# Analysis of Search in an Online Clinical Laboratory Manual

*Michael Blechner, MD,[1] Joshua Kish, MD,[2] Vivek Chadaga, MFA,[2] and Anand S. Dighe, MD, PhD[2]*

**Key Words:** Informatics; Management; Administration; Clinical chemistry

DOI: 10.1309/PHAMDT8NN44WMBDF

## Abstract

*Online laboratory manuals have developed into an important gateway to the laboratory. Clinicians increasingly expect up-to-date laboratory test information to be readily available online. During the past decade, sophisticated Internet search technology has developed, permitting rapid and accurate retrieval of a wide variety of content. We studied the role of search in an online laboratory manual. We surveyed the utilization of search technology in publicly available online manuals and examined how users interact with the search feature of a laboratory handbook. We show how a laboratory can improve its online handbook through insights gained by collecting information about each user's activity. We also discuss future applications for search-related technologies and the potential role of the online laboratory manual as the primary laboratory information portal.*

Laboratories have long recognized the need to provide their customers with an updated, reliable source of information about their diagnostic testing services.[1,2] Historically, this was accomplished by using printed materials that laboratories have found challenging to keep up-to-date owing to the ever-changing nature of laboratory policies, test menus, specimen requirements, and reference ranges. An online laboratory manual has been shown to be a useful means for storing and widely distributing the latest laboratory information and can serve as an intuitive portal for clinician inquiries to the laboratory.[3-5] However, with the increasing number and sophistication of clinical laboratory tests, rapid and accurate retrieval of information from an online manual requires a robust search mechanism. The promise of search technology is that it can provide a simple, familiar interface that permits users to quickly find and access the specific content they seek.

In this report, we discuss the role of search technology in an online laboratory manual. We assessed how the search process has been implemented in publicly available handbooks, and we show that capture and analysis of user search information can result in an improved handbook. Furthermore, we describe future applications of search technology for the laboratory.

## Materials and Methods

### Setting

The Massachusetts General Hospital (MGH) is an 898-bed tertiary care academic medical center in Boston, MA. The MGH clinical laboratories support all of the inpatient medical, surgical, pediatric, and obstetric services of the hospital, as well as extensive primary care and specialty outpatient practices


DOI: 10.1309/PHAMDT8NN44WMBDF

extending into the greater Boston community. The clinical laboratories include the core lab (chemistry-hematology) microbiology, blood transfusion services, and various specialty laboratories (immunology, diabetes, health center laboratories, and neurochemistry).

### Online Manual

The MGH online laboratory handbook is available on the Internet at http://mghlabtest.partners.org. The information for each test in the handbook is stored in a Microsoft SQL Server 2000 database (Microsoft, Redmond, WA). The handbook page for a given test is created dynamically from the database using Active Server Pages technology (Microsoft).

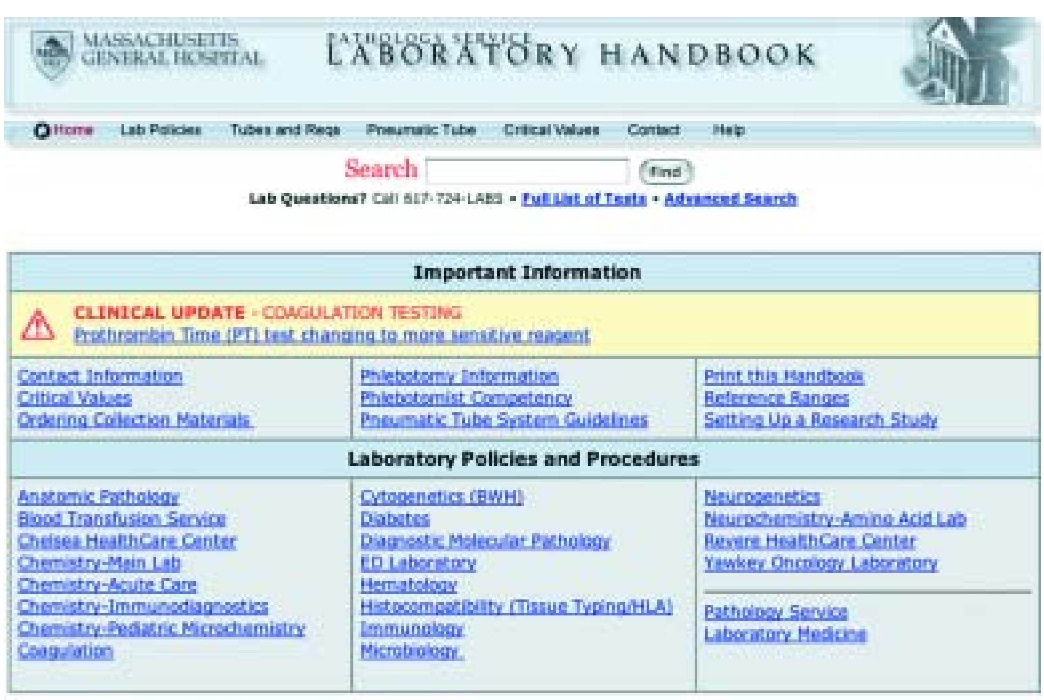
There are 2 methods that can be used to search for a laboratory test in our online handbook **❚Figure 1❚**. The first is an alphabetical listing of tests and test synonyms. The second pathway is free text search. In free text search mode, the user types into a search box and hits the "Find" button. For each test in the laboratory handbook database, there is a list of key words associated with the test. The free text search feature that we implemented takes what the user types in the search box and does the following: (1) removes extraneous punctuation (commas, semicolons); (2) removes common words (eg, "of," "for," "and," "the," and "or"); (3) breaks apart user input into strings (sequences of characters) based on the space character; and (4) retrieves tests that have each of these strings present in the starting position of at least 1 of their key words. For example, the user-entered text "CMV for antigen" results in a query to the database retrieving tests that contain at least 1 key word starting with "CMV" and at least 1 key word starting with "antigen." The rationale for requiring that all user-entered key words are present is to permit the algorithm to quickly narrow the search results based on user input.

### Storage of Search Information

We store detailed information each time a user performs a free text search or uses the alphabetical test listing. The stored data include user login alias, type of search (free text or alphabetical list), date and time of activity, search text (what the user typed in if using free text search), the number of search results retrieved, and a unique session identifier. Given the login alias, the role of the user can be identified subsequently by using the hospital directory. All individual user information was deidentified after the user's role (eg, nurse, physician, or phlebotomist) was determined. User information was obtained in accordance with the hospital's online Web site privacy policy. Web use statistics were generated with the use of code provided by StatCounter (StatCounter.com, Dublin, Ireland).

### Analysis of Search Capabilities

We identified publicly available laboratory handbooks by searching with Google (http://www.google.com). We identified 5 reference laboratory Web sites and 15 large academic medical center Web sites for the study. We excluded laboratory manuals that we identified as incomplete or poorly functional or that were self-identified as "under construction."




**❚Figure 1❚** Front page of the Massachusetts General Hospital Online Laboratory Handbook (http://mghlabtest.partners.org).



DOI: 10.1309/PHAMDT8NN44WMBDF

### Data Analysis

Data analysis was performed using Microsoft Access and Excel (Microsoft).

## Results

### Survey of Laboratory Manuals

To understand how online laboratory manuals use search, we surveyed publicly available laboratory handbooks from a variety of institutions. The results of the survey are given in **Table 1**. Free text search (allowing the user to type words and then hit a "Find" button) was available in 9 (60%) of 15 academic medical center Web sites and all 5 of the reference laboratory Web sites. For the 14 Web sites that used free text search, we determined how the Web sites had implemented their search feature by searching for words present in the test name, test synonym, and/or other test-related fields for each test. From this analysis, we were able to determine that all 14 of the search engines operated by matching user input with text strings found in the test name, test synonyms, or other test related fields. Of the 14, 3 (21%) searched only the test name field; 10 (71%) searched the test name and test synonym fields; and 1 (7%) searched all test-related fields, including test description. The response to common misspellings (for example, "Willebrand" misspelled as "Willibrand" or "gentamicin" misspelled as "gentamycin") also was assessed. Of the 14 search engines, 12 (86%) retrieved no results when queried with misspelled words.

### Analysis of Laboratory Manual Use

In an effort to better understand the intentions of our users and to obtain information to facilitate rapid improvement of our search capability, we designed our laboratory handbook to capture information about users as they searched for laboratory tests. We launched our online laboratory manual in March 2005 with an awareness campaign for physicians, nurses, and laboratory staff. Since July 2005, we have had 150 to 250 unique users per day executing 300 to 500 searches per day. An analysis of 1,000 consecutive users in August 2005 demonstrated that nurses and nursing support staff represented 60% of our users **Figure 2**. Within the nursing group, registered

**Table 1**
**Search Characteristics of Online Laboratory Handbooks***

| | Academic Medical Centers | Reference Laboratories | Overall |
|---|---|---|---|
| Alphabetical listing of tests | 13/15 (87) | 4/5 (80) | 17/20 (85) |
| Free text search permitted | 9/15 (60) | 5/5 (100) | 14/20 (70) |
| Free text search method | | | |
| Test name only searched | 3/9 (33) | 0/5 (0) | 3/14 (21) |
| Test name and test synonyms searched | 6/9 (67) | 4/5 (80) | 10/14 (71) |
| All test related fields searched | 0/9 (0) | 1/5 (20) | 1/14 (7) |

* Data are given as number/total (percentage).

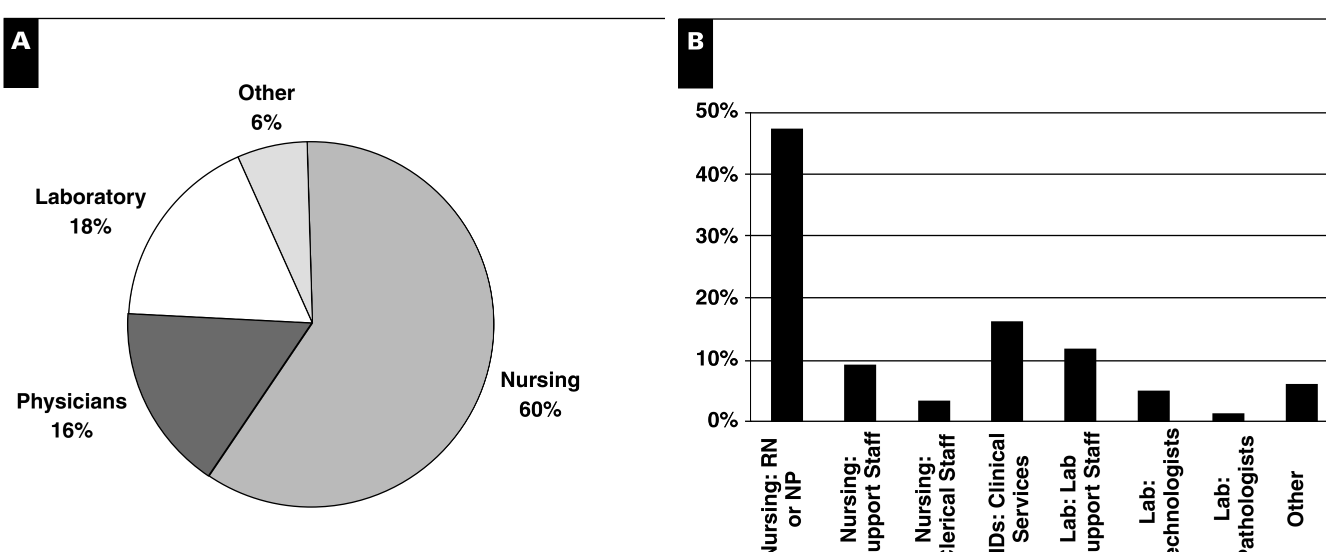



**Figure 2** Laboratory handbook users. Categories and roles for 1,000 consecutive laboratory handbook users were determined. **A**, Overall categories of users. **B**, Roles of users. NP, nurse practitioner; RN, registered nurse.


DOI: 10.1309/PHAMDT8NN44WMBDF

nurses and nurse practitioners were the heaviest consumers, together accounting for 79% of nursing users. Physicians on clinical services accounted for 16% of total searchers. Laboratory personnel represented 18% of the total users. The predominant laboratory users, comprising 65% of laboratory users, were members of our laboratory customer service call center, who used the handbook as a source of information for answering clinician test–related inquiries. Our call center also has had an important role in creating awareness of the handbook by informing callers of the existence of the online handbook.

We examined 5,000 consecutive searches in August 2005 to better understand how users were searching for laboratory tests. For all groups of users, the predominant type of search performed was free text searching, representing 87% of all searches, with the remaining users using the alphabetical listing of tests. This heavy bias toward free text searching may be due in part to free text searching being the default pathway for the handbook, whereas the alphabetical listing of tests is less prominently displayed (Figure 1). For each free text search, we examined the user-entered text. The majority (90.7%) of users typed 1 or 2 words in the search box **Table 2**.

### Analysis of Search Productivity

We analyzed search productivity by examining the number of results retrieved for a given free text search **Figure 3**. We analyzed 5,000 consecutive free text searches in July and August 2005. In July 2005, we found that 31% of searches were nonproductive, yielding zero search results. We identified the most prevalent reasons for a nonproductive search by analyzing the user-entered search terms for each of the 1,550 nonproductive free text searches **Table 3**. The most common reason (40%) for a nonproductive search was the inclusion of words that were not associated with a particular laboratory test. Common examples were searches for "serum GGT" or "celiac sprue." The word "serum" was not included in the key words associated with the GGT test, whereas the words "celiac" and "sprue" were not included in the key words associated with antiendomysial antibodies. The second most common reason for a search yielding zero results was a misspelled word (31%). Commonly misspelled words included ferritin, cytomegalovirus, and hepatitis.

### Use-Driven Updates Enhance Search Capability

On July 31, 2005, based on review and analysis of the 5,000 consecutive searches, we instituted changes to the handbook with the intent to decrease the percentage of nonproductive searches. Thus, additional key words were added for certain tests, including words that users had entered on prior nonproductive searches. We also added common misspellings that had been entered in prior nonproductive searches.

In August 2005, following the changes to the laboratory handbook key words, we again examined search productivity by analyzing 5,000 consecutive free text searches. Following the update, the number of nonproductive searches decreased from 31% to 18% ($P < .001$; $\chi^2$ analysis), whereas the number of searches returning 1 to 4 results increased (Figure 2). Because the decrease in nonproductive searches could be due to the confounding factor of increased user familiarity with the site, we performed all of the August 2005 searches using a July 2005 copy of the laboratory handbook database and a local copy of the Web site as it existed in July 2005, before the update. We found that the number of nonproductive searches was 29%, similar to the number of nonproductive searches in July 2005 (31%). This result suggested that the majority of the reduction in nonproductive searches was due to the key word updates. Thus, we concluded that our review and analysis of prior user searches had allowed us to systematically improve the search capabilities of the laboratory handbook.

**Table 2**
**Number of Words Entered per Search**

| No. of Words Entered by User | Percentage of Searches |
|---|---|
| 0 (all tests) | 3.4 |
| 1 | 69.0 |
| 2 | 21.7 |
| 3 | 5.3 |
| >3 | 0.7 |

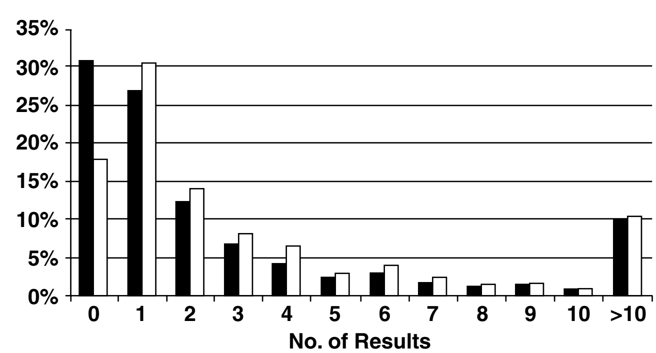



**Figure 3** Analysis of search productivity. The number of search results returned for 5,000 consecutive user searches was determined for periods in July (black bars) and August (white bars) 2005.

**Table 3**
**Reasons for Nonproductive (Zero Results) Searches**

| Reason | Percentage of Searches |
|---|---|
| Test synonym not mapped to test | 40 |
| Misspelled test | 31 |
| Nonexistent test | 19 |
| Other | 10 |




DOI: 10.1309/PHAMDT8NN44WMBDF

## Discussion

The sole product of a medical laboratory is information. The importance of improving access to this information cannot be overstated. Rapid and accurate access to laboratory test information is crucial to efficient workflow for health care providers and laboratory staff. Inaccurate information can lead to improper test selection or specimen collection that can compromise patient care. Failed attempts to find the desired information impede clinician workflow and impact the clinicians' perceptions of the laboratory. Moreover, if health care providers determine that online or printed materials are not easily accessible and accurate, calls to the laboratory will increase and negatively impact laboratory workflow.

Simply making the laboratory handbook available online can result in improved clinician access to laboratory information. Furthermore, the ability to provide immediate and frequent updates of an online laboratory handbook can ensure highly accurate information. Online manuals also provide an easy and effective way for users to report errors and provide other helpful feedback. Nevertheless, the creation of an online handbook does not automatically guarantee that users will be able to efficiently retrieve the information they seek. The increasing number and complexity of laboratory tests has made finding the correct test for a given clinical scenario challenging. Therefore, the information retrieval mechanisms of an online laboratory handbook need to be efficient, and the handbook must be intuitive to use.

Some of the laboratory manuals we surveyed relied on alphabetical lists of tests as the predominant or sole mode of information retrieval. In addition to being cumbersome to use and maintain, alphabetical lists cannot provide the efficient information retrieval that health care workers require. An alphabetical list of tests requires the user to conjecture how a given test is listed in the index. For example, a user looking for aspartate aminotransferase would not know whether to look in the alphabetical test list under "aspartate aminotransferase," "AST," "liver function tests," "SGOT," "serum glutamic oxaloacetic transaminase," or "transaminases."

The increasing use of search engines (>250 million searches per day at Google alone) has made search an integral part of the online experience. The sophistication of modern search engines has drastically simplified the process of finding information on the Internet. Online search has transformed the music, news, and entertainment industries.[6] The rapid progress in online search has been due to detailed cataloging of the Web and the development of algorithms designed to understand users' intent.[7] All search engines use "free text" search as the predominant mode of searching for information. In free text search, the user types in words and clicks the "Find" button, engaging the search engine and displaying links to Web pages that the search engine deems to be of interest to the user. The text that a user types before hitting the "Find" button has come to be recognized by search engine companies as a valuable commodity.[6] It is through analysis of this user-entered text that the users' intentions can be understood. By storing and analyzing user searches, it is possible to generate a "database of intentions," thus arriving at an understanding of how and why users interact with a search engine.[6] This information has permitted search engine companies to make rapid advances and dramatically improve search engine efficiency and quality.

For an online laboratory handbook, a free text search function may provide a flexible and potentially efficient interface, but it does not guarantee quality information retrieval. There are numerous approaches to implementing free text search for a laboratory handbook, and the selected method has major implications for the quality of information retrieval. The simplest approach is to search the test name for the user-inputted text. In our survey of laboratory handbook Web sites that have a free text search feature, 21% (3/14) used matching on the test name only. This approach leads to some major quality issues and errors of omission. For example, in one handbook, this approach resulted in searches for "cytomegalovirus" not finding the test for cytomegalovirus antigen because the test name in the handbook was written as "CMV antigen."

Another approach to free text searching is to look for matches within all of the fields associated with each laboratory test (eg, test name, test description, performing laboratory, test requirements, and other test-related fields). This approach is likely to lead to errors of commission, in which the results returned to the user include many unwanted tests in addition to the test of interest. For example, searching a manual for "liver," and expecting to get a short list of tests relevant to liver disease, a user would instead get a long list of tests, many unrelated to liver disease testing, because the word "liver" is a common word included in the description of many unrelated tests.

Associating each test with specific key words is a common approach to improving free text searching. In this approach, each test is associated with a series of terms (key words) synonymous with the test. For example, the key words for the test "serum potassium" could include "K," "potassium," "serum," "electrolytes," and "lytes." Searching a sparsely keyed database can lead to errors of omission in which appropriate tests are not included in the results. For example, if the key words for the AST test do not include "SGOT" (a synonym for the AST test), this test will not be found when the user types in "SGOT." Another aspect of key word–driven free text searching is that it introduces the potential for nonproductive search results owing to user misspellings and use of nonstandard abbreviations (eg, using the query "LFTs" to retrieve the AST and ALT tests). A search engine can be made more


DOI: 10.1309/PHAMDT8NN44WMBDF

robust by adding key words that correspond to such abbreviations and common misspellings. However, owing to the increasing number and sophistication of clinical laboratory tests and the variety of users who interact with a laboratory handbook, the a priori identification of all appropriate key words for a given test is virtually impossible. Thus, a system is needed to improve the identification and selection of test key words.

In this report, we show that ongoing improvement in the quality of a laboratory handbook can be accomplished by monitoring user queries. Each user query provides data that can be used to improve the search engine's accuracy. The first attempt at identifying key words, before clinical use of the handbook, likely will result in relatively sparse indexing. We demonstrate an approach to improving key word–driven search by regular monitoring of user search data followed by updates to the test-associated key words. Such an active update process requires easy access to the laboratory handbook database for monitoring and updates. To facilitate the monitoring and updating process, we link each of our SQL Server database tables to Microsoft Access via a standard database access method (Open Database Connectivity, or ODBC) to permit simplified updating of test information by nontechnical staff.

Recommendations for enhancing the search capability of an existing online laboratory handbook include making free text search the most prominent pathway, monitoring user searches, tracking the percentage of unsuccessful searches (searches with the number of results equal to zero), and using search monitoring data as the basis for iterative updates to test key words and search algorithms. These enhancements to the laboratory handbook can result in rapid improvement in the quality of an online handbook.

An additional component of online handbook development that may easily be overlooked is "marketing" of the handbook to users. An important strategy for increasing clinician awareness and use of the laboratory handbook is having links to the handbook from clinician-centric applications such as online hospital Web sites, provider order entry, and clinical information systems. By making the online handbook available in the daily workflow of clinicians, it can be accessed in a "just in time" manner to inform and improve clinical decision making.[8] In addition to awareness campaigns, it is equally important that detailed Web site use statistics are kept. This enables the laboratory to understand who is actually using the handbook and where they are coming from (the referring page or application), both of which are useful in the assessment of marketing effectiveness. Moreover, the tracking of additional user settings has proven useful in the design of our site. For example, we observed that 85% of our users had their monitor screen resolution set to 800 × 600 pixels, with the remaining users having a higher resolution. Knowledge of this parameter has been important for page layout and other design considerations for the Web site.

The information needs of health care professionals vary widely. A physician looking for the testing options for "celiac disease" may want to see a very different set of results and information than a phlebotomist looking to find the specimen requirements for the antiendomysial antibody test, despite the fact that these clinicians essentially may be looking for the same test. The present iteration of our key word–based search engine is unable to provide such context specificity. The ideal search engine would use readily available contextual data to determine the most appropriate search results. Knowledge of who the user is and why the user is searching (whether for specimen requirements, to decide on the value of a test, or to find out what testing options are available) could be used to change the output of the search engine to better match the context. In striving to achieve this, we must first understand how search intentions differ across populations of health care workers. In the online world, we can capture information about users based on their network login, their physical location, where they came from (the referring page or application), their personal history of searches, and their "clickstream" (the exact sequence of pages that a user views as he or she visits a site). We then can use these additional attributes to develop an understanding of these patterns of intention and customize search results accordingly.

As more sophisticated search techniques begin to provide health care workers with flexible and robust access to testing information, we envision the laboratory handbook becoming an intuitive portal for clinician inquiries to the laboratory. Clinician queries typically are test-centric. Questions may be generated when the test is ordered, when results are viewed initially, or when results are interpreted. Furthermore, tests can be linked to department policies, individual pathologists who may provide interpretive advice, laboratory guidelines, or external Web sites that may provide additional information. Linkage to the handbook can occur directly from computerized physician order entry applications, hospital and outpatient clinical information systems, and wireless handheld devices. A well-trafficked online laboratory handbook provides an efficient channel for disseminating laboratory information to clinical staff. This can assist in communicating such important issues as changes to test reference ranges, tube recalls, or laboratory policy updates. The laboratory handbook database also is well positioned to evolve into a clearinghouse for all types of laboratory information that may be useful to a wide variety of clinical applications, including computerized provider order entry, clinical information systems, and clinical decision support systems. Web services, software platforms that connect applications, can be programmed to facilitate data exchange between a laboratory handbook application and other clinical applications.[9]



DOI: 10.1309/PHAMDT8NN44WMBDF

Despite the increasing prevalence of search in our daily lives, search technology and user interfaces remain in their early developmental stages. It has only been during the past 5 years that robust search engines have been available to search the Internet. It will be essential for customer-driven laboratories to keep up with advances in search methods and interfaces, as they will undoubtedly provide improved ways for distributing information to our customers.

*From [1]Decision Systems Group, Brigham & Women's Hospital and Harvard-MIT Division of Health Sciences & Technology, and [2]Department of Pathology, Massachusetts General Hospital, Boston, MA.*

*Dr Blechner is a research fellow in biomedical informatics, supported in part by training grant T15 LM007092 from the National Library of Medicine, Bethesda, MD.*

*Address reprint requests to Dr Dighe: Massachusetts General Hospital, Bigelow Bldg, Room 510, 55 Fruit St, Boston, MA 02114.*

## References

1. Hardy R, Grubbs FC. Revised laboratory manual helps solve problems. *Am J Med Technol*. 1976;42:441-444.
2. Croft SE, Koehler MA. New concept: a lab manual that withstands changing times. *Am J Med Technol*. 1978;44:1090-1092.
3. Lazinger B, Steif J, Granit E. An online tests catalog for clinical laboratories. *J Med Syst*. 1989;13:187-192.
4. Fine JS, Ching AK, Schneider JB, et al. BibleCard: network-based virtual database for laboratory information. *Clin Chem*. 1995;41:1349-1353.
5. Beckwith B, Schwartz R, Pantanowitz L. Analysis of on-line clinical laboratory manuals and practical recommendations. *Arch Pathol Lab Med*. 2004;128:476-479.
6. Battelle J. *The Search: How Google and Its Rivals Rewrote the Rules of Business and Transformed Our Culture*. New York, NY: Penguin Group; 2005.
7. Brin S, Page L. The anatomy of a large-scale hypertextual Web search engine. *Comput Netw ISDN Syst*. 1998;30:107-117.
8. Davenport TH, Glaser J. Just-in-time delivery comes to knowledge management. *Harvard Bus Rev*. 2002;80:107-111, 126.
9. Saadawi G, Harrison JH Jr. XML syntax for clinical laboratory procedure manuals [abstract]. *AMIA Annu Symp Proc*. 2003:993.


DOI: 10.1309/PHAMDT8NN44WMBDF