# Dynamics of relevance judgment during physicians' online information search process for patient treatment

**Soon Kim[1], EunKyung Chung[2*]** and **Nahyun Kwon[3]**
[1]Research Institute for Social Science, Ewha Womans University,
11-1 Seodaemun-Gu Daehyun-Dong, Seoul 120-750, KOREA
[2]Department of Library and Information Science, Ewha Womans University,
11-1 Seodaemun-Gu Daehyun-Dong, Seoul 120-750, KOREA
[3]Department of Library and Information Science, Myongji University ,
34 Geobukgol-ro, Seodaemun-Gu, Seoul 120-728, KOREA
e-mail: soonkim0916@gmail.com; *echung@ewha.ac.kr (corresponding author);
nahyun.kwon@gmail.com

**ABSTRACT**

*Physicians rely on online sources of information more so than ever, to find the information they need when treating patients, but little is known about determining the relevance criteria when searching for online sources. Relevance judgment criteria can change with a user's background knowledge on a search topic. Physicians are a subset of users who have received insufficient attention in this relevance criteria research topic. Due to the dynamic and multidimensional nature of relevance judgments, it is important to understand how relevance judgment criteria change in the course of information search. This study aims to determine the dynamics in relevance judgment criteria in physicians' information search process. We observed ten internal medicine physicians while they searched the online clinical information resource to resolve a clinical question concerning patient treatment. They rated on a checklist to assess relevance judgment criteria at each of the three stages in the search process, namely, problem recognition (before searching), system interaction (searching for online resources), and document interaction (selecting the final documents after searching). We conducted pre-interview and post-interview to probe physicians search behaviours and reasons for the assessments. Topical relevance, accuracy, and credibility were found to be the most important criteria for all three search stages; however, ratings of some relevance judgment criteria, such as title, obtainability, personal preference, and understandability, changed markedly across the search process. The criteria regarding the information content tended to be more important during the problem recognition stage, while the criteria in the information format became more important as the search progressed. Physicians preferred recently published review articles in highly reliable journals with well-written article outlines, which allowed them to have quick and clear understanding of the contents. By identifying specific relevance criteria prominent at each search stage, our study revealed that physicians employ a unique set of criteria when judging the relevance. Given the fact that physicians are relying on online resources to find answers to patient care, this study can help in developing the effective clinical information systems for physicians and integrating into the Electronic Health Record (EHR) systems.*

**Keywords***:* Information search process; Online information search; Information seeking behaviour; Dynamics of relevance judgment; Online clinical information systems.

## INTRODUCTION

Traditionally, physicians have asked colleagues or relied on their own clinical experience or personal knowledge for answers to clinical questions (Clarke et al. 2013); however, physicians are now using more online sources of clinical information (Daei et al. 2020a; Pluye et al. 2005). Much of this online information, for example drug information from pharmaceutical companies, lacks credibility, thus selecting the most relevant and authoritative information can be difficult for physicians (Mikalef, Kourouthanassis, and Pateli 2017). Many clinical questions have gone unresolved because physicians have insufficient time to search for new medical information and obtain objective evidence for optimal treatment of patients, due to their busy schedules (Davies 2011).

Physicians face many questions during patient care, as patients have the right to information about their health status, treatment options and reasonable alternatives, and the likely benefits and risks of proposed treatment and non-treatment. Therefore finding high-quality evidence provides a good opportunity to improve patient care (Daei et al. 2020b). Preventing medical errors resulting from a lack of information and providing high-quality medical services to patients requires an information search system that allows physicians to easily select needed medical information and immediately apply it to treatment. Such a system could greatly reduce information search time and provide prompt assistance in making important clinical decisions (Westbrook, Gosling and Westbrook 2005), thereby can improve the quality of patient treatment (Del Fiol, Workman and Gorman 2014). As clinicians need easy access to the evidence and clinical data every day (Daei et al. 2020a), it is necessary to uncover how physicians perform information searches to resolve daily work problems and how they choose relevant information at different stages of the search process. Previous studies have empirically investigated the dynamic characteristics of relevance judgment during the information search process (Bruce 1994; Spink, Greisdorf and Bateman 1998; Vakkari, Pennanen and Serola 2003). These studies found that relevance judgment criteria can change with a user's background knowledge on a search topic. However, little is known about how physicians determine the relevance judgment criteria during the information search process, except those reported in a few studies on text retrieval for precision medicine literature search (Hollis et al. 2020; Qu, Arguello and Wang 2020; Uprety et al. 2018).

Saracevic (1997) referred to the dynamic and multidimensional characteristics of relevance judgments as "relevance dynamics," and stressed the importance of understanding how relevance criteria change across information search stages. Tang and Solomon (2001) found that the most important selection criteria are topicality which is related to the search questions. However, when selecting the final documents, the novelty of the subject (freshness) is perceived as more important than in the search stage. Vakkari, Pennanen and Serola (2003) classified three search stage - pre-focus, focus formulation and post-focus - and suggested that as users acquire new knowledge, they learn more sophisticated search-term selections and search techniques. According to Vakkari, Pennanen and Serola (2003), a more detailed investigation of relevance dynamics might help in designing a more effective information search system.

Due to the dynamic and multidimensional nature of relevance judgments, it is important to understand how relevance judgment criteria change in the course of information search. This study aims to determine the dynamics in relevance judgment criteria in physicians' information search process, such as problem recognition (before searching), system interaction (searching for online resources), and document interaction (selecting final

documents after searching). Considering the fact that physicians are heavily using online resources to find answers for patient treatment, this study could help medical librarians and system designers to make the clinical information system easier to accommodate for physicians' information demand.

## LITERATURE REVIEW

### Relevance Dynamics in the Information Search Process

Bateman (1998) investigated the relevance dynamics of the information search process using a 6-stage Information Search Process (ISP) model developed by Kuhlthau (1991). Specifically, Bateman asked graduate students to write a search journal using 40 relevance judgment criteria in each information search stage. The results showed that topicality (such as information quality, credibility, completeness) was the most important judgment criterion.

Wang and Soergel (1998) used the think-aloud protocol to examine relevance judgment and derived 11 relevance judgment criteria (suc as topicality, orientation, quality, novelty, recency, authority). In a follow-up study with the same users, Wang and White (1999) divided article citation behaviours into three stages - selecting, reading, and citing relevant articles - to examine how relevance judgment criteria changed across these stages. Topicality, currency, and novelty were the most critical relevance criteria, while "cognitive requisite" was added in the citing stage. When selecting articles, users tended to consider many more articles as relevant than they did during the citing stage, wherein they were more judicious and employed diverse judgment criteria. This finding indicates that, in the early stage of the search process, users perceive more articles to be relevant because they are less skillful in judging relevance.

Taylor et al. (2007) found that specificity and currency were considered important criteria during the stage of identifying and learning about a topic, and coverage and novelty during the stage of writing the paper. Taylor et al. (2007) found the following criteria - understandability, clarity, depth/coverage, accuracy, and specificity - as important across all search stages, but the preferred criteria changed as the search progressed. Based on their statistical analysis, the authors suggested several ways of improving search system by incorporating relevance dynamics.

The numerous qualitative studies examined information search system users have shown that relevance judgment criteria change during the information search process. Most changes are contextual or relate to background knowledge - the essential relevance judgment criteria (e.g., topicality, credibility) do not greatly change. Yet, a limitation of these studies is that they were conducted in the academic setting, sampling college professors and students, which highlights the need for further research on how specific work-task-oriented professional knowledge or experience influences relevance judgments during a work-oriented information search process, that is, search context for patient treatment.

### User-Centered Relevance Judgment Criteria in Medical Fields

Hersh (1994) argued that contextual relevance should be considered alongside topical relevance, reporting that topical relevance judgment criteria were used earlier on in the search, whereas contextual relevance was more influential for selecting relevant articles. Hughes, Wareham and Joshi (2010) had 35 physicians compose an online search journal

and explored their relevance judgments criteria. Information quality (such as usefulness and excellence) and recognized authority (such as credibility and reputation) were selected as important criteria. Additionally, past use experience and colleague recommendations were influential in the evaluation of a website as an information source.

Vibert et al. (2009) investigated relevance judgments by using PubMed, a free medical online database, sampling life scientists and neuroscientists. The result demonstrated that both cognitive ability and specialized knowledge of the topic influence individuals' search skill and relevance judgment. All participants were pressured by time, they all considered information search time as more important than information usability. In a follow-up study, Macedo-Rouet et al. (2012) examined how life scientists and neuroscientists apply relevance judgment criteria during information searches. While the article search process was largely similar across participants, they showed considerable individual differences in selecting articles according to their specialized knowledge or preferences. The relevance judgment criteria used most frequently were currency and wide coverage, followed by journal quality or reputation, highly diverse references, open access to the article, language, research method, and author reputation.

Overall, studies on physicians' relevance criteria in the context of an information search for patient treatment is lacking, particularly compared to the exponential increase in use of online clinical information resources in medicine. Accordingly, research on the online information search process of physicians is necessary, in particular the relevance judgment criteria they employ. Results of the research could be utilized to design search interfaces for online clinical information systems that are appropriate for physicians' search progress.

## METHOD

We investigated clinical information search process of physicians who seek to answer clinical questions to diagnose or treat their patients. We employed multiple data collection techniques, including pre and post in-depth interview, observation of information search process and self-assessment checklist responses.

Regarding the information search process frameworks, we applied Grad et al. (2011)'s three-stage model because it specifically concerns physicians' search process. It is a modified version of Saracevic and Kantor (1997)'s ACA (acquisition-cognition-application) model. Very few studies have used a three-stage search model. For instance, Bruce (1994) divided the search into before searching (after recognizing the problem), after finishing the search (interaction with the system), and after examining the selected literature, while Wang and White (1999) divided article citation behaviours into selecting, reading, and citing relevant articles to examine how relevance judgment criteria changed across these stages. We referred to the three stages in this study as (1) problem recognition (before the search); (2) system interaction (completion of information search); and (3) document interaction (reading and interacting with final articles).

The physicians were recruited using a purposive sampling method. At first, we asked the librarian of the hospital to introduce a chief physician at the internal medicine department. After interviewing the chief physician, we asked her/him to introduce other physicians with varying years of experience, gender, and searching skills. In order to avoid bias caused by differences in specialty and hospital environment, we recruited only internal medicine specialists from a large university hospital in Korea (Y University hospital, having 3,362

beds and 2,500 physicians). Physicians working in the university hospital environment are more likely to use online clinical information resources to search for case studies of unusual patients or up-to-date treatment guidelines (Davies 2011). Regarding sample size, Barry (1994) reported that interviewing around 9 participants led to redundancies and theoretical saturation (i.e., no more new data after the 9th interview); indeed, data saturation was achieved in our study at ten (10) physicians.

Details of participants are presented in Table 1. To determine participants' search skills, we asked them whether they use "Advanced" search menu of PubMed. Only two participants reported using the advanced search functions at least once a week; the rest used only the basic search menu.

Table 1: Characteristics of Research Participants (N = 10)

| ID | Gender | Specialty | Age | Position | Clinical experience (years) | Training on database use | Use of PubMed Advanced search |
|---|---|---|---|---|---|---|---|
| P1 | Female | Gastroenterology | 30 | Clinical Lecturer | 6 | Not received | No |
| P2 | Male | Gastroenterology | 35 | Clinical Lecturer | 8 | Received | Yes |
| P3 | Male | Gastroenterology | 33 | Clinical Lecturer | 9 | Received | Yes |
| P4 | Male | Gastroenterology | 36 | Clinical Lecturer | 9 | Not received | No |
| P5 | Female | Endocrinology | 30 | Clinical Lecturer | 6 | Not received | No |
| P6 | Male | Nephrology | 59 | Professor | 34 | Not received | No |
| P7 | Male | Endocrinology | 47 | Associate Professor | 21 | Not received | No |
| P8 | Male | Gastroenterology | 58 | Professor | 32 | Not received | No |
| P9 | Male | Cardiology | 46 | Professor | 20 | Received | No |
| P10 | Female | Gastroenterology | 43 | Associate Professor | 15 | Not received | No |

To make a self-assessment checklist to get insight of the relevance judgement, we used the relevance judgment criteria reported in Barry and Schamber (1998) which is commonly used as a basic criteria in relevance studies (Tang and Solomon 2001). Given the numerous relevance judgment criteria considered, we extracted a total of 30 criteria and categorized them into five categories according to their characteristics. The categories are information content, bibliographic information, presentation style, searcher characteristics, and physical environment (detailed in Table 2).

To check the physicians' preferred relevance judgment criteria across all three stages, we asked the 10 participants to rate the importance of 30 relevance judgment criteria in each search stage, before the start of searching, completion of searching and after selecting the final articles. We asked the participants to assess each of the 30 criteria on a scale of 0–100, with 0 indicating the least important and 100 the most. The magnitude estimation method with 0–100 is cited as more reliable method of evaluating relevance criteria than rating on either 1-5 or 1-10 scales (Bruce 1994; Janes 1991).

The overall research procedures are illustrated in Figure 1.

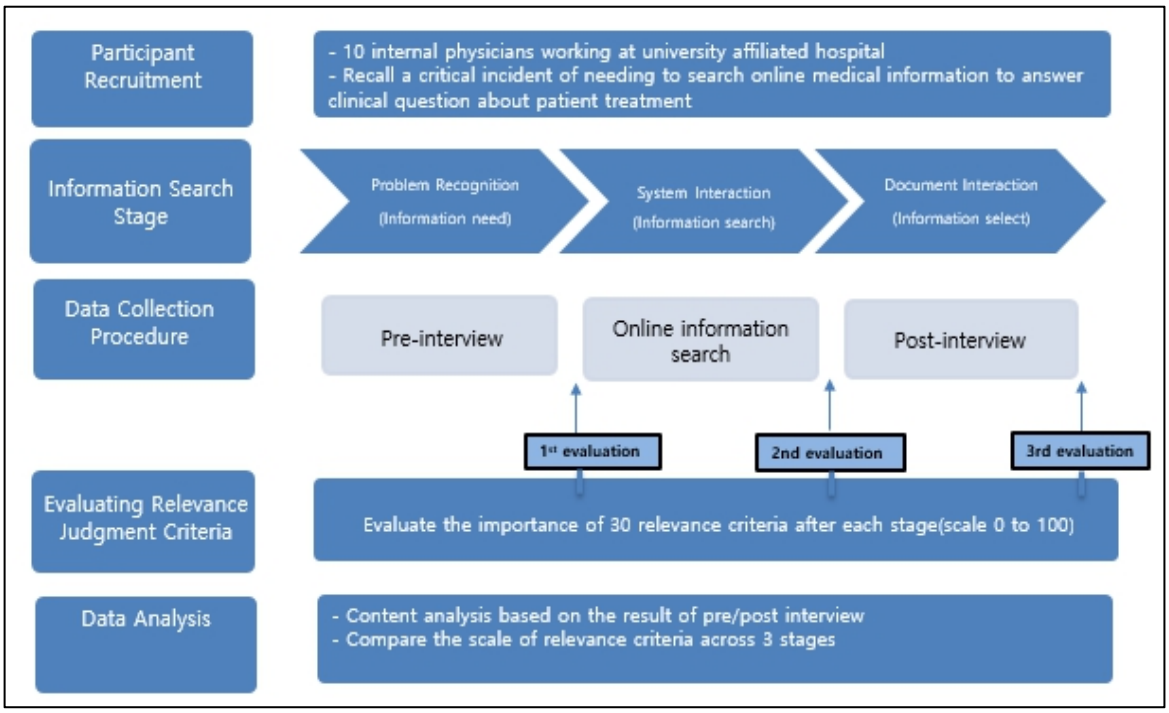


Figure 1: The Overall Research Procedures

Pre and post interviews and observation of information searching were conducted from at the physicians' offices, where they have easy access to the library's homepage, and each session took 40–90 minutes. We asked the participants to recall a critical incident of needing to search online medical information to answer clinical questions. The participants freely choose the medical information resources they prefer for their information searching. This hospital library provides various paid electronic resources such as UpToDate, Embase, Cochrane Library and ClinicalKey.

To observe the information retrieval process more accurately, during their online searching, each participant was asked to think aloud why they selected the specific article among others and to explain why that article was appropriate for the search task.

In the initial problem recognition stage, physicians encountered a patient treatment situation and recognized the need to search for information to solve the problem; however, the actual search had not yet begin in this stage. As a pre interview, we asked each participant to recall a memorable incident of needing to search online information resources to resolve a clinical question. Subsequently, we asked them about their information needs and search strategies during this incident, and then to rate each of the 30 relevance criteria for its importance before starting the information search.

In the system interaction stage, participants were asked to use the information search system to resolve their clinical question and to select relevant articles. While conducting the search, participants were told to think aloud about the keywords they were using and the number of returned articles. To observe the information retrieval process more accurately, they were asked the following questions:

(1) Why do they chose particular online clinical information resources;
(2) Why do they chose particular articles on screen;

(3) Why the articles were relevant to the search task;
(4) If the searched articles were satisfactory, why they were satisfactory, and what relevance judgment criteria they used; and
(5) Why they had disregarded unselected articles and why they were dissatisfied with these articles.
After finding all related information and finish searching, they completed the 30-item checklist again.

In the document interaction stage, participants selected final documents which could be applied to patient care among the retrieved documents. After that, they rated on the relevance criteria checklist a third time.

Finally, we conducted a post-search interview to debrief the information search process and probe participants' search behaviours and relevance assessments in the third stage and the previous two. Follow-up interviews were conducted by phone calls or e-mails whenever additional information was needed.

With interviewees' consent, the interviews were audio-taped during the search process. We employed NVivo 10 (QSR International Pty Ltd., Melbourne, Australia) to analyze the transcribed interview data. We used codes obtained from previous literature (Davies 2011) and generated possible new codes from the interview data. The study was approved by the institutional review board of the university.

## RESULTS

### Physicians' General Search behaviour for Patient Treatment

We initially asked participants about clinical questions and the current use of clinical information resources, to identify the characteristics of physicians' clinical information needs as a pre-interview. The participants typically used online information sources at least once a week for an average of 30–40 minutes. For patient treatment, physicians searched for information that would help them quickly decide on a patient treatment procedure or update themselves on the latest treatment trends. They preferred recently published review articles and case reports in highly reliable journals with well-written article outlines, thereby enabling ready understanding.

PubMed was the preferred choice for searching because of its familiarity, credibility, and vast information pool. P1 remarked, "*If the interface is not familiar; I would not use it. As I have been using PubMed since I was a medical student, I usually go to PubMed"*.

Google was also used frequently for its easy access, image search capability, and fast retrieval of results. P9 said, "*In Google, the images are comprehensively panoramic and provide pretty good pictures. If I find a good image, I click the paper corresponding to this image. It is much easier than reading the text only*".

In selecting information sources, physicians found an intuitive interface and convenient access highly important because of their time constraints. During their actual online search, most physicians used the basic search function due to their lack of search skills and they used certain targeted information resources familiar to them, such as PubMed. One such commonality was observed in the problem recognition stage, that is, most physicians went to Google or UptoDate to obtain background information on the disease. P5 confirmed this

in her response, "*If the disease is unfamiliar to me, I search Google a lot to grasp* [its] *basic concepts. A summary of lung cancer is available on Google. Here is a brief definition of the disease…*". Similarly P2 said, "*If I have a chance to search for a new disease that I don't know about, I'm the first to go into UpToDate. It's very useful when I get background information about a disease that I don't know because it's very well organized".*

Upon grasping the basic concept, physicians then visited PubMed to search for case reports and review articles dealing with cases like their patient's. P5 said, "*I often use the search filters* [of PubMed] *to find review articles and clinical trials first. In the PubMed search results, the types of the document we select are either reviews or clinical trials*".

When selecting the final documents, participants acknowledged scanning 10–20 relevant documents, selected two to five of them and saved these as PDF files for reference. Whether an article deals with a similar case as their patient's was a critical criterion for its being selected for the final set. The most preferred articles were either review articles or case reports from highly reliable journals, which can be understood at one glance.

The interviews with physicians revealed that the information search process itself was an important learning process. Many physicians reported that they could widen their background knowledge on diseases and collect ideas for patient treatment when scanning through the article contents, particularly the figures and tables.

## Physicians' Preferred Relevance Judgment Criteria Across All Three Stages

In each search stage, we asked the 10 participants to rate the importance of 30 relevance judgment criteria on a scale of 0–100. We then combined their scores for each criterion in each stage. Figure 2 shows the resulting total score of each of the 30 criteria, in order of importance.

Accuracy, topicality, and credibility were considered the most important relevance criteria throughout the information search process. The least important was the length of the document, author name/affiliation, and publisher name. P9 emphasized, "*The author's name is not so important. I sometimes see a well-known author's paper; however, the journal name is more important"*.

The importance of some criteria aligns with findings from previous studies (Barry 1994; Barry and Schamber 1998; Schamber 1991), however, there were some uniquely important criteria for physicians, including obtainability, time to obtain, understandability, article type, and journal name, as P6 remarked, "*The journal name is most important. The impact factor of the journal is also important. As I already know the impact factor of major journals, as soon as I see the name of the journal, I can recognize its credibility".*

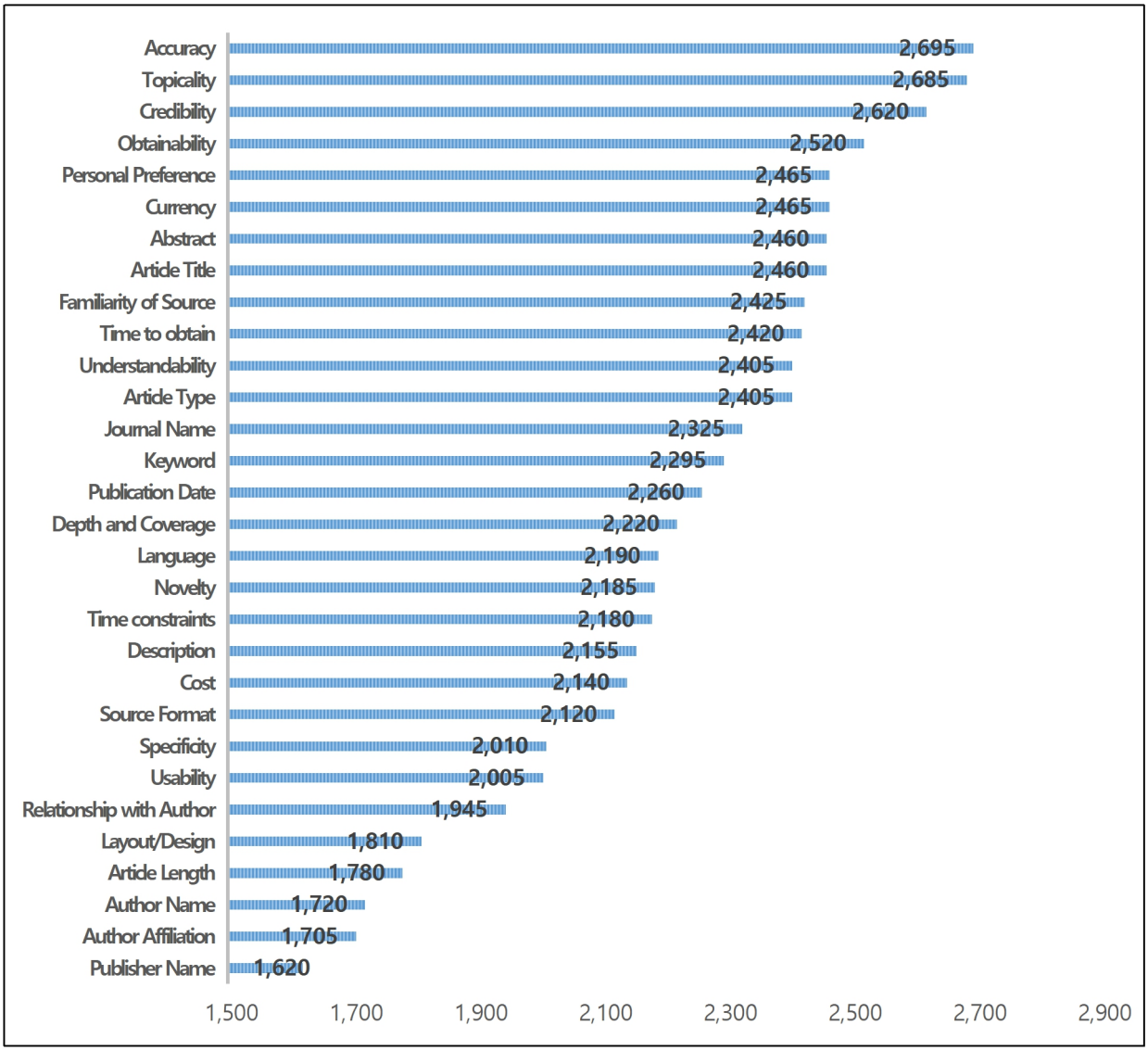


Figure 2: Physicians' Preferred Relevance Judgment Criteria

## Dynamics of Relevance Judgment Criteria

When reviewing the total scores of the participants' relevance assessment, we found that, unquestionably, accuracy, topicality, and credibility were the three most important criteria throughout the entire search process (Barry 1994; Schamber 1991; Barry and Schamber 1998). Conversely, the importance of most other criteria fluctuated considerably across the three stages. Table 2 depicts these relevance dynamics across search stages.

Besides the three most important criteria (accuracy, topicality, and credibility), the following criteria were considered important (in descending order of importance) in the problem recognition stage: currency, article title, time to obtain information, article type, and journal name. In the system interaction stage, obtainability and the abstract were considered important, while in the document interaction stage, personal preference, and understandability were considered important. In the following sections, we further detail the dynamics of individual relevance judgment criteria across the search stages based on five categories - information content, bibliographic information, presentation style, searcher characteristics, and physical environment

Table 2: Changes in Relevance Judgment Criteria across Search Stages

| Category | Criteria*** | | Importance level: Problem Recognition Score | Problem Recognition Rank | System Interaction Score | System Interaction Rank | Document Interaction Score | Document Interaction Rank | Total Score * | Rank ** |
|---|---|---|---|---|---|---|---|---|---|---|
| Information Content | **Accuracy** | How accurate is the information content? | **905** | **1** | **865** | **3** | **925** | **2** | **2,695** | **1** |
| | **Credibility** | Is the information content credible? | **845** | **3** | **860** | **4** | **915** | **3** | **2,620** | **3** |
| | **Currency** | Is the information content the most recent? | **835** | **4** | 785 | 12 | 845 | 7 | 2,465 | 6 |
| | Depth and Coverage | Is the amount of information too much or too little? | 780 | 12 | 725 | 18 | 715 | 22 | 2,220 | 16 |
| | Novelty | Is the information content new and fresh? | 755 | 15 | 715 | 20 | 715 | 22 | 2,185 | 18 |
| | **Topicality** | Is it topically relevant to my clinical question? | **855** | **2** | **900** | **1** | **930** | **1** | **2,685** | **2** |
| Bibliographic Information | **Article Title** | Are the article title, abstract, and keywords relevant to my search topic? | **825** | **5** | 800 | 8 | 835 | 9 | 2,460 | 7 |
| | **Abstract** | | 795 | 10 | **815** | **5** | 850 | 6 | 2,460 | 7 |
| | Keyword | | 775 | 13 | 715 | 20 | 805 | 11 | 2,295 | 14 |
| | Article Length | Is the number of pages, article type (e.g., review), or language used in the article satisfactory? | 465 | 30 | 630 | 25 | 685 | 25 | 1,780 | 27 |
| | Article Type | | 810 | 6 | 800 | 8 | 795 | 13 | 2,405 | 12 |
| | Language | | 685 | 20 | 755 | 16 | 750 | 15 | 2,190 | 17 |
| | Author Affiliation | How important is the author name or author affiliation to relevance judgment? | 540 | 27 | 600 | 27 | 565 | 29 | 1,705 | 29 |
| | Author name | | 535 | 28 | 600 | 27 | 585 | 27 | 1,720 | 28 |
| | Journal Name | Does the journal name or publisher name influence relevance judgment? | 810 | 6 | 775 | 14 | 740 | 16 | 2,325 | 13 |
| | Publisher Name | | 535 | 28 | 570 | 29 | 515 | 30 | 1,620 | 30 |
| | Publication Date | Does the publication date have an important impact on relevance judgment? | 700 | 18 | 775 | 13 | 785 | 14 | 2,260 | 15 |
| Presentation Style | Detailed Description | Is the information concrete and described in detail? | 685 | 20 | 735 | 17 | 735 | 17 | 2,155 | 20 |
| | Layout/Design | Is the layout or design of the documents satisfactory? | 610 | 26 | 615 | 26 | 585 | 27 | 1,810 | 26 |
| | Source Format | Is it a journal? Does the preference differ depending on source format? | 675 | 22 | 715 | 20 | 730 | 18 | 2,120 | 22 |
| | Specificity | Is the provided information properly presented and described with appropriate focus? | 630 | 25 | 660 | 24 | 720 | 20 | 2,010 | 23 |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Relationship with Author | Is the author my colleague or respect the author? | 655 | 23 | 570 | 29 | 720 | 20 | 1,945 | 25 |
| | Familiarity of Source | Do I know well or trust the information source? | 795 | 10 | 795 | 10 | 835 | 9 | 2,425 | 9 |
| **Searcher Characteristic** | **Personal Preference** | Am I interested in the topic or having fun personally? | 770 | 14 | **815** | **5** | **880** | **4** | **2,465** | **5** |
| | **Understandability** | Is it written in a manner I can understand? | 730 | 16 | 795 | 10 | **880** | **4** | 2,405 | 11 |
| | Usability/ Subjective Value | Is it meaningless because it is similar to an article I have already read? Can I use it later? | 645 | 24 | 675 | 23 | 685 | 25 | 2,005 | 24 |
| | Cost | Is the information free to obtain or do I need to pay? | 715 | 17 | 720 | 19 | 705 | 24 | 2,140 | 21 |
| | **Obtainability** | Is the information easily obtainable? | 805 | 9 | **870** | **2** | 845 | 7 | **2,520** | **4** |
| **Physical Environment** | Time Constraints | Do I have constraints regarding information search time? | 690 | 19 | 765 | 15 | 725 | 19 | 2,180 | 19 |
| | Time to Obtain | How long does it take to obtain the information? | 810 | 6 | 810 | 7 | 800 | 12 | 2,420 | 10 |

*Note.* * Total scores of all ten interviewees on 30 relevance judgment criteria (score range: 1–100), Maximum score is 3,000 (100 score X 10 interviewees X 3 search stages)
**Total Score Rank
*** Bolded top 5 important criteria on each search stage

## Information Content Category

The information content is an item that measures how accurate and reliable the search content is, and whether the information searchers feel it is important to their research. Information content category includes accuracy, credibility, currency (most recent), depth and coverage, novelty (new and fresh) and topicality (topical relevance) of information content.

Accuracy was considered the most important by physicians (2,695 points, 1st rank), followed by topicality (2,685, 2nd) and credibility (2,620, 3rd). Other important criteria in the information content category were currency, depth/scope, and novelty (in descending order of importance).

Many interviewees reported that articles published in a reputable journal with a high impact factor were guaranteed to be credible because they had undergone expert review. P1 said, "*When deciding on the final articles to use in treating patients, I think credibility and accuracy are the most important* [characteristics]".

Currency was important because interviewees preferred articles on new treatment methods or that had been recently published. If two articles addressing a similar topic were published in different years, participants tended to look at the more recent one because the newer one would be more accurate, as it would be built on previous findings and address previous limitations. P6 said, "*New practice is important. Whenever I find an article about a new treatment, I always click* [on it]".

Interestingly, the scores of most information content criteria fluctuated during the search process, as shown in Figure 3. In the problem recognition stage, accuracy was considered the most important criterion, whereas topicality was first ranked in the document interaction stage. As the search progressed from the system interaction stage to the document interaction stage, scores in topicality (900 to 930), accuracy (865 to 925), and credibility (860 to 915) all showed an increasing trend.

Conversely, currency ranked high in the problem recognition stage (835, 4th), but its importance had declined by the system interaction stage (785, 12th). This suggests that currency is favoured as an initial screener earlier in the information search than in the later stages.

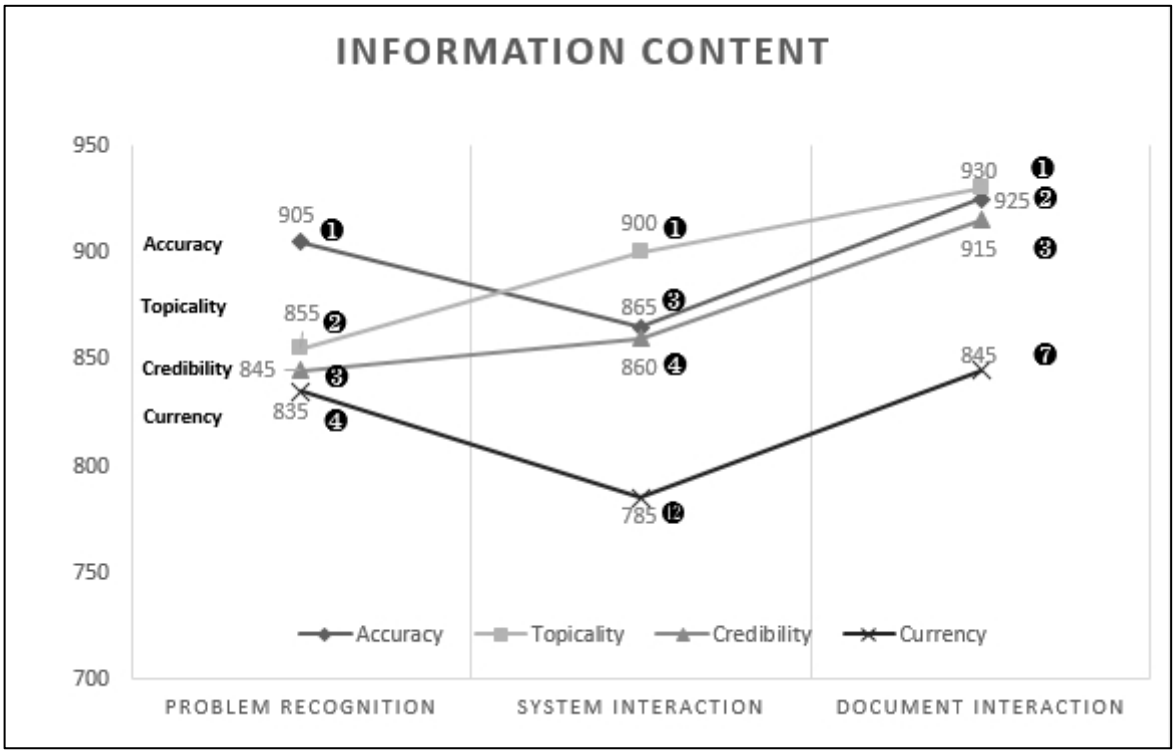


Figure 3: Changes in Relevance Criteria in the Information Content Category

.

## Bibliographic Information Category

The bibliographic information is a tool designed to make it easier to find the required literature. It can also be referred to as a list of works cited. It is a list of bibliographies of a number of literature or books arranged and compiled in a certain way. The bibliographic information contained criteria relating to document characteristics such as article title, author name and publication date. All physicians rated both article title and abstract as most important, followed by article type, journal name, keyword, publication date, language, article length, author name, author affiliation, and publisher name, in that order. Notable criteria at the problem recognition stage were article title (825, 5th), article type, and journal name (810 each, shared 6th) as shown in Figure 4.

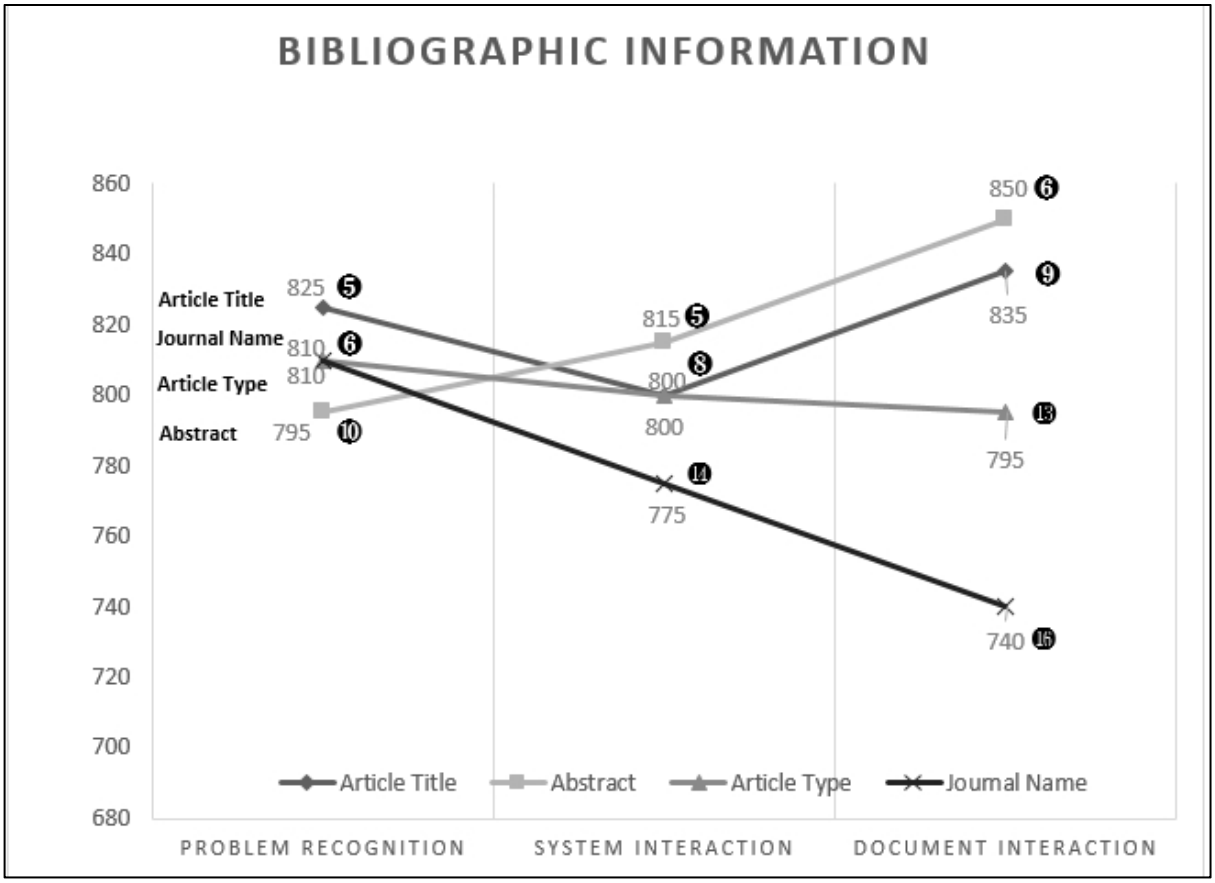


Figure 4: Changes in Relevance Criteria in the Bibliographic Information Category

The article title was the most important evaluation criteria in the problem recognition stage, the reason being that the title helped physicians determine the article's topical relevance to the area of interest. In fact, most interviewees responded that they initially glanced at the title and judged its relevance to the topic to determine whether to continue reading the abstract or main text. P2 remarked, "*If the title looks valid to some extent; I go deeper. The title is the most crucial.* [If it has] *a strange title, I don't go further*".

Figure 4 shows the changes in the importance of these criteria over the three search stages. In the problem recognition stage, article title received the highest score (825, 5th), while abstract (795, 10th) had a much lower score. However, the importance of the abstract increased to 815 (5th) in the system interaction stage and 850 (6th) in document interaction stage. Abstract became the highest scoring criterion in the bibliographic information category in the system and document interaction stages. Importance of article type markedly decreased from the problem recognition stage to the document interaction stage(from 810, 6th to 795, 13th rank), and this is reflected in P1's response, "*I usually use a limit function to filter review and case reports only before I browse the searched documents to save time*."

Journal name plunged regarding importance from 810 points (6th) in the problem recognition stage to 775 points (14th) in the system interaction stage and 740 points (16th) in the document interaction stage. The results imply that the bibliographical information was considered important only during the early search stage. P3's response is a case in point, "*So, I tend to look at the journal name a lot in the beginning. If the journal is well*

*known or familiar to me, generally the articles look fine*".
In contrast, the article length was not considered important in the problem recognition stage (465 points, 30th), but its score rose substantially to 630 points (25th) in the system interaction stage and 685 points (25th) in the document interaction stage.

Summarizing the previous two sections, it is evident that the content of the information, such as accuracy and topicality, is the most prominent at the beginning of the search, whereas the format of the information, such as article length, becomes more important as the search progresses. The latter tendency was observed in presentation style category as well, as described in the following section.

## Presentation Style Category

Presentation style refers to the way in which the information resources are presented, including description level, layout, design, and source format (e.g., journal, book). The most important criteria in this category were, in descending order, detailed description, source format, specificity, and layout/design, as shown in Figure 5.

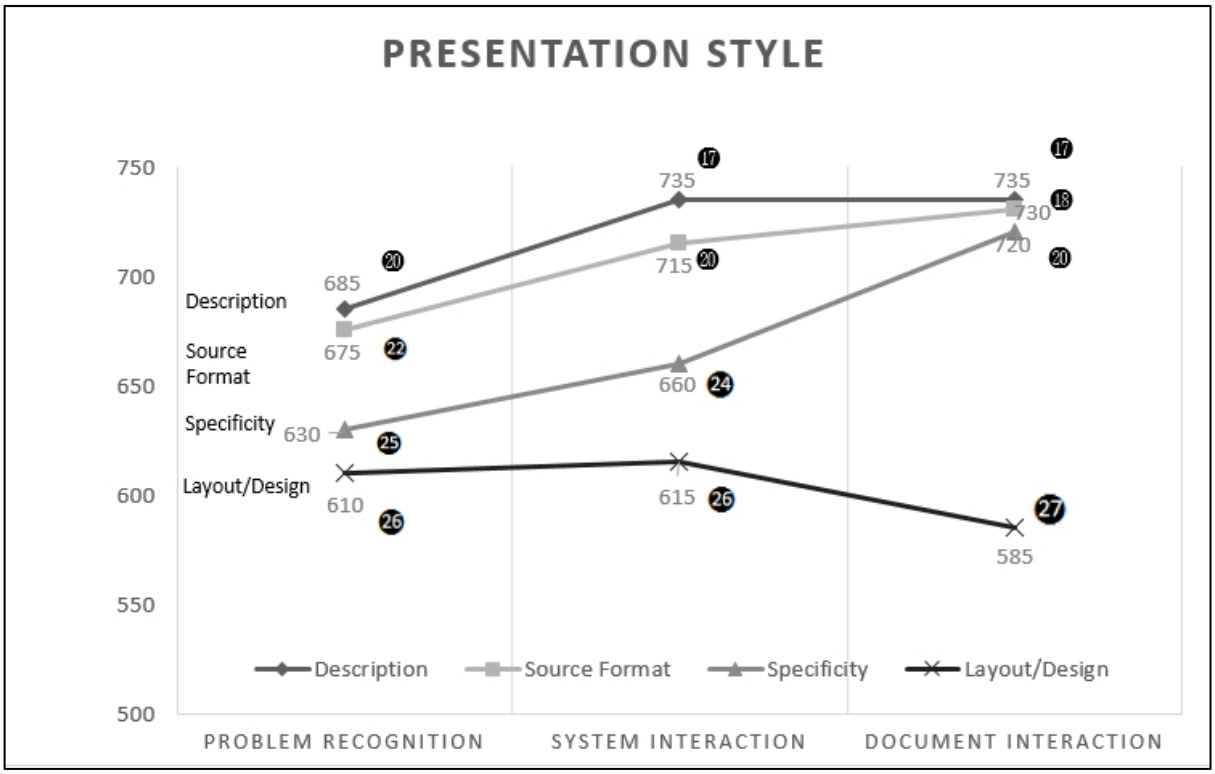


Figure 5: Changes in Relevance Criteria in the Presentation Style Category

A notable criterion in this category was specificity (i.e., "Is the information properly presented and described with appropriate focus?"), which rose substantially in its importance over time. Specifically, it increased from 630 points (25th) at the problem recognition stage to 720 (20th) at the document interaction stage. This finding is reasonable considering that specificity cannot be fully confirmed until physicians pay close attention to and interact with the particular document, as indicated by P4, "*The content is too easy. It provides very simple information only.* [So] *let's see other articles. I like more detailed explanations when I apply* [article information] *to my patient treatment"*

This finding implies that information presentation style is influential when physicians are making their final document selections for patient treatment. Physicians also mentioned that the figure and tables were crucial information, as P9 illustrated, "*I always see tables and graphs first… I don't download and read all the words in the articles, but just skim the article for patient treatment*."

## Searcher Characteristics Category

Searcher characteristics refer to searchers' personal experience, affectiveness, and knowledge influences their perception of a particular information resource. The highest score in this category was personal preference, followed by source familiarity (reputation/visibility), understandability, and personal usability/subjective values as depicted in Figure 6.

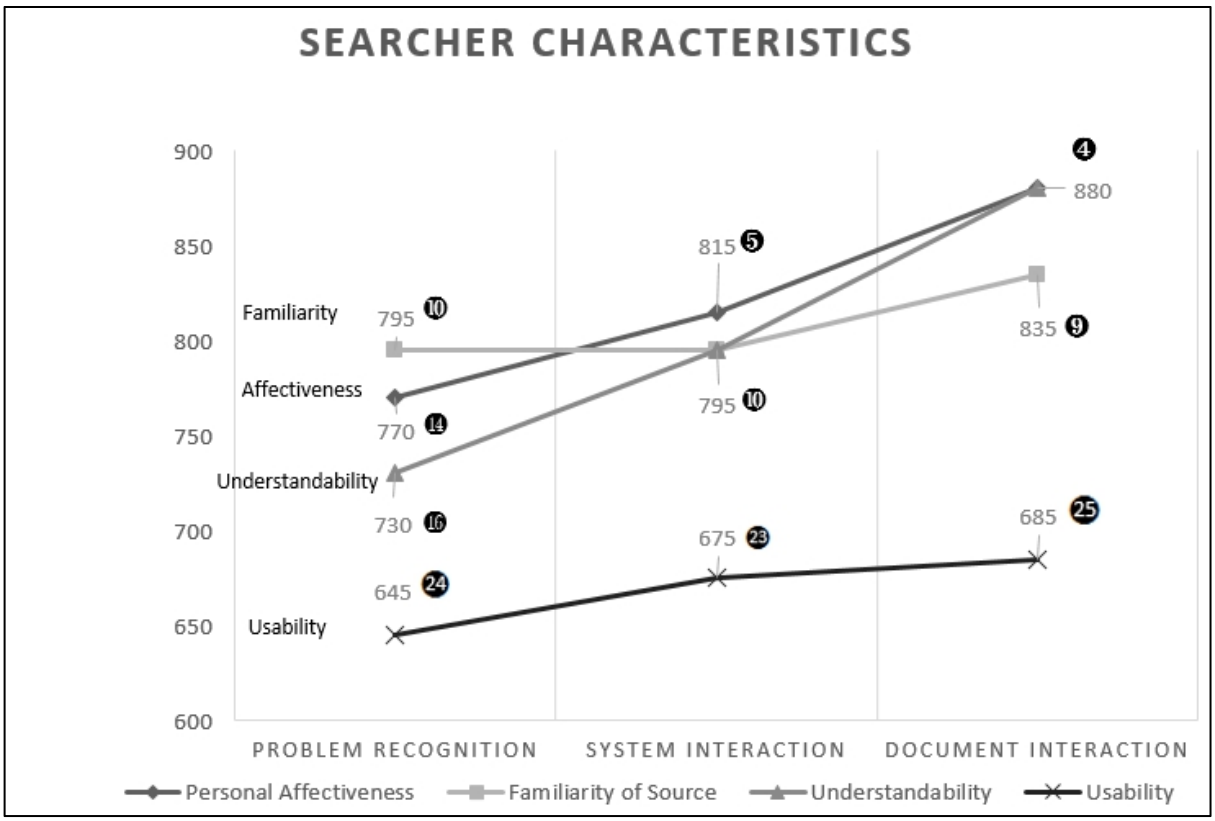


Figure 6: Changes in Relevance Criteria in the Searcher Characteristics Category

None of the criteria in this category ranked particularly high in either the problem recognition or system interaction stage; however, in the document interaction stage, both personal preference(affectiveness) and understandability reached 4th rank (880 points) (Figure 6). The importance of personal preference (770:14th → 815:5th → 880:4th) and understandability (730:16th → 795:10th → 880:4th) increased as the search progressed; during physicians' selection of the final articles for application to patient treatment, both criteria were considered important, as refelected in P9's remarks, "*I just browse the paper at the time of patient treatment, especially focusing on tables and graphs. For patient treatment, a well-organized paper is good for quick understanding*".

## Physical Environment Category

Physical environment relates to external factors affecting the obtainability of information resources, such as cost and availability. The most important relevance criterion in this

category across all three stages was obtainability (i.e., the possibility of acquiring information), followed by time to obtain, time constraints, and cost as shown in Figure 7.

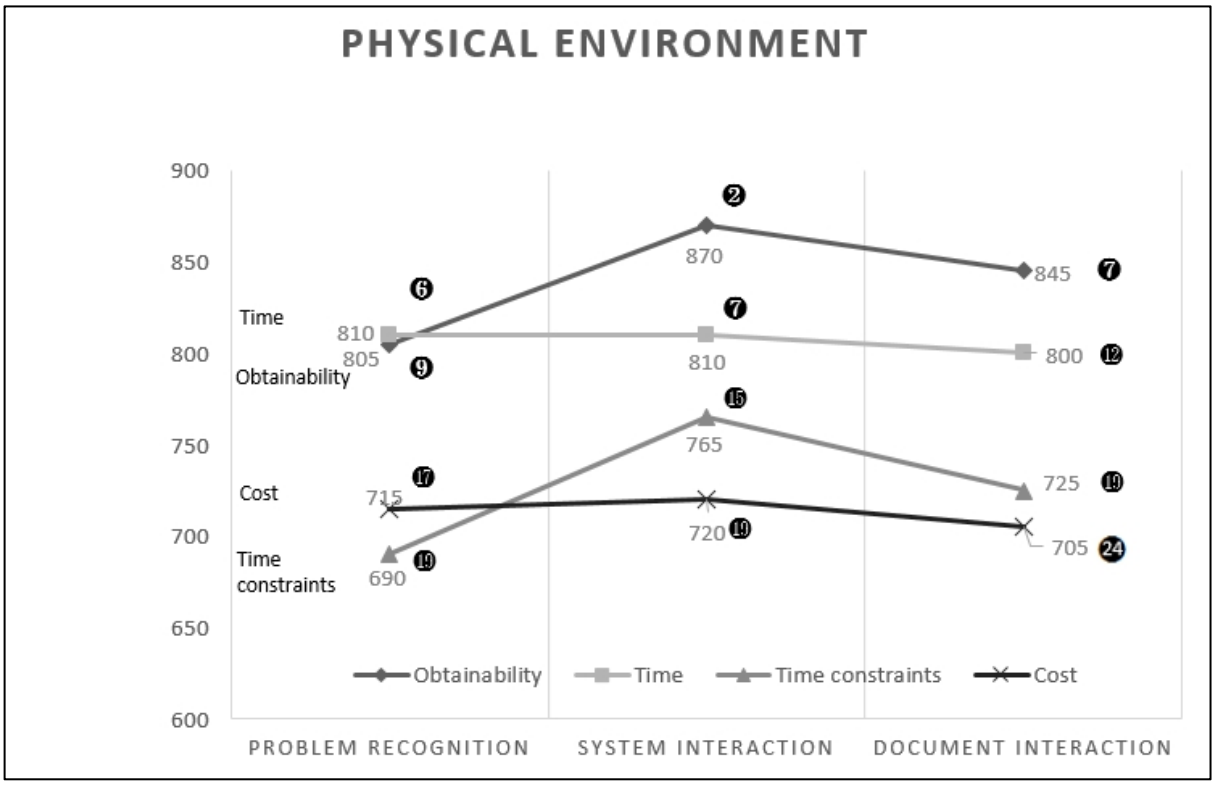


Figure 7: Changes in Relevance Criteria in the Physical Environment Category

Time constraints was considered as most important during system interaction, compared with other stages. Especially, obtainability was particularly crucial during the system interaction stage. Furthermore, as shown in Figure 7, both obtainability (870, 2nd) and time to obtain the information (810, 7th) were important criteria in the system interaction stage but declined slightly in the document interaction stage.

As participants progress into the system interaction stage, obtainability emerged as an important evaluation criterion. P4 mentioned, "*Searching PubMed on the library homepage is very convenient because there is a direct link to the full-text article. However, sometimes I cannot get an article immediately because our library has not subscribed to this journal. It is very inconvenient because I cannot see the full-text immediately*."

## DISCUSSION

Although topical relevance was revealed as the most important relevance criterion across all three stages in this study, certain criteria were more important in certain stages than in others. In the system interaction stage, information obtainability was scored higher than in other search stages. This indicates that situational relevance criteria are more important in this stage. In the document interaction stage, affective relevance criteria (e.g. personal preference) were considered important. Taken together, the findings suggest that the relevance criteria change throughout the search process.

In the problem recognition stage, physicians favoured article type and journal name. A unique preference for physicians was well-summarized review articles, as they valued contents that could be instantly referred to at the time of patient treatment; such a preference was not observed in previous studies with general users. These findings are congruent with Vakkari, Pennanen and Serola (2003)'s suggestion of the need for background information when users deal with unfamiliar topics (via sources such as textbooks, encyclopedias, and review papers). The findings also support Ely et al. (2005)'s recommendation that the ideal medical information search system for physicians emphasizes credibility, accessibility, and a presentation style suitable for a clinical setting. Author name, author's affiliated organization, and novelty were not considered important by physicians, whereas these criteria were considered crucial by general or academic users (Barry and Schamber 1998; Wang and White 1999).

In the system interaction stage, physicians preferred obtainability. The inability to access information and insufficient time were the main factors preventing successful online clinical information search. While most criteria in the information content category were considered more important: obtainability ranked 4th out of 30 items and time to obtain the information ranked 10th; thus, they were important criteria for busy physicians. These findings are consistent with Ely et al. (2005)'s study targeting physicians, which found that convenient access and time to obtain the information were important. These findings also support Davies (2011)'s study, which indicated that information inaccessibility and insufficient time were key factors hampering successful online clinical information searches among physicians.

In the document interaction stage, twelve criteria had scores of 800 or higher. Given that only nine criteria received 800 or more points in the two earlier stages, it can be inferred that physicians might assess relevance in a stricter manner during final article selection. These findings are largely consistent with results of previous relevance research conducted with general or academic users (Barry and Schamber 1998; Maglaughlin and Sonnenwald 2002). When physicians selected the documents to apply to their patients, personal preference and understandability were evaluated as the important relevance criteria. Physicians prefer well-organized and detailed descriptions, especially when applying information to patient treatment; this is markedly different from what was found with general information users in previous studies (Barry and Schamber 1998; Schamber 1991).

Drawing on these results, a model of the dynamics of relevance assessment in physicians' clinical information search process (Figure 8) was constructed. The model illustrates characteristics that influence physicians' clinical online information search process. Moving from problem recognition towards system interaction and document interaction, there are evident changes in physicians' cognitive status, important relevance criteria, and recommendations for designing clinical information search systems.

Findings of this study have some practical implications for designing interactive information systems that effectively accommodate relevance dynamics in the clinical information search process. First, in the problem recognition stage, it was found that bibliographic information, such as article title, journal name, and article type, was ranked highly by physicians. An effective information system might display search results with focusing on the article title and list complete journal names (not abbreviated ones) together with the journal impact factor, as this could help users understand the reputation or credibility of the journals.

Second, in the system interaction stage, physicians like to read the full-text article to test their hypotheses concerning patient treatment immediately. Because of time constraints, obtainability becomes critical in this stage. Thus, search convenience might be improved by showing, with the search results, whether the institution's library currently subscribes to the journal, an option to request an interlibrary loan, and, if possible, a link to a site where the article is freely available.

Third, the document interaction stage represents a state to acquire greater understanding of the targeted disease and treatment. In this stage, personal preference and understandability had much higher importance than in earlier stages. Therefore, when selecting the final articles, allowing users to customize the search interface according to their personal preference, such as preferred journals or article types, would enhance their search satisfaction. Most participants in this study utilized only basic search functions, it might be necessary to make these customized functions more intuitive and visible.

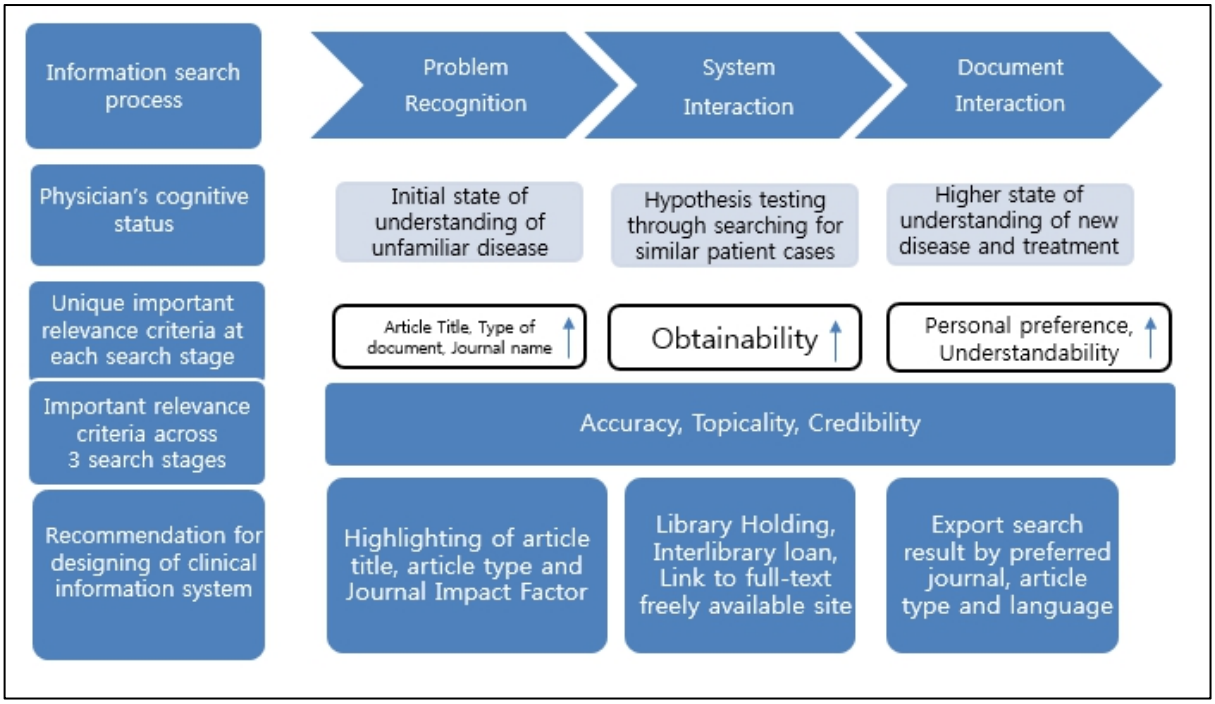


Figure 8: Modeling the Dynamics Of Relevance Assessment in Physicians' Clinical Information Search Process

## CONCLUSIONS

We examined ten internal medicine physicians' relevance assessments during the three stages of the information search process using multiple data collection techniques. While each physician had their idiosyncratic behaviours, we did find some commonalities in their information search behaviour and relevance assessments in each of the three stages (problem recognition, system interaction, and document interaction). Accuracy, topicality, credibility, and currency were considered important because physicians were searching for information to treat patients. Article type, journal name, obtainability, personal preference, understandability also ranked as the important category comparing other general user groups, such as professors and students in other subject area. The results show that the

relevant criteria change throughout the search process. In the system interaction phase, information acquisition received a higher rating than other search phases. At the document interaction stage, emotional relevance criteria (e.g. personal preference) were considered important.

While the present study offers meaningful implications for both relevance research and system design, it is not without limitations. One such limitation is the lack of validation - only ten internal medicine specialists from the internal medicine department were interviewed, as we wanted to limit possible differences in information search behaviour by specialty. Another limitation relates to the fact that the findings cannot be generalized beyond hospitals that have similar information services as Y university hospital. Thus, future studies could expand to other hospitals and other specialties such as dermatology or medium-sized hospitals, to further validate the findings and the proposed model.

## ACKNOWLEDGEMENT

This study draws on some parts of Soon Kim's Ph.D. thesis: *An investigation of physicians' online information search process and the dynamics of relevance judgment*. Seoul: Department of Library and Information Science, Ewha Womans University; 2016. No grant from any public, commercial, or non-profit funding agency was offered for the conduct of this research.

## REFERENCES

Barry, C.L. 1994. User-defined relevance criteria: An exploratory study. *Journal of the American Society for Information Science,* Vol. 45, no. 3: 149–59.

Barry, C.L and Schamber, L. 1998. Users' criteria for relevance evaluation: A cross-situational comparison. *Information Processing & Management,* Vol. 34, no.2–3: 219–36.

Bateman, J. 1998. Changes in relevance criteria: A longitudinal study. *Proceedings of the 61st American Society for Information Science Annual Meeting* (ASIS' 98), Vol. 8, no. 23-32.

Bruce, H.W. 1994. A cognitive view of the situational dynamism of user-centered relevance estimation. *Journal of the American Society for Information Science,* Vol. 45, No. 3: 142–148.

Clarke, M.A., Belden, J,L,, Koopman, R.J., Steege, L.M., Moore, J.L., Canfield, S.M. and Kim, M.S. 2013. Information needs and information-seeking behaviour analysis of primary care physicians and nurses: a literature review. *Health Information & Libraries Journal,* Vol.30, no.3: 178–90.

Daei, A., Soleymani, M.R., Ashrafi-rizi, H., Zargham-Boroujeni, A. and Kelishadi, R. 2020a. Clinical information seeking behavior of physicians: A systematic review. *International Journal of Medical Informatics.* Available at: https://doi.org/10.1016/ j.ijmedinf.2020.104144.

Daei, A., Soleymani, M. R., Ashrafi-rizi, H., Kelishadi, R., and Zargham‐Boroujeni, A. 2020b. Personal, technical and organisational factors affect whether physicians seek answers to clinical questions during patient care: A literature review. *Health Information & Libraries Journa*l. Available at: https://doi.org/10.1111/hir.12323.

Davies, K. 2011. Information needs and barriers to accessing electronic information: hospital-based physicians compared to primary care physicians. *Journal of Hospital Librarianship,* Vol.11, no.3: 249–60.

Del Fiol, G., Workman, T.E. and Gorman, P.N. 2014. Clinical questions raised by clinicians at the point of care: a systematic review. *JAMA Internal Medicine*, Vol. 174, no. 5: 710–718.

Ely, J., Sheroff, J., Hambliss, M., Bell, M. and Osenbaum M. 2005. Answering physicians' clinical questions: Obstacles and potential solutions. *Journal of the American Medical Informatics Association,* Vol. 12, no. 2: 217–24.

Hollis, K. F., Roberts, K., Bedrick, S., and Hersh, W. R. 2020. Addressing the search challenges of precision medicine with information retrieval systems and physician readers. *30th Medical Informatics Europe Conference* (MIE 2020) (pp. 813-817). IOS Press.

Grad, R., Pluye, P., Granikov, V., Johnson-lafleur, J., Shulha, M., Sridhar, S.B., Moscovici, J.L., Bartlett, G., Vandal, A.C., Marlow, B. and Kloda, L. 2011. Physicians' assessment of the value of clinical information: operationalization of a theoretical model. *Journal of the American Society for Information Science and Technology,* Vol. 62, no. 10: 1884–1891.

Hersh, W. 1994. Relevance and retrieval evaluation: Perspectives from medicine. *Journal of the American Society for Information Science,* Vol. 45, no.3: 201-206.

Hughes, B., Wareham, J. and Joshi, I. 2010. Doctors' online information needs, cognitive search strategies, and judgments of information quality and cognitive authority: How predictive judgments introduce bias into cognitive search models. *Journal of the American Society for Information Science and Technology,* Vol. 61, no.3, 433-452.

Janes, J.W. 1991. Relevance judgments and the incremental presentation of document representations. *Information Processing & Management,* Vol. 27, no.6: 629–46.

Kuhlthau, C. C. 1991. Inside the search process: Information seeking from the user's perspective. *Journal of the American Society for Information Science,* Vol. 42, no. 5: 361-371.

Macedo-Rouet, M., Rouet, J. F., Ros, C. and Vibert, N. 2012. How do scientists select articles in the PubMed database? An empirical study of criteria and strategies. *European Review of Applied Psychology*, Vol. 62, no. 2: 63-72.

Maglaughlin, K.L. and Sonnenwald, D.H. 2002. User perspectives on relevance criteria: A comparison among relevant, partially relevant, and not-relevant judgments. *Journal of the American Society for Information Science and Technology,* Vol. 53, no. 5: 327–42.

Mikalef, P., Kourouthanassis, P. E. and Pateli, A. G. 2017. Online information search behaviour of physicians. *Health Information & Libraries Journal,* Vol. 34, no. 1: 58-73.

Pluye, P., Grad, R.M., Dunikowski, L.G. and Stephenson, R. 2005. Impact of clinical information-retrieval technology on physicians: a literature review of quantitative, qualitative and mixed methods studies. *International Journal of Medical Informatics,* Vol.74, no. 9: 745–68. Available at: https://doi.org/10.1016/j.ijmedinf.2005.05.004.

Qu, J., Arguello, J., and Wang, Y. 2020. Towards explainable retrieval models for precision medicine literature search. *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval* (pp. 1593-1596). Available at: https://doi.org/10.1145/3397271.3401277.

Saracevic T. 1997. The stratified model of information retrieval interaction: Extension and applications. *Proceedings of the Annual Meeting-American Society for Information Science*, Vol. 34: 313–27.

Saracevic. T. and Kantor, K.B. 1997. Studying the value of library and information services: Part I. Establishing a theoretical framework. *Journal of the American Society for Information Science*, Vol. 48, no.6: 527–542.

Schamber, L. 1991. Users' criteria for evaluation in a multimedia environment. *Proceedings of the ASIS Annual Meeting*, Vol.28: 126–33.

Spink, A., Greisdorf, H. and Bateman J. 1998. From highly relevant to not relevant: examining different regions of relevance. *Information Processing & Management,* Vol.34, no. 5: 599–621.

Tang, R. and Solomon P. 2001. Use of relevance criteria across stages of document evaluation: On the complementarity of experimental and naturalistic studies. *Journal of the American Society for Information Science and Technology,* Vol. 52, no. 8: 676–85.

Taylor, A.R., Cool, C., Belkin, N.J., and Amadio, W.J. 2007. Relationships between categories of relevance criteria and stage in task completion. *Information Processing & Management,* Vol. 43, no. 4: 1071–1084.

Uprety, S., Su, Y., Song, D. and Li, J. 2018. Modeling multidimensional user relevance in IR using vector spaces. *41st International ACM SIGIR Conference on Research & Development in Information Retrieval* (pp. 993-996). Available at: https://doi.org/10.1145/3209978.3210130

Vakkari, P., Pennanen, M. and Serola, S. 2003. Changes of search terms and tactics while writing a research proposal. *Information Processing & Management*, Vol. 39, no. 3: 445–63.

Vibert, N., Ros, C., Bigot, L. L., Ramond, M., Gatefin, J. and Rouet, J. F. 2009. Effects of domain knowledge on reference search with the PubMed database: An experimental study. *Journal of the American Society for Information Science and Technology,* Vol. 60, no. 7: 1423-1447.

Wang, P. and Soergel D. 1998. A cognitive model of document use during a research project. Study I. Document selection. *Journal of the American Society for Information Science,* Vol.49, no. 2: 115–33.

Wang, P.L. and White, M.D. 1999. A cognitive model of document use during a research project. Study II. Decisions at the reading and citing stages. *Journal of the American Society for Information Science,* Vol.50, no. 2: 98–114.

Westbrook, J.I, Gosling, A.S, and Westbrook, M.T. 2005. Use of point-of-care online clinical evidence by junior and senior doctors in New South Wales public hospitals. *Internal Medicine Journal,* Vol. 35, no. 7: 399–404.