Developing and using a rubric for evaluating evidence-based medicine point-of-care tools

Suzanne Shurtz, MLIS, AHIP; Margaret J. Foster, MS, MPH

See end of article for authors’ affiliations. DOI: 10.3163/1536-5050.99.3.012

Objective: The research sought to establish a rubric for evaluating evidence-based medicine (EBM) point-of-care tools in a health sciences library.

Methods: The authors searched the literature for EBM tool evaluations and found that most previous reviews were designed to evaluate the ability of an EBM tool to answer a clinical question. The researchers’ goal was to develop and complete rubrics for assessing these tools based on criteria for a general evaluation of tools (reviewing content, search options, quality control, and grading) and criteria for an evaluation of clinical summaries (searching tools for treatments of common diagnoses and evaluating summaries for quality control).

Results: Differences between EBM tools’ options, content coverage, and usability were minimal. However, the products’ methods for locating and grading evidence varied widely in transparency and process.

Conclusions: As EBM tools are constantly updating and evolving, evaluation of these tools needs to be conducted frequently. Standards for evaluating EBM tools need to be established, with one method being the use of objective rubrics. In addition, EBM tools need to provide more information about authorship, reviewers, methods for evidence collection, and grading system employed.

INTRODUCTION

Health care providers require quick retrieval of information to efficiently answer questions related to patient care [1]. Evidence-based medicine (EBM) tools, designed for use at the point of care (POC), meet these needs by providing high-quality and synthesized information at the patient’s “bedside” [2]. To support clinical information needs, medical and hospital libraries may wish to subscribe to multiple POC tools, owing to the variety in topics covered, the options provided, and the audiences targeted by each tool. To avoid redundancy and to offer the best products for their users, libraries need to periodically reassess available EBM POC resources [2]. What criteria should librarians utilize to evaluate these POC tools? Because these resources are used to make clinical decisions, resources providing the best evidence relevant to patient care would be most desirable. Those who use these tools “need to know how much confidence they can place in the recommendations” [3]. One role of medical librarians is to evaluate the quality of EBM resources so that users can have confidence in the information upon which they base clinical decisions. This is a weighty responsibility, and librarians have an important challenge determining which EBM POC resources to make available to users.

BACKGROUND

The Texas A&M University Medical Sciences Library (MSL) is 1 of 5 libraries on the College Station campus of Texas A&M University. MSL houses 120,000 print titles and 1,600 serials, with a collection budget of more than $1.8 million. One of the library’s user groups is the Texas A&M Health Science Center (TAMHSC), which includes the TAMHSC College of Medicine, College of Nursing, School of Graduate Studies, Rangel College of Pharmacy, and School of Rural Public Health. The director of the MSL charged a team of 3 librarians in 2008 to evaluate and determine which EBM POC subscription-based tools the library should keep or purchase, in order to reduce overlap and to ensure that user needs were met. In 2008, MSL subscribed to 5 EBM POC tools: Stat!Ref’s ACP PIER, DynaMed, Cochrane Library, Essential Evidence Plus, and MD Consult’s First Consult. The librarians were asked to evaluate all...
subscription EBM POC resources. No directions were given as to how to perform this evaluation. To determine which tools to evaluate and to find best practices for evaluation, the librarians conducted a literature review.

**LITERATURE REVIEW**

The literature review included searching MEDLINE (PubMed), CINAHL (EBSCO), Library Information Science Technology Abstracts (LISTA) (EBSCO), and Google Scholar. Search terms included combinations of "evidence based medicine tools," "point of care tools," "bedside tools," "evaluation," and names of specific tools. The researchers were particularly interested in studies reviewing EBM tools to which the library currently subscribed. UpToDate, Epocrates, and Ovid’s Clin-eguide were tools in which users had expressed interest and were also used as search terms. Articles excluded from the search were opinion papers, either letters to editors or journal columns.

Fourteen research-based articles were found evaluating specific features of EBM POC tools for clinical medicine (Table 1). Found studies focused mainly on EBM tools that physicians use. One article was found on tools for clinical laboratory specialists [4]. The current status of each tool mentioned in the articles was verified. Some tools’ names had been changed since the article was written, such as InfoPOEMs/InfoRetriever, which is now Essential Evidence Plus [2, 5–7]. Other products included in studies no longer exist, such as MAXX [8]. Negative reviews of the tools were often later resolved. For instance, a 2003 review of UpToDate found the information for patients weak [9], and UpToDate has since created a version specific for patients [10]. In another example, DynaMed was criticized for not clearly grading its evidence [11], and DynaMed producers responded by creating a description of the levels of evidence used in the product [12].

Retrieved articles varied by who acted as the tool reviewer (the researcher versus the user) and by the platform used during the evaluation (computer versus mobile device). Nearly 80% of the articles included 1 or more clinicians (MDs) as reviewers. Backgrounds of authors varied among medical librarians, practicing physicians, clinical faculty, medical informaticians, graduate students, and research coordinators. The evaluations also took place in a wide range of locations including hospitals, private practice, laboratories, and libraries. Some authors evaluated both free tools, such as Google, and subscription-based tool, such as DynaMed [2, 7, 8, 13–19]. Differences were evident in how EBM tools were defined and compared. One study divided tools into two categories [5]. The first category, "evidence-based resources," included bibliographic databases and summary resources. Resources such as PubMed, Cochrane Database of Systematic Reviews, First Consult, InfoPoems, and DynaMed were referenced in this group. The second group, defined as "review resources," incorporated "evidence-based content" without an "explicit process to ensure that all relevant literature is selected and appraised" [5]. Examples included AccessMedicine, MD Consult, Stat!Ref, and UpToDate [5]. Other evaluations did not differentiate between bibliographic versus synthesized tools, comparing resources like PubMed, MEDLINE, and Cochrane to UpToDate [5, 7, 15, 17–19]. Another evaluation defined three EBM tools as "e-textbooks" [20].

Most evaluation metrics in these studies were designed to determine whether answers to clinical questions were retrieved in the tools. Either preselected questions or those formulated at the point of care were used in the evaluations. Methods varied as to how preselected questions were gathered. Some studies additionally factored in the average search time required to locate the answer. One article used a scoring system created by the authors to assess overall content in each tool.

Due to the wide variation in the studies’ researchers, methods, and evaluated tools, comparisons between studies and their respective recommendations proved difficult. Overall conclusions varied. Some reviews rated no product higher than another [18, 20], multiple studies rated UpToDate highest [2, 9, 14, 16, 19], and one rated DISEASEDEX highest [6]. UpToDate’s easy-to-use interface was cited as a reason for its popularity. Campbell acknowledged that participants in her study had used UpToDate previously and therefore were already familiar with the search interface [2]. Kim explained that "bibliographic databases such as PubMed and Medline were] difficult to navigate, and current evidence-based synopses or summary sources [had] small databases that covered a limited content" [5]. Two of the studies that found DynaMed valuable included the product’s creator as one of the authors [8, 13].

**OBJECTIVES**

While most previous evaluations of EBM tools focused on health professionals’ and researchers’ perspectives on the ability of a tool to answer a clinical question, the librarians in this study did not feel qualified to determine if a tool sufficiently answered a clinical question. Also, the researchers wanted to ultimately develop a process with which they could regularly evaluate the tools in their collection. Due to these factors, the goal of this study was to assess the attributes, the standards and clarity of evidence grading, and the level of transparency (authorship or review processes) of the tools. The following objectives were identified to achieve these goals:

- to create a list of subscription-based resources claiming to be EBM POC
- to develop and use an evaluation form to systematically compare the features of these resources
- to methodically evaluate the content and quality control of these resources using commonly searched clinical topics
- to make recommendations about subscriptions to these resources based on these evaluations
<table>
<thead>
<tr>
<th>Article</th>
<th>Tools evaluated</th>
<th>Reviewer academic qualifications</th>
<th>Evaluation methods</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alper et al., 2001* [8]</td>
<td>StatRef, MD Consult, DynaMed, MAXX†, MDChoice.com</td>
<td>MDs</td>
<td>20 preselected clinical questions searched by 2 family medicine clinicians</td>
<td>StatRef, DynaMed, MAXX, MD Consult answered 95% of questions</td>
</tr>
<tr>
<td>Alper et al., 2005* [13]</td>
<td>Dynamed</td>
<td>MDs, medical researcher</td>
<td>698 clinical questions at point of care searched by 52 clinicians</td>
<td>DynaMed answered more questions than other sources</td>
</tr>
<tr>
<td>Banzi et al., 2010 [18]</td>
<td>5-min Clinical Consult, ACP PIER, BestBETs, CKS, Clinical Evidence, DynaMed, eMedicine, tEG complete, EBM Guidelines, First Consult, GP Notebook, Harrison's Practice, Health Gate, Map of Medicine, Micromedex, Pepid, UpToDate, ZynxEvidence</td>
<td>Pharmacological researchers</td>
<td>Scoring system used by authors to evaluate overall content</td>
<td>No product rated highest</td>
</tr>
<tr>
<td>Campbell and Ash, 2006‡ [2]</td>
<td>ACP's PIER, DISEASEDEX, FIRST Consult, InfoRetriever (Essential Evidence Plus), UpToDate</td>
<td>Medical librarian, PhD</td>
<td>Preselected clinical questions searched by 8 clinicians, 3 pharmacists, 5 medical informatics students, 1 MRI technologist, 1 registered nurse</td>
<td>UpToDate answered more questions than other sources and preferred for usability</td>
</tr>
<tr>
<td>Fenton and Badgett, 2007 [16]</td>
<td>National Guideline Clearinghouse, UpToDate</td>
<td>MD, PhD</td>
<td>752 preselected NLM clinical questions searched by authors</td>
<td>UpToDate answered more questions than National Guideline Clearinghouse</td>
</tr>
<tr>
<td>Fox and Moawad, 2003 [9]</td>
<td>UpToDate</td>
<td>MDs</td>
<td>Clinical questions at point of care searched by authors</td>
<td>UpToDate answered most questions quickly</td>
</tr>
<tr>
<td>Goodyear-Smith et al., 2008 [20]</td>
<td>DynaMed, MD Consult, UpToDate</td>
<td>PhDs</td>
<td>4 preselected clinical questions searched by 200 clinicians</td>
<td>No product rated highest</td>
</tr>
<tr>
<td>Hoogendam et al., 2008 [19]</td>
<td>UpToDate, PubMed</td>
<td>MDs</td>
<td>2,986 clinical questions at point of care searched by 40 residents, 30 internists</td>
<td>UpToDate answered more questions than PubMed</td>
</tr>
<tr>
<td>Jones and Emery, 2004 [6]</td>
<td>DISEASEDEX, DynaMed, UpToDate, InfoRetriever, First Consult, Clin-eguide, DynaMed, eMedicine, StatRef</td>
<td></td>
<td>Preselected clinical questions searched by authors</td>
<td>DISEASEDEX rated highest</td>
</tr>
<tr>
<td>Kim et al., 2008 [5]</td>
<td>DynaMed, InfoPoems, UpToDate, MD Consult, eMedicine, StatRef</td>
<td>MDs, medical librarian, PhD</td>
<td>Clinical vignettes in pre- and posttests searched in EBM class of 50 residents</td>
<td>Ovid resources** and InfoPoems used most by EBM class; UpToDate, MD Consult used most by control group</td>
</tr>
<tr>
<td>Koonce et al., 2004 [17]</td>
<td>UpToDate, Cochrane, HEALTHGate, Clinical Guidelines</td>
<td>Librarians, MD</td>
<td>40 complex clinical questions (randomly selected); 40 general care clinical questions (selected)</td>
<td>Resources completely answered 20% of complex questions and 47% of general care used most</td>
</tr>
<tr>
<td>McCord et al., 2007 [14]</td>
<td>UpToDate, Epocrates, MD Consult, DynaMed, eMedicine</td>
<td>MDs, medical students, research coordinator</td>
<td>532 clinical questions at point of care searched by 25 family medicine residents and 37 clinical faculty</td>
<td>UpToDate rated highest, but Epocrates used most</td>
</tr>
<tr>
<td>Schilling et al., 2005 [15]</td>
<td>MEDLINE, UpToDate, MD Consult, Cochrane Library, National Guideline Clearinghouse, Micromedex</td>
<td>MDs, MBA</td>
<td>158 clinical questions at point of care searched by 43 internal medicine residents</td>
<td>MEDLINE and UpToDate used most</td>
</tr>
<tr>
<td>Schwartz et al., 2003 [7]</td>
<td>TRIP, InfoRetriever</td>
<td>MDs, PhD/MPH</td>
<td>92 clinical questions at point of care searched by 3 clinical faculty</td>
<td>TRIP database used to answer most questions</td>
</tr>
</tbody>
</table>

* Founder of Dynamed.
† Tool no longer exists.
‡ Tools evaluated were only those in which the library was interested.
§ Tools evaluated were only those with which the users were familiar.
** Ovid resources included MEDLINE, ACP Journal Club, Cochrane, and DARE.
METHODS

Tool selection process

Screening criteria. In the spring of 2009, a list of EBM tools was created from those found in the literature review, from tools that patrons had requested, and from tools to which Texas A&M University subscribed. As illustrated in the literature review, definitions for and categories of EBM tools varied. Tools that were subscription based, claimed to be evidence based, claimed to be POC, and focused on clinical content were chosen for the review. Because the original charge from the library director was to compare subscription-based products, free resources were excluded. Evidence of the first three criteria could be found on the product’s website or brochures. Emphasis on clinical content was determined by reading product descriptions on the site. A sample of academic medical and hospital libraries’ holdings was also reviewed to ensure tools were not missed and to compare the researchers’ holdings with that of the sample. The South Central Academic Medical Libraries consortium (SCAMEL), of which MSL is a member, was selected as the sample. Libraries in the consortium vary in size and users or programs supported, providing a spectrum of medical libraries appropriate for the sample that the researchers desired. Each SCAMEL library’s website was searched, based on the screening criteria, to find a listing of EBM tools available to their users. No additional tools were found when the SCAMEL holdings were compiled.

The tools found were then entered into a form made using Zoho Creator, a web-based software product that creates databases with forms and tables [21]. Each tool was screened by at least two reviewers, using an online form created in Zoho Creator [21]. Three screeners discussed discrepancies before coming to a consensus about tools to exclude. This form was completed in 2009 and again in 2010 to show changes in subscriptions.

Tools selected for evaluation. Based on the screening criteria, the researchers chose to include the following tools in their evaluation: ACP PIER, BMJ Point of Care, Clin-eguide, DynaMed, Epocrates, Essential Evidence Plus, First Consult, and UpToDate. The authors chose to exclude any tool that was strictly a pharmaceutical resource, as the target for this evaluation was clinical medicine.

Evaluation

Each tool was evaluated in two ways: by rating common attributes on a general evaluation form and by determining the tool’s ability to provide a treatment summary on commonly diagnosed conditions.

Part 1: general evaluation. The general evaluation form, produced in Zoho Creator, was developed to analyze all parts of the resource and initially included four sections: general characteristics, scope and content, search options, and results presentation. The evaluation form questions were adapted from a presentation by Trumble in 2004 (in which one of the authors participated) [22]. As the evaluation progressed, it became apparent that an important difference between these tools was how the evidence was graded. Evidence that has been compiled and evaluated by authoritative bodies, or in other words, “graded,” has the most weight in EBM [2, 23]. Because the reviewers wanted to examine more closely how the evidence was graded, a fifth section was added to the form.

The final evaluation form included the following sections: Section A included measures of general characteristics of the tool, including availability of platform compatibility, simultaneous users, mobile devices accommodated, free version, and available customizations. Section B parameters were the scope and content, including availability of patient handouts, continuing education credits, practice guidelines, and help screens. Section C covered the search options (browsing, guided searches for pharmaceuticals, advanced), navigation, and presentation of the results (readability, adjustments, print options, and other outputs). Section D focused on quality control, including listing of authors or reviewers’ names, credentials, and affiliations; the process of becoming an author or reviewer; updating schedule; and any biases (affiliations were considered to constitute a conflict of interest). Section E included the standards used for grading and the clarity of the grading levels.

Part 2: clinical summaries evaluation. During the general evaluation, the researchers also noted that the consistency with which the evidence summaries were graded varied between tools. In summer of 2010, the investigators wanted a snapshot of the extent to which grades were provided for common clinical topics in each tool. A list of common primary diagnoses was located in the Centers for Disease Control and Prevention’s “National Ambulatory Medical Care Survey: 2005 Summary” [24]. The investigators selected the top ten diagnoses to search across each EBM tool. For those terms that were very broad (such as malignant neoplasms), librarians selected the most common form of the disease (such as skin cancer). The final list of ten terms searched was: hypertension, upper respiratory infection, arthritis, skin cancer, diabetes, low back pain, ear infection, asthma, high cholesterol, and sinusitis.

A Zoho Creator form was made for each clinical topic addressing the following criteria: Was a section on the topic provided? Was the evidence graded? Was it summarized? Were summary tables provided? For each question, a yes or no option was listed, as well as the number of clicks to find the answer, and a notes field. The number of clicks was counted from the main page to the appropriate section of the results page. Two librarians searched each term in every tool.
RESULTS

Part 1: general evaluation

Table 2 summarizes the results of the general evaluation. In terms of platforms available for tools, all but two tools were compatible with PC and Mac computers. Several tools mentioned mobile availability, with DynaMed listing the most mobile platforms. Customizations for all tools were limited to changing the font size and customizing patient handouts. Subscription options were fairly consistent across tools, with only ACP PIER using a charge-per-seat model, while others had switched to unlimited seat access. All evaluated tools included some practice guidelines in their summaries. Six offered some continuing education credits, mostly to physicians. Little variation was seen in search options and presentation of results across tools. However, one obvious difference noted in the search for pharmaceuticals was that some tools offered additional options, such as pill identification, indications, and counter indications.

The quality control of the evidence in the resources was of particular interest to investigators. Most tools listed the authors of individual summaries. DynaMed and UpToDate specified the process to become an author. Most of the tools provided the authors’ credentials and affiliations. Two resources, DynaMed and Essential Evidence Plus, explained the process to become an editor, and five of the resources (BMJ Point of Care, DynaMed, Epocrates, First Consult, and UpToDate) listed the editors’ names and credentials.

Six of the tools claimed to update their summaries in six months or less. The standard of grading was described in five of the tools: One used Strength of Recommendation Taxonomy (SORT) [25]; two used Grading of Recommendations Assessment, Development and Evaluation (GRADE) [23]; and two created their own grading systems. Three of the tools did not specify a grading process or standard.

Part 2: clinical summaries evaluation

Table 3 summarizes the results of searching for 10 specific diagnoses treatment summaries in each resource with the exception of BMJ Point of Care, which provides the disease content for Epocrates. Rather than evaluating the same content twice, the investigators chose to search the content solely in Epocrates. A total of 70 treatment summaries were retrieved, 10 for each of the 7 tools. Evidence supporting possible treatments was often summarized in narrative form (average 8.1 times per tool), but not often in table form (average 2.3 times per tool). Transparency in authorship varied, with 5 out of 7 listing individual authors, 4 providing credentials, and 4 providing affiliations. One bias was noted while reviewing the affiliations, with a regional director for Merck listed as an author for ACP PIER. The average number of authors listed in a summary was 1.6 for those 5 tools that listed authors. Fifty-two of the 70 retrieved searches (74%) graded the evidence. DynaMed had the most up-to-date summaries (updated on average within 19 days), while First Consult had...
the least up to date (updated on average within 449 days). Six tools claimed to update summaries within 6 months or less. For the 10 topics searched, however, only DynaMed met this claim.

**DISCUSSION**

**Principal findings**

While some articles in the literature review favored one EBM tool, the major finding of this review was that current EBM POC tools were quite similar relative to their content coverage and search options. Reporting of quality control measures was where researchers noted the greatest difference between tools. Level of transparency of authorship, evidence selection process, and grading standards varied greatly between tools. Transparency in processes is vital to evaluating reliability and completeness to determine possible biases in the evidence. This need for transparency in evidence summaries has been expressed by those developing standards for practice guidelines and systematic reviews [23, 26].

**Evaluation of rubrics**

Two rubrics for the review were created: one to evaluate tool characteristics and another to assess the evidence presented in treatment summaries. The Zoho Creator software provided a user friendly and efficient way to develop and complete the rubrics. Overall, the general evaluation rubric was easy to develop due to the resource assessment skill sets common to most librarians. The elements in sections A to C were quickly agreed upon and the information for them found. However, elements in sections D and E were much more difficult to define and locate. The reviewers needed to contact customer service providers in some cases to receive clarification on grading standards. The second rubric, clinical summaries evaluation, was more challenging to create. Determining the clinical topics to search required some research and discussion. Some measures included in the rubric were ultimately omitted from the final assessment. For instance, because the number of clicks was difficult to standardize between screeners, the number of clicks to find the treatment section and to find the evidence grading were omitted from the findings.

**Strengths and weaknesses**

Limitations of the study lie in what was absent from the review. Understanding the factors affecting SCAMEl institutions’ selection of EBM POC tools could have been valuable. Factors such as product cost, budget, and user input might have impacted the selection process more than attributes of the tools themselves. Another omission from this review was free EBM tools. If assessed, these tools might have changed the findings of the study. Lastly, one element that should have been included in the tool evaluation was an analysis of potential bias based on editor affiliations. Only the affiliations of the evidence summary authors were considered.

The span of time, eighteen months, covered by this review was both a strength and a limitation. Researchers had an opportunity to observe how quickly EBM POC tools can change and to see trends in the SCAMEl collections and in tool development. Due to product changes during this time frame, both the literature review and information about the tools were updated multiple times.

**Strengths of this review** included systematic evaluation of tools’ characteristics and quality control processes, as well as in-depth assessment of grading systems. By focusing on those elements that affect the reliability and value of the synthesized evidence, the reviewers highlighted the need for EBM POC tools to be more transparent.

**Recommendations**

The rubrics developed in this study can be used by other librarians to evaluate EBM tools. Because the rubrics themselves do not weigh one element over another, the importance of each element could be adapted according to the priorities and needs of specific libraries. As products change often, these evaluations need to be conducted on a regular basis,
Developing a rubric for point-of-care tools

Christina Seeger, AHIP, previously at Texas A&M’s MSL, now the pharmacy librarian at Feik School of Pharmacy, University of the Incarnate Word, assisted with this research project.

REFERENCES


19. Hoogendam A, Stalenhoef AF, Robbe PF, Overbeke AJ. Answers to questions posed during daily patient care are more likely to be answered by UpToDate than PubMed. J Med Internet Res. 2008 Oct 3;10(4):e29. DOI: 10.2196/jmir.1012.


to discover the similarities and differences in tools and to make more informed subscription decisions. For example, in this case, the reviewers’ recommendation to the library director was to drop at least one of the currently subscribed tools due to redundancy.

Future developments

One further area of investigation is to determine how clinicians and students perceive the grading of evidence. How important is it to EBM POC users that the evidence is graded, and how well do they understand the grading scales? If the meaning of the grade is unclear, users may be hesitant to apply the recommendations. What are the mental processes by which users apply the evidence found in EBM POC tools? Focus groups may be helpful to determine users’ preferences of tools and impressions of evidence grading.

CONCLUSIONS

EBM POC tools remain an important resource for health care practitioners, librarians, and patients. Due to the dynamic nature of these tools, evaluation needs to be continuous. Many tools have similar audiences and content. The grading of the evidence, however, varies among resources and requires standards for evaluation. Tool creators need to more openly communicate guidelines for authorship, reviewers, methods of evidence collection, and grading systems. Medical librarians play an important role in evaluating the quality of POC tools and can develop assessment tools, such as rubrics, to aid in this evaluation.

ACKNOWLEDGMENTS

Christina Seeger, AHIP, previously at Texas A&M’s MSL, now the pharmacy librarian at Feik School of Pharmacy, University of the Incarnate Word, assisted with this research project.


AUTHORS’ AFFILIATIONS

Suzanne Shurtz, MLIS, AHIP (corresponding author), sshurtz@library.tamu.edu, Instructional Service Librarian, Medical Sciences Library, Texas A&M University, 4462 TAMU College Station, TX 77843-4462; Margaret J. Foster, MS, MPH, margaretfoster@library.tamu.edu, Social Sciences and Education Librarian, Sterling C. Evans Library, Texas A&M University, 5000 TAMU College Station, TX 77843-5000

Received October 2010; accepted January 2011