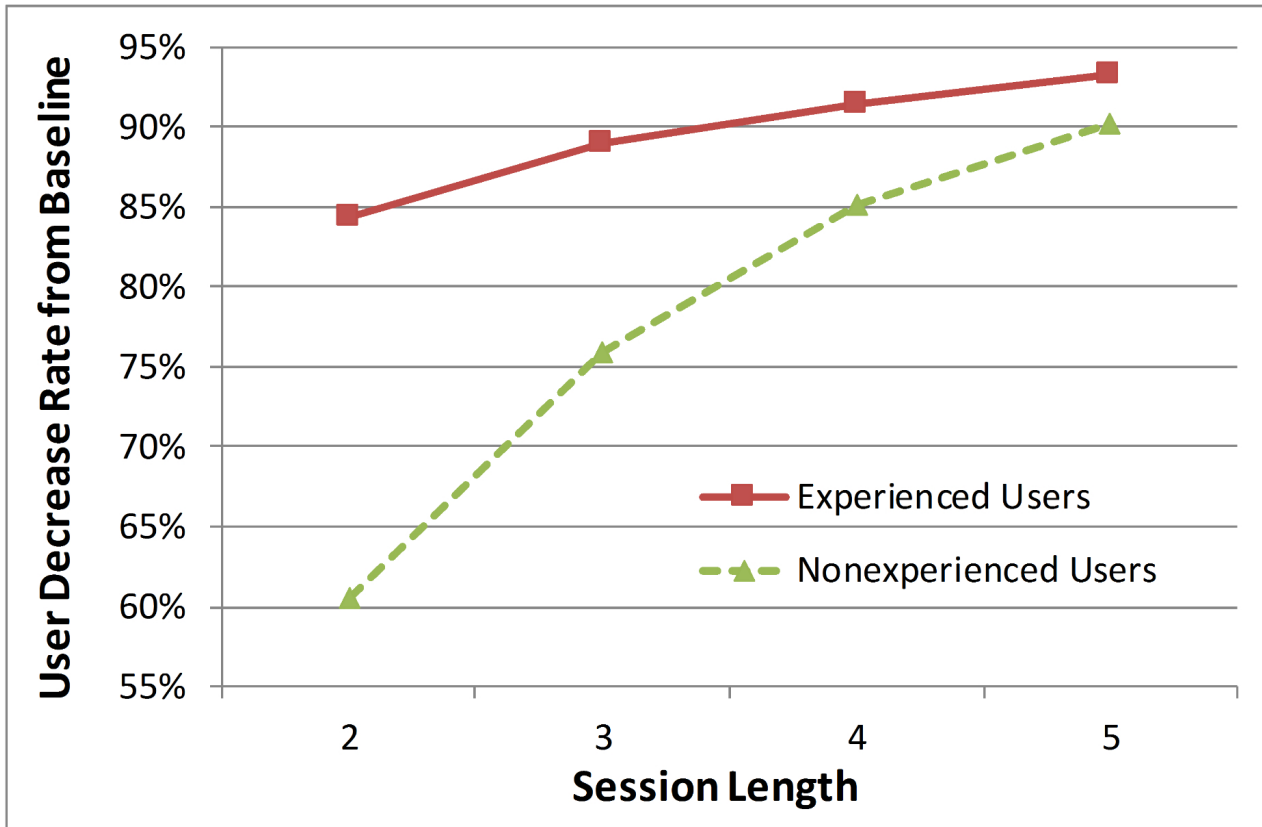**Figure 4.** Query types and session types.
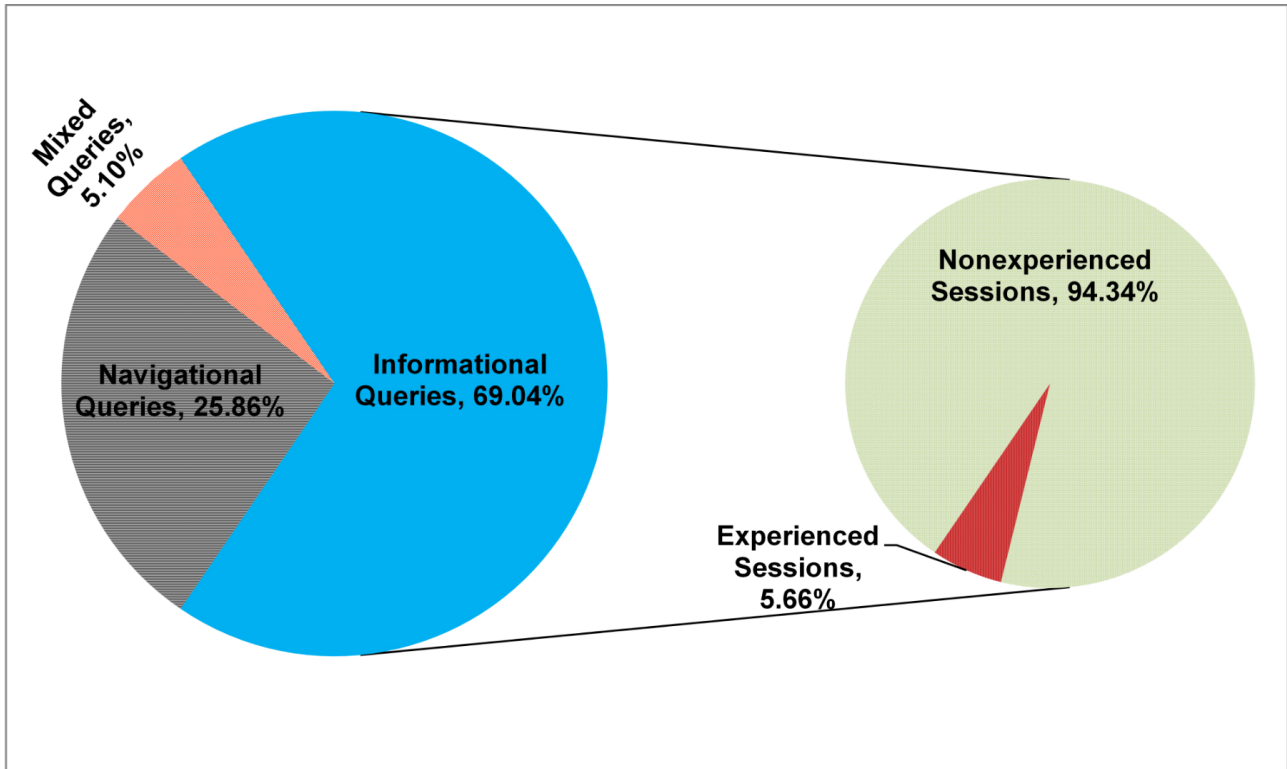
Figure 5. Percentages of experienced and nonexperienced users per session length (# of queries per session).

Figure 6. Decrease rates of experienced and nonexperienced users by session length (# of queries per session).

Discussion

Principal Findings

In bibliographic searches like PubMed searches, procedural knowledge is an important factor to improve the overall performance of information retrieval. Procedural knowledge includes experience using online search systems and their search functions. Earlier studies demonstrated that PubMed users perform searches with higher recall and precision if PubMed search functions are used [25,26,85-89]. These studies used at most tens of human subjects for their experiments. In this study, to check the effect of IR functions on PubMed searches, we performed an analysis on a very large scale. The full-day PubMed log data we used contained nearly 3 million user queries issued by more than 0.6 million users. To our knowledge, this study is the first in the field of biomedical and health informatics to use log data containing nearly 3 million queries to compare search performance and behavior of experienced and nonexperienced users. For the analysis, we first categorized queries into informational or navigational based on their underlying intentions, and then identified 0.7 million informational query sessions from more than 2 million informational queries. An informational query session consisted of one or many informational queries in a row within a 20-minute session window. Sessions were further categorized into experienced and nonexperienced user sessions. To test our hypotheses, we compared experienced and nonexperienced users, and found that experienced PubMed users quickly retrieved relevant documents and nonexperienced PubMed users had longer search sessions than experienced users.

Limitations

There are some limitations of this study. First, the PubMed query log data used in this study could have been biased in terms of IR function usage because the data contained search queries for one day only. Second, we used a predetermined time cutoff (20 minutes) for determining search sessions since the log data did not contain any session-related information. It is possible for a PubMed user to perform more than one session in 20 minutes. However, according to recent studies [6,84], most users complete their search session within 20 minutes. At the same time, it is not common that PubMed users spend more than 20 minutes on a search session; more than 65% of PubMed users perform one to three queries per session (see Figure 3). Third, the classification of users based on the use of search tags is not always correct. In other words, the user classification names (ie, experienced and nonexperienced user groups) do not always necessarily indicate that, for example, all the users in the nonexperienced user group are PubMed novice users. At the same time, we believe the group included some experienced users. There are two reasons why experienced users sometimes do not use search functions: first, in order to find "recently published" articles one must use natural language (nonMeSH terms) because those articles are not indexed yet (indexing lag); second, using MeSH terms requires one to search the MeSH database first before conducting PubMed searches (this is an additional step).

Fourth, we assumed if a session was closed within a few queries, the session was successful (meaning that their information needs were fulfilled) even if a session close does not always mean successful IR. This assumption is based on the fact that nearly

77% of users had only 1 to 3 queries in a session. We believe that most searches are successful. If most searches were unsuccessful, one would expect that most users would not use PubMed again. However, according to the NLM, the number of PubMed users has been increasing. In fact, there is no way to know if a session has been successful using the log data; using web log information is the only solution to this problem but this information is not available. We believe that some sessions that are closed within a few queries are unsuccessful. However, the gaps between the decrease rates of the experienced and nonexperienced users (especially at the session length of 2, see Figures 5 & 6) clearly indicate that most sessions that are closed within a few queries are successful. In fact, these limitations are related to the use of log data, rather than direct data from human subjects, for the analysis. In other words, the limitations are simply drawbacks of using log data that we cannot readily overcome.

Current Applicability of the Log Data Analysis to PubMed

It is unknown when the PubMed query data were collected, for confidentiality reasons. However, they are at least 9 years old. One might argue that this study based on old log data is still currently applicable, because the NLM has added many features to improve the performance and user interface of PubMed. Some examples are related citations, automatic term mapping, and PubMed Clinical Queries. PubMed is significantly different from how it was 9 years ago, in terms of the user interface and internal processes for better information retrieval. However, it is imperative to ascertain whether the new features and user interface retrieve documents that are more relevant or lead to better PubMed searches. Studies have found that most PubMed users *still* have difficulty finding relevant documents for patient care in PubMed and do not want to use PubMed for their information needs (instead they want to use UpToDate and/or Google).

There are many recent studies (published in 2010 or later) that found that physicians prefer UpToDate and/or Google to PubMed, and that UpToDate and/or Google provide more answers to clinical questions. Thiele and colleagues (2010) [90] evaluated four search tools (Google, Ovid, PubMed, and UpToDate) widely used to answer clinical questions. They found that Google was the most frequently used search engine for patient care, and Google and UpToDate were faster and brought more clinical answers than PubMed and Ovid. Shariff and colleagues (2013) [91] compared the performance of searches in PubMed and Google Scholar by evaluating the recall and precision of the searches (the first 40 search result records were analyzed) to determine how well search engines answered nephrological questions. The recall of Google Scholar was two times higher than that of PubMed (indicating documents twice as relevant) while the precision of Google Scholar was slightly higher than that of PubMed (indicating less irrelevant documents in the search result). Another advantage of Google Scholar was that it provided nearly three times more links to full-text documents than PubMed. Duran-Nelson and colleagues (2013) [92] carried out a survey to uncover how internal medicine residents use resources (such as UpToDate, Google/Google Scholar, and PubMed) for point-of-care (POC) clinical decision

making. The top two resources the residents used daily at the POC were UpToDate and Google. Of interest, although the residents thought both UpToDate and PubMed provided trustworthy information for patient care, only 20 residents used PubMed daily while nearly 140 residents used UpToDate daily. In addition, the biggest barrier to using PubMed was speed (it took more time to find clinical answers with PubMed). Cook and colleagues (2013) [93] performed a study similar to Duran-Nelson's (Duran-Nelson et al, 2013) [92]. This focus group study (based on a brief survey) showed that physicians used UpToDate two times as much as PubMed, and physicians regarded PubMed as less useful in POC learning due to the time required to find relevant information through PubMed searches. Sayyah Ensan and colleagues (2011) [94] compared PubMed Clinical Queries and UpToDate to determine their ability to answer clinical questions and the time required to find answers. Their findings were that (a) physicians obtain more answers using UpToDate (76%) than PubMed Clinical Queries (43%), and (b) the median times spent retrieving answers using UpToDate and PubMed Clinical Queries were 17 minutes and 29 minutes, respectively. Nourbakhsh and colleagues (2012) [95] evaluated PubMed and Google Scholar with four clinical questions. The first 20 citations/results were analyzed and classified into three relevance groups (clearly relevant, possibly relevant, and not relevant). They found Google Scholar retrieved more relevant documents than PubMed (80% vs 67.6%). Thiele and colleagues (2010) [96] conducted a survey of medical students, residents, and attending physicians on computer use and four search engines widely used to answer clinical questions (Google, Ovid, PubMed, and UpToDate), and compared the search engines in terms of accuracy, speed, and user confidence. Results showed that 33% and 32% of physicians used UpToDate and Google, respectively, for answering their clinical questions, while only 13% of physicians used PubMed. The authors found that Google and UpToDate answered more clinical questions correctly and more quickly than PubMed.

In sum, the findings of these recent studies indicate that the information retrieval features of PubMed are inferior to other electronic resources or search engines such as UpToDate and Google. In other words, most PubMed users still have considerable difficulty obtaining relevant documents/information despite its many new features. As a result, physicians spend more time finding relevant information with PubMed. This problem is critical for PubMed because recent studies still show that the main barrier to POC learning is lack of time [90] [91] [92] [93] [97] [98]. We believe, based on these recent studies that virtually nothing has changed in terms of information-seeking behavior and PubMed from the user's perspective.

Conclusions

The PubMed log analysis indicated that experienced PubMed users quickly retrieved relevant documents in terms of session length and nonexperienced PubMed users had longer search sessions than experienced users. We believe there are a few potential solutions to this problem. First, the NLM could design and provide a novel PubMed user interface for nonexperienced users so that they can readily utilize advanced search functions without special training in PubMed. Second, because it is

imperative for health professionals (especially physicians) to learn the system functions and MeSH vocabulary for better PubMed searches, the NLM could award grant funding only to institutes that regularly train health professionals in PubMed search skills. Third, the NLM could develop a sophisticated relevance-sorting algorithm similar to Google's, so that PubMed users can quickly find relevant documents. Currently, PubMed provides a relevance sorting option. However, it is not the default sorting option as of 17 June 2015 and we believe there should be a significant improvement to the sorting algorithm. This PubMed search problem is not just an information retrieval issue but also a health care practice matter, because health professionals, especially physicians, could significantly improve the quality of patient care and effectively educate chronic patients using clinical and medical information and knowledge obtained from PubMed searches.

Acknowledgments

Conflicts of Interest

None declared.

References

1. Downing R, Moore J, Brown S. The effects and interaction of spatial visualization and domain expertise on information seeking. Comput Human Behav 2005;21(2):195-209.

2. Rouet J. The Skills of Document Use: From Text Comprehension to Web-Based Learning. Mahwah, NJ: Lawrence Erlbaum Associates, Inc; 2006.

3. Tenopir C, King D, Boyce P, Grayson M, Zhang Y, Ebuen M. Patterns of journal use by scientists through three evolutionary phases. D-Lib Magazine 2003 May;9(5) [FREE Full text] [doi: 10.1045/may2003-contents]

4. De Groote SL, Dorsch JL. Measuring use patterns of online journals and databases. J Med Libr Assoc 2003 Apr;91(2):231-240 [FREE Full text] [doi: 10.1043/0025-7338(2003)091<0231:MUPOOJ>2.0.CO;2] [Medline: 12883574]

5. Vibert N, Rouet J, Ros C, Ramond M, Deshoullieres B. The use of online electronic information resources in scientific research: The case of neuroscience. Libr Inf Sci Res 2007 Dec;29(4):508-532. [doi: 10.1016/j.lisr.2007.05.001]

6. Vibert N, Ros C, Bigot L, Ramond M, Gatefin J, Rouet J. Effects of domain knowledge on reference search with the PubMed database: An experimental study. J Am Soc Inf Sci 2009 Jul;60(7):1423-1447. [doi: 10.1002/asi.21078]

7. Mattox DE. Welcome to ARCHIVES CME. Arch Otolaryngol Head Neck Surg 2000 Jul;126(7):914. [Medline: 10889012]

8. Muin M, Fontelo P, Liu F, Ackerman M. SLIM: an alternative Web interface for MEDLINE/PubMed searches - a preliminary study. BMC Med Inform Decis Mak 2005;5:37 [FREE Full text] [doi: 10.1186/1472-6947-5-37] [Medline: 16321145]

9. Bar-Ilan J, Fink N. Preference for electronic format of scientific journals—A case study of the Science Library users at the Hebrew University. Libr Inf Sci Res 2005 Jun;27(3):363-376. [doi: 10.1016/j.lisr.2005.04.011]

10. Sood A, Ghosh A. Literature search using PubMed: an essential tool for practicing evidence- based medicine. J Assoc Physicians India 2006 Apr;54:303-308. [Medline: 16944614]

11. Bahaadinbeigy K, Yogesan K, Wootton R. MEDLINE versus EMBASE and CINAHL for telemedicine searches. Telemed J E Health 2010 Oct;16(8):916-919. [doi: 10.1089/tmj.2010.0046] [Medline: 20925566]

12. Yoo I, Marinov M. Recent research for MEDLINE/PubMed: short review. USA: ACM; 2010 Presented at: Proceedings of the ACM fourth international workshop on Datatext mining in biomedical informatics; October 26 - 30, 2010; Toronto, ON, Canada p. 69-70. [doi: 10.1145/1871871.1871887]

13. Haux R, Grothe W, Runkel M, Schackert HK, Windeler HJ, Winter A, et al. Knowledge retrieval as one type of knowledge-based decision support in medicine: results of an evaluation study. Int J Biomed Comput 1996 Apr;41(2):69-85. [Medline: 8803668]

14. Nankivell C, Wallis P, Mynott G. Networked information and clinical decision making: the experience of Birmingham Heartlands and Solihull National Health Service Trust (Teaching). Med Educ 2001 Feb;35(2):167-172. [Medline: 11169091]

15. Baker NC, Hemminger BM. Mining connections between chemicals, proteins, and diseases extracted from Medline annotations. J Biomed Inform 2010 Aug;43(4):510-519 [FREE Full text] [doi: 10.1016/j.jbi.2010.03.008] [Medline: 20348023]

16. Herskovic JR, Tanaka LY, Hersh W, Bernstam EV. A day in the life of PubMed: analysis of a typical day's query log. J Am Med Inform Assoc 2007;14(2):212-220 [FREE Full text] [doi: 10.1197/jamia.M2191] [Medline: 17213501]

17. Islamaj DR, Murray G, Névéol A, Lu Z. Understanding PubMed user search behavior through log analysis. Database (Oxford) 2009;2009:bap018 [FREE Full text] [doi: 10.1093/database/bap018] [Medline: 20157491]

18. Lacroix E, Mehnert R. The US National Library of Medicine in the 21st century: expanding collections, nontraditional formats, new audiences. Health Info Libr J 2002 Sep;19(3):126-132. [Medline: 12390234]

19. Hersh W, Hickam D. Use of a multi-application computer workstation in a clinical setting. Bull Med Libr Assoc 1994 Oct;82(4):382-389 [FREE Full text] [Medline: 7841907]

XSL•FO

RenderX

20. McKibbon KA, Haynes RB, Dilks CJ, Ramsden MF, Ryan NC, Baker L, et al. How good are clinical MEDLINE searches? A comparative study of clinical end-user and librarian searches. Comput Biomed Res 1990 Dec;23(6):583-593. [Medline: 2276266]

21. Pao ML, Grefsheim SF, Barclay ML, Woolliscroft JO, McQuillan M, Shipman BL. Factors affecting students' use of MEDLINE. Comput Biomed Res 1993 Dec;26(6):541-555. [Medline: 8112055]

22. Gallagher PE, Allen TY, Wyer PC. How to find evidence when you need it, part 2: a clinician's guide to MEDLINE: the basics. Ann Emerg Med 2002 Apr;39(4):436-440. [Medline: 11919532]

23. Chapman D. Advanced search features of PubMed. J Can Acad Child Adolesc Psychiatry 2009 Feb;18(1):58-59 [FREE Full text] [Medline: 19270851]

24. Hersh WR, Crabtree MK, Hickam DH, Sacherek L, Friedman CP, Tidmarsh P, et al. Factors associated with success in searching MEDLINE and applying evidence to answer clinical questions. J Am Med Inform Assoc 2002;9(3):283-293 [FREE Full text] [Medline: 11971889]

25. Richter RR, Austin TM. Using MeSH (medical subject headings) to enhance PubMed search strategies for evidence-based practice in physical therapy. Phys Ther 2012 Jan;92(1):124-132 [FREE Full text] [doi: 10.2522/ptj.20100178] [Medline: 21979271]

26. Darmoni SJ, Soualmia LF, Letord C, Jaulent M, Griffon N, Thirion B, et al. Improving information retrieval using Medical Subject Headings Concepts: a test case on rare and chronic diseases. J Med Libr Assoc 2012 Jul;100(3):176-183 [FREE Full text] [doi: 10.3163/1536-5050.100.3.007] [Medline: 22879806]

27. Hubert G, Cabanac G, Sallaberry C, Palacio D. Query operators shown beneficial for improving search results. In: Gradmann S, editor. Research and Advanced Technology for Digital Libraries: International Conference on Theory and Practice of Digital Libraries. Berlin: Springer Berlin Heidelberg; 2011:118-129.

28. Xie I, Joo S. Factors affecting the selection of search tactics: Tasks, knowledge, process, and systems. Inf Process Manag 2012 Mar;48(2):254-270. [doi: 10.1016/j.ipm.2011.08.009]

29. Hölscher C, Strube G. Web search behavior of Internet experts and newbies. Comput Netw 2000 Jun;33(1-6):337-346. [doi: 10.1016/S1389-1286(00)00031-1]

30. Oldroyd B. Study of strategies used in online searching 5: differences between the experienced and the inexperienced searcher. Online Rev 1984 Mar;8(3):233-244. [doi: 10.1108/eb024156]

31. Harter S. Online searching styles: An exploratory study. Coll Res Libr 1984 Jul 01;45(4):249-258. [doi: 10.5860/crl_45_04_249]

32. Penniman D. Modeling and Evaluation of On-Line User Behavior. Final Report to the National Library of Medicine. Dublin, OH: OCCL; Sep 1981.

33. Hsieh-Yee I. Effects of search experience and subject knowledge on the search tactics of novice and experienced searchers. J Am Soc Inf Sci 1993 Apr;44(3):161-174 [FREE Full text] [doi: 10.1002/(SICI)1097-4571]

34. Drabenstott K. Do nondomain experts enlist the strategies of domain experts? J Am Soc Inf Sci Technol 2003 Jul;54(9):836-854. [doi: 10.1002/asi.10281]

35. Tabatabai D, Shore B. How experts and novices search the Web. Libr Inf Sci Res 2005 Mar;27(2):222-248. [doi: 10.1016/j.lisr.2005.01.005]

36. Gardois P, Calabrese R, Colombi N, Deplano A, Lingua C, Longo F, et al. Effectiveness of bibliographic searches performed by paediatric residents and interns assisted by librarians. A randomised controlled trial. Health Info Libr J 2011 Dec;28(4):273-284. [doi: 10.1111/j.1471-1842.2011.00957.x] [Medline: 22051126]

37. Fields B, Keith S, Blandford A. Designing for expert information finding strategies. In: Fincher S, editor. People and Computers XVIII - Design for Life: Proceedings of HCI 2004 (Pt. 18). London, UK: Springer London; 2005:89-102.

38. US National Library of Medicine. National Institutes of Health. 2005 May 17. MEDLINE/PubMed Data Element (Field) Descriptions URL: http://www.nlm.nih.gov/bsd/mms/medlineelements.html [accessed 2014-07-30] [WebCite Cache ID 6RRV0Kbq8]

39. Marchionini G. Information Seeking in Electronic Environments. Cambridge: Cambridge University Press; 1995.

40. Chen C, Rada R. Interacting with hypertext: A meta-analysis of experimental studies. Hum Comput Interact 1996 Jun 1;11(2):125-156. [doi: 10.1207/s15327051hci1102_2]

41. Palmquist R, Kim K. Cognitive style and on-line database search experience as predictors of Web search performance. J Am Soc Inf Sci 2000;51(6):558-566. [doi: 10.1002/(SICI)1097-4571(2000)51:6<558::AID-ASI7>3.0.CO;2-9]

42. Marchionini G, Dwiggins S, Katz A, Lin X. Information seeking in full-text end-user-oriented search systems: The roles of domain and search expertise. Libr Inf Sci Res 1993;15(1):35-69.

43. Hembrooke H, Granka L, Gay G, Liddy E. The effects of expertise and feedback on search term selection and subsequent learning. J Am Soc Inf Sci 2005 Jun;56(8):861-871. [doi: 10.1002/asi.20180]

44. Jenkins C, Corritore C, Wiedenbeck S. Patterns of information seeking on the Web: A qualitative study of domain expertise and Web expertise. Inf Technol Soc 2003;1(3):64-89.

45. Ju B. Does domain knowledge matter: Mapping users' expertise to their information interactions. J Am Soc Inf Sci 2007 Nov;58(13):2007-2020. [doi: 10.1002/asi.20680]

XSL•FO

RenderX

46.   Patel S, Drury C, Shalin V. Effectiveness of expert semantic knowledge as a navigational aid within hypertext. Behav Inf Tech 1998 Jan;17(6):313-324. [doi: 10.1080/014492998119274]

47.   Wildemuth B, de Bliek R, Friedman C, File D. Medical students' personal knowledge, searching proficiency, and database use in problem solving. J Am Soc Inf Sci 1999;46(8):590-607. [doi: 10.1002/(SICI)1097-4571(199509)46:8<590::AID-ASI7>3.0.CO;2-#]

48.   Hersh WR, Hickam DH. How well do physicians use electronic information retrieval systems? A framework for investigation and systematic review. JAMA 1998 Oct 21;280(15):1347-1352. [Medline: 9794316]

49.   Saracevic T, Kantor P. A study of information seeking and retrieving. II. Users, questions, and effectiveness. J Am Soc Inf Sci 1988 May;39(3):177-196. [doi: 10.1002/(SICI)1097-4571(198805)39:3<177::AID-ASI3>3.0.CO;2-F]

50.   Allen B. Topic knowledge and online catalog search formulation. Libr Q 1991 Apr;61(2):188-213.

51.   Aula A, Nordhausen K. Modeling successful performance in Web searching. J Am Soc Inf Sci 2006 Oct 1;57(12):1678-1693. [doi: 10.1002/asi.20340]

52.   Egan D. Individual differences in human-computer interaction. In: Handbook of Human-Computer Interaction. Amsterdam, Netherlands: Elsevier; 1988:543-568.

53.   Cullen RJ. In search of evidence: family practitioners' use of the Internet for clinical information. J Med Libr Assoc 2002 Oct;90(4):370-379 [FREE Full text] [Medline: 12398243]

54.   Davies KS. Physicians and their use of information: a survey comparison between the United States, Canada, and the United Kingdom. J Med Libr Assoc 2011 Jan;99(1):88-91 [FREE Full text] [doi: 10.3163/1536-5050.99.1.015] [Medline: 21243061]

55.   Markey K. Twenty-five years of end-user searching, Part 1: Research findings. J Am Soc Inf Sci 2007 Jun;58(8):1071-1081. [doi: 10.1002/asi.20462]

56.   Macedo-Rouet M, Rouet J, Ros C, Vibert N. How do scientists select articles in the PubMed database? An empirical study of criteria and strategies. Eur Rev Appl Psychol 2012 Apr;62(2):63-72. [doi: 10.1016/j.erap.2012.01.003]

57.   Lasserre K. Expert searching in health librarianship: a literature review to identify international issues and Australian concerns. Health Info Libr J 2012 Mar;29(1):3-15. [doi: 10.1111/j.1471-1842.2011.00974.x] [Medline: 22335285]

58.   McKibbon KA, Haynes RB, Johnston ME, Walker CJ. A study to enhance clinical end-user MEDLINE search skills: design and baseline findings. Proc Annu Symp Comput Appl Med Care 1991:73-77 [FREE Full text] [Medline: 1807700]

59.   Silverstein C, Marais H, Henzinger M, Moricz M. Analysis of a very large web search engine query log. SIGIR Forum 1999 Sep 01;33(1):6-12. [doi: 10.1145/331403.331405]

60.   Jansen B, Spink A, Saracevic T. Real life, real users, and real needs: a study and analysis of user queries on the web. Information Processing & Management 2000 Mar;36(2):207-227. [doi: 10.1016/S0306-4573(99)00056-4]

61.   Facca F, Lanzi P. Mining interesting knowledge from weblogs: a survey. Data Knowl Eng 2005 Jun;53(3):225-241. [doi: 10.1016/j.datak.2004.08.001]

62.   Murray G, Teevan J. Query log analysis: social and technological challenges. SIGIR Forum 2007 Dec 01;41(2):112-120. [doi: 10.1145/1328964.1328985]

63.   Doğan R, Murray G, Névéol A, Lu Z. Characterizing user search behavior in PubMed. : American Medical Informatics Association; 2010 Presented at: AMIA 2010 Annual Symposium; 2015 Nov 16; Washington, DC p. 1094.

64.   Lu Z, Wilbur W. Improving accuracy for identifying related PubMed queries by an integrated approach. J Biomed Inform 2009 Oct;42(5):831-838 [FREE Full text] [doi: 10.1016/j.jbi.2008.12.006] [Medline: 19162232]

65.   Lu Z, Kim W, Wilbur WJ. Evaluation of query expansion using MeSH in PubMed. Inf Retr Boston 2009;12(1):69-80 [FREE Full text] [doi: 10.1007/s10791-008-9074-8] [Medline: 19774223]

66.   Névéol A, Islamaj DR, Lu Z. Semi-automatic semantic annotation of PubMed queries: a study on quality, efficiency, satisfaction. J Biomed Inform 2011 Apr;44(2):310-318 [FREE Full text] [doi: 10.1016/j.jbi.2010.11.001] [Medline: 21094696]

67.   Lu Z, Wilbur WJ, McEntyre JR, Iskhakov A, Szilagyi L. Finding query suggestions for PubMed. 2009 Nov 14 Presented at: AMIA Annu Symp Proc; 2009; San Francisco, USA p. 396-400.

68.   Lu Z, Xie N, Wilbur WJ. Identifying related journals through log analysis. Bioinformatics 2009 Nov 15;25(22):3038-3039 [FREE Full text] [doi: 10.1093/bioinformatics/btp529] [Medline: 19734155]

69.   Mosa ASM, Yoo I. A study on PubMed search tag usage pattern: association rule mining of a full-day PubMed query log. BMC Med Inform Decis Mak 2013;13:8 [FREE Full text] [doi: 10.1186/1472-6947-13-8] [Medline: 23302604]

70.   Broder A. A taxonomy of web search. SIGIR Forum 2002 Sep 01;36(2):3-10. [doi: 10.1145/792550.792552]

71.   PubMed. U.S. National Library of Medicine, National Institute of Health. PubMed Help URL: http://www.ncbi.nlm.nih.gov/books/NBK3827/pdf/Bookshelf_NBK3827.pdf [accessed 2015-06-18] [WebCite Cache ID 6ZOCemft9]

72.   Spink A, Bateman J, Jansen B. Searching heterogeneous collections on the Web: behaviour of Excite users. Inf Res 1998 Oct;4(2):12-27.

73.   Swanson D. Information retrieval as a trial-and-error process. Libr Q 1977;47(2):128-148.

74.   Downey D, Dumais S, Horvitz E. Models of searching and browsing: languages, studies, and applications. In: Proceedings of the 20th international joint conference on Artifical intelligence. San Francisco, CA: Morgan Kaufmann Publishers Inc; 2007 Presented at: IJCAI; 2007 Jan 6-12; Hyderabad, India p. 2740-2747.

75. Huang C, Chien L, Oyang Y. Relevant term suggestion in interactive web search based on contextual information in query session logs. J Am Soc Inf Sci 2003 May;54(7):638-649. [doi: 10.1002/asi.10256]

76. Buzikashvili N, Jansen B. Limits of the web log analysis artifacts. 2006 Presented at: Workshop on Logging Traces of Web Activity: The Mechanics of Data Collection; 2006; Edinburgh, Scotland.

77. He D, Göker A. Detecting session boundaries from web user logs. 2000 Apr 05 Presented at: Proceedings of the 22nd Annual Colloquium on Information Retrieval Research; 2000; Cambridge p. 57-66.

78. Radlinski F, Joachims T. Query chains: learning to rank from implicit feedback. New York, NY: ACM Press; 2005 Presented at: Proceedings of the eleventh ACM SIGKDD international conference on knowledge discovery in data mining; 2005 Aug 21-24; Chicago, IL, USA p. 239-248. [doi: 10.1145/1081870.1081899]

79. Shi X, Yang C. Mining related queries from search engine query logs. New York, NY, USA: ACM Press; 2006 Presented at: Proceedings of the 15th international conference on World Wide Web - WWW; May 23-26, 2006; Edinburgh, Scotland p. 943-944. [doi: 10.1145/1135777.1135956]

80. Gayo-Avello D. A survey on session detection methods in query logs and a proposal for future evaluation. Inf Sci 2009 May 30;179(12):1822-1843. [doi: 10.1016/j.ins.2009.01.026]

81. Catledge L, Pitkow J. Characterizing browsing strategies in the World-Wide Web. Proc Int World Wide Web Conf 1995 Apr;27(6):1065-1073. [doi: 10.1016/0169-7552(95)00043-7]

82. Jansen B, Spink A. An analysis of Web information seeking and use: Documents retrieved versus documents viewed. 2003 Presented at: Proceedings of the 4th International Conference on Internet Computing; 2003 Jun 23-26; Las Vegas, Nevada, USA p. 65-69.

83. Jansen B, Spink A, Blakely C, Koshman S. Defining a session on Web search engines. J Am Soc Inf Sci 2007 Apr;58(6):862-871. [doi: 10.1002/asi.20564]

84. Ahmadi S, Faghankhani M, Javanbakht A, Akbarshahi M, Mirghorbani M, Safarnejad B, et al. A comparison of answer retrieval through four evidence-based textbooks (ACP PIER, Essential Evidence Plus, First Consult, and UpToDate): a randomized controlled trial. Med Teach 2011;33(9):724-730. [doi: 10.3109/0142159X.2010.531155] [Medline: 21854150]

85. Haynes RB, McKibbon KA, Walker CJ, Ryan N, Fitzgerald D, Ramsden MF. Online access to MEDLINE in clinical settings. A study of use and usefulness. Ann Intern Med 1990 Jan 1;112(1):78-84. [Medline: 2403476]

86. Bernstam E, Kamvar S, Meric F, Dugan J, Chizek S, Stave C. Oncology patient interface to MEDLINE. 2001 Presented at: Proceedings of the American Society of Clinical Oncology; 2001; USA p. 244a.

87. Agoritsas T, Merglen A, Courvoisier DS, Combescure C, Garin N, Perrier A, et al. Sensitivity and predictive value of 15 PubMed search strategies to answer clinical questions rated against full systematic reviews. J Med Internet Res 2012;14(3):e85 [FREE Full text] [doi: 10.2196/jmir.2021] [Medline: 22693047]

88. Shariff SZ, Sontrop JM, Haynes RB, Iansavichus AV, McKibbon KA, Wilczynski NL, et al. Impact of PubMed search filters on the retrieval of evidence by physicians. CMAJ 2012 Feb 21;184(3):E184-E190 [FREE Full text] [doi: 10.1503/cmaj.101661] [Medline: 22249990]

89. Ugolini D, Neri M, Casilli C, Bonassi S. Development of search filters for retrieval of literature on the molecular epidemiology of cancer. Mutat Res 2010 Aug 30;701(2):107-110. [doi: 10.1016/j.mrgentox.2010.05.016] [Medline: 20685344]

90. Thiele RH, Poiro NC, Scalzo DC, Nemergut EC. Speed, accuracy, and confidence in Google, Ovid, PubMed, and UpToDate: results of a randomised trial. Postgrad Med J 2010 Aug;86(1018):459-465. [doi: 10.1136/pgmj.2010.098053] [Medline: 20709767]

91. Shariff SZ, Bejaimal SA, Sontrop JM, Iansavichus AV, Haynes RB, Weir MA, et al. Retrieving clinical evidence: a comparison of PubMed and Google Scholar for quick clinical searches. J Med Internet Res 2013;15(8):e164 [FREE Full text] [doi: 10.2196/jmir.2624] [Medline: 23948488]

92. Duran-Nelson A, Gladding S, Beattie J, Nixon LJ. Should we Google it? Resource use by internal medicine residents for point-of-care clinical decision making. Acad Med 2013 Jun;88(6):788-794. [doi: 10.1097/ACM.0b013e31828ffdb7] [Medline: 23619072]

93. Cook DA, Sorensen KJ, Hersh W, Berger RA, Wilkinson JM. Features of effective medical knowledge resources to support point of care learning: a focus group study. PLoS One 2013;8(11):e80318 [FREE Full text] [doi: 10.1371/journal.pone.0080318] [Medline: 24282535]

94. Sayyah EL, Faghankhani M, Javanbakht A, Ahmadi S, Baradaran HR. To compare PubMed Clinical Queries and UpToDate in teaching information mastery to clinical residents: a crossover randomized controlled trial. PLoS One 2011;6(8):e23487 [FREE Full text] [doi: 10.1371/journal.pone.0023487] [Medline: 21858142]

95. Nourbakhsh E, Nugent R, Wang H, Cevik C, Nugent K. Medical literature searches: a comparison of PubMed and Google Scholar. Health Info Libr J 2012 Sep;29(3):214-222. [doi: 10.1111/j.1471-1842.2012.00992.x] [Medline: 22925384]

96. Thiele RH, Poiro NC, Scalzo DC, Nemergut EC. Speed, accuracy, and confidence in Google, Ovid, PubMed, and UpToDate: results of a randomised trial. Postgrad Med J 2010 Aug;86(1018):459-465. [doi: 10.1136/pgmj.2010.098053] [Medline: 20709767]

97. Cook DA, Sorensen KJ, Hersh W, Berger RA, Wilkinson JM. Features of effective medical knowledge resources to support point of care learning: a focus group study. PLoS One 2013;8(11):e80318 [FREE Full text] [doi: 10.1371/journal.pone.0080318] [Medline: 24282535]

98.   van Dijk N, Hooft L, Wieringa-de WM. What are the barriers to residents' practicing evidence-based medicine? A systematic review. Acad Med 2010 Jul;85(7):1163-1170. [doi: 10.1097/ACM.0b013e3181d4152f] [Medline: 20186032]

Abbreviations

ATM: Automatic Term Mapping
IR: information retrieval
MeSH: Medical Subject Heading
NLM: National Library of Medicine