Published in final edited form as:

Med Ref Serv Q. 2014 ; 33(3): 241–252. doi:10.1080/02763869.2014.925379.

The Use of Research Electronic Data Capture (REDCap) Software to Create a Database of Librarian-Mediated Literature Searches

JENNIFER A. LYON, ROLANDO GARCIA-MILIAN, and HANNAH F. NORTON

University of Florida Health Science Center Libraries, Gainesville, Florida, USA

MICHELE R. TENNANT

University of Florida Health Science Center Libraries; and University of Florida Genetics Institute, Gainesville, Florida, USA

Abstract

Expert-mediated literature searching, a keystone service in biomedical librarianship, would benefit significantly from regular methodical review. This paper describes the novel use of Research Electronic Data Capture (REDCap) software to create a database of literature searches conducted at a large academic health sciences library. An archive of paper search requests was entered into REDCap, and librarians now prospectively enter records for current searches. Having search data readily available allows librarians to reuse search strategies and track their workload. In aggregate, this data can help guide practice and determine priorities by identifying users' needs, tracking librarian effort, and focusing librarians' continuing education.

Keywords

Data capture; evidence-based librarianship; expert literature searching; mediated literature searching; REDCap

INTRODUCTION

Expert searching, the process in which a user with an information need seeks consultation and assistance from an acknowledged expert, is a recognized role of health sciences librarians^{1,2} and is offered by many medical libraries. However, multiple barriers to effective provision of this service exist, including the lack of objective measures of quality in expert searching education and practice.³ One approach to providing the data and

Address correspondence to Jennifer A. Lyon, University of Florida Health Science Center Libraries, P.O. Box 100206, Gainesville, FL 32610. jalyon@ufl.edu.

ABOUT THE AUTHORS

Jennifer A. Lyon, MS, MLIS, AHIP (jalyon@ufl.edu) is Clinical Research Librarian Rolando Garcia-Milian, MLIS, AHIP (rolando.milian@ufl.edu) is Basic Biomedical Sciences Librarian Hannah F. Norton, MSIS, AHIP (nortonh@ufl.edu) is Reference and Liaison Librarian; all at University of Florida Health Science Center Libraries, P.O. Box 100206, Gainesville, FL 32610. Michele R. Tennant, PhD, MLIS, AHIP (tennantm@ufl.edu) is Assistant Director, University of Florida Health Science Center Libraries, P.O. Box 100206, Gainesville, FL 32610-0206 and Bioinformatics Librarian, University of Florida Genetics Institute, P.O. Box 103610, Gainesville, FL 32610.

This article is based on a poster presented at the Medical Library Association Annual Meeting, Seattle, WA, May 20, 2012.

evidence required for quality improvement in expert searching is the collection and analysis of search strategies and related data, including requester demographic information. Using an electronic tool to gather these data can facilitate search practice analysis, workload tracking, and foster expertise-sharing and quality improvement⁴ and may lead to the development of a more rigorous approach to conducting complex literature searching.

Electronic data collection has many advantages to paper, as it allows the material to be re-used and re-purposed as well as easily analyzed and reported.^{4,5} Librarians who wish to view or count the searches conducted over a specific time period can run a report in moments. Searching patterns by groups of librarians or patron categories can also be identified rapidly and examined, both for internal library use and for communicating to institutional leadership about librarians' work in relation to institutional mission and goals.⁶ Previous data can be located and then re-used for a multitude of purposes including adaptation for identical or similar patron questions, teaching and sharing information among librarians, examples for patron instruction, and determining the most commonly asked questions by a specific client group.

The University of Florida's (UF) Health Science Center Library (HSCL) provides expert searching services to its clientele. Although each librarian traditionally maintained individual statistics on the number of searches completed for annual reporting, there was no established procedure to digitally record and analyze complex literature searches and mediated results. Search histories and demographic information were stored on paper forms, making it impossible to pull reports of interest (requesting individual or department, search topics/strategies, hours spent per search), and cumbersome to locate a specific strategy for training purposes.

In spring 2010, the UF's Clinical and Translational Science Institute (CTSI) joined the REDCap Consortium,⁷ which supports the Research Electronic Data Capture (REDCap) web application. Developed at Vanderbilt University's Institute for Clinical and Translational Research, REDCap is a web-based system for collecting, storing, securing, managing, and reusing research data.⁸ REDCap requires information technology and data storage support from each participating institution and is designed to hold regulated and protected data, including clinical trial results, in stable, secured, long-term storage. As far as can be ascertained, no one has reported using the REDCap system for operational metrics and record-keeping in a library setting, although doing so offers the following potential advantages: rapid point-and-click form development; control of user rights; the ability to protect confidential data; easy data entry, reporting, and analysis; and the ability to feed specified data selections to other software tools and user interfaces. Once registered in the system, the REDCap user has access from any Internet-connected computer by simple log-in, allowing use from anywhere.

This paper describes the novel use of the data capture system REDCap by HSCL librarians to create a database of the mediated searches conducted for patrons.

METHODS

With support and advice from UF's CTSI REDCap technical support team, the HSCL team utilized REDCap's built-in web-based form developer to create a database of mediated literature searches. Data collection forms were developed so that for each search request, information was captured on the patron; the type of search requested; the subject of the search; the search process; time spent and turn-around time; and deliverables returned to the patrons. During the development and implementation phases, all HSCL librarians had input into form design, as well as the opportunity to test data entry and contribute feedback. Additionally, the forms were examined by members of the UF CTSI's REDCap team, resulting in the identification and resolution of several issues with internal logic and database structure. Once the forms were fully tested and corrected by UF HSCL librarians and the CTSI's REDCap team, the database went into production in October 2010.

The database record consists of four forms, collecting the following information:

- *Patron Information Form*: data on the search requestor including position (e.g., faculty, student, patient); affiliation (i.e., UF college and department); and search request logistics, including method of contact, request date, due date, and billing status (free, paid, grant-funded).
- *Search Question and Limits Form*: data on the details of the search question including its purpose (e.g., patient care, education, research, policy) and relevant limits such as language, age groups, publication types, research subjects (e.g., human, animal, microbes), and publication dates.
- *Search Strategies and Databases Form*: data on the search process including useful search terms, databases searched, up to six PubMed searches, single searches for other databases, and relevant notes (see Figure 1).
- *Stats and Deliverables Form*: data on the search results including the time spent divided by task, the completion date, the type(s) and formats of deliverables. This form allows the upload of up to three documents.

Once the database went into production, 14 librarians at the HSCL and Borland Medical Library (UF-Jacksonville) were trained in the use of REDCap and introduced to the specific data forms relevant to this project. They began entering their mediated searches prospectively and continue to do so as a regular part of their work process. Additionally, two grant-funded students entered the complete print archive of past literature searches conducted by HSCL librarians; this included more than 1,100 paper request forms from 1998–2010. The back-entry of the paper forms was completed May 1, 2011.

RESULTS

Combining the retrospective data entry with continuing prospective recording of search requests, the database contains 1,511 records as of February 1, 2014. The project highlighted a number of advantages to utilizing REDCap for this purpose and illuminated some challenges faced in the development of the electronic data collection forms.

One of the major advantages of the REDCap system is the ease of form construction. A point-and-click web interface is available for creating forms; it is also possible to create and upload data dictionaries externally constructed in Excel. Therefore, creating initial forms for testing took only a few hours of work. Additionally, the opportunity to communicate quickly through e-mail and phone with the CTSI REDCap support personnel allowed rapid resolution of questions regarding the appropriate use of branching logic and record numbering.

Another advantage of the REDCap software is the ease of producing customized reports and exporting the data (see Figure 2). It allows export in multiple data formats for use in various statistical programs. Each librarian has created a customized, personalized report that automatically updates whenever searches are added or revised; individuals can select fields to be displayed and in what order, so that report structures can be individualized. Librarians have been pleased with their ability to produce their own reports, list their ongoing searches, and rapidly generate statistics for annual reporting and evaluation. In addition to the individual reports, the team created specialized reports to look at search requests overall by UF college, department, patron type, type of question, time spent on search, etc. (see examples in Figures 2–5). These reports are in the process of being analyzed for service and patron usage patterns; such patterns may be useful in determining appropriate HSCL liaison librarian staffing to match user need.

Although there were typical retrospective data collection issues in the older paper-stored searches, such as incomplete, illegible, or unverifiable information, these problems were avoided for current searches through the use of REDCap. The electronic forms were set to require the entry of specific data elements in specific formats. Additionally, rapid quality control assessment, built into the software, facilitates the identification of outliers in the data, so that they can be verified or fixed by librarians on a continuing basis (see Figure 6).

Also of value is the system's ability to assign and regulate user privileges including assignment of temporary and/or limited privileges, such as those provided to student workers. It is also possible to lock records that contain confidential information. These attributes provide data safeguards that are significant in the health sciences where protection of patient health information (PHI) is federally mandated.

Nonetheless, there were some challenges and limitations identified during this project. Additional entry form problems were discovered when active entering of data began. These problems required the database to be temporarily taken down, errors fixed, and the database re-started. Other findings were unexpected, such as the high use of resources not previously considered worth including in the check-box menus, for example, the provision of branched-logic-based open text boxes for entry of "other" choices proved to be valuable. Therefore, taking time to carefully plan out the data forms in advance and subject them to rigorous and imaginative testing before going into production is strongly recommended.

Another challenge was inducing librarians to continue entering searches on a permanent basis. Workflow changes can be difficult to implement, particularly those requiring time and effort. Regular encouragement is provided through verbal and e-mail reminders.

One significant limitation is the absence of a simple keyword searching capability in the database interface. This can be worked around using the report feature or by exporting the data into another program that allows keyword searching.

However, these challenges and limitations are minor in comparison to the easy access to the analysis, data sharing, and reporting capabilities the resulting REDCap database provides. This project demonstrates that REDCap can be utilized effectively for recording organizational metrics, even though it was originally designed to collect research and clinical data.

DISCUSSION

Recording searches in this database has become a regular part of the HSCL librarians' workflow. Permanent electronic storage in REDCap of librarian-mediated searches and related statistics addresses the organizational needs of individual librarians and the department as a whole. Having mediated search data readily available for analysis allows librarians to reuse and revamp search strategies and results where appropriate and to track their workload. Further, librarians are able to share and reuse each other's products, to identify subject topics for proactive preparation of materials for patrons, and to recognize potential areas of collaboration. REDCap benefits the department by facilitating better understanding of users and their expectations, tracking time and effort expenditures, determining which types of deliverables are most effective, and more efficiently targeting of librarians' continuing education efforts.⁵ Use of the system supports the mission and goals of the library by enhancing librarian proficiency and improving library services to the health science center.

For example, all UF HSCL librarians participated in four continuing education searching practice sessions in 2013 based on selected search requests from the REDCap database. Searches were chosen from REDCap-generated reports based on the field "Purpose of the Search." The coordinating librarian was able to choose search questions from four distinct categories: Patient Care, Bioscience Research, Systematic Reviews, and Medical Policy; the questions were of varying quality. Utilization of these stored search questions from REDCap allowed the team to practice search strategy development as a group and compare their results with each other and with the original searches stored in the database. In the future, a more systematic approach may prove fruitful for examining the issue of the quality of the searches within the database, thus providing an evidence base for both tracking and improving librarians' searching skills.

Iterative analysis of the collected data over time will allow the team to continually review outcomes and evaluate the success or failure of changes in practice instituted as a result of the initial data examination, thus facilitating evidence-based decision making. Project data may be used in the future to tailor adjustments to HSCL collections, identify new users and user needs, and make the stored resources available to users who may find it valuable in their own workflow. Results of the data analysis may lead to adaptations in how the library delivers mediated searches to HSCL user groups, and how library liaisons are allocated to campus units.

CONCLUSION

Given the successful implementation of REDCap for organizational data capture, the HSCL has adopted this tool for use in recording reference statistics, and librarians are considering using the tool for instructional statistics as well. The forms can be modeled after those used by the other campus libraries, and the results can be easily transferred into Excel spreadsheets for merging with other data sets. These changes will be designed to save effort, time, and cost to the library. Although training and skills development for librarians may involve short-term investments of time and effort, it is the project team's expectation that any short-term costs will lead to improved librarian skills and further efficiency in the future. In addition to future uses of REDCap at the HSCL, the project team recommends that librarians at any of the more than 800 REDCap Consortium⁷ partner institutions consider using this robust tool for their own organizational tracking and data capture.

REFERENCES

1. Role of Expert Searching in Health Sciences Libraries: Policy Statement by the Medical Library Association adopted September 2003. *Journal of the Medical Library Association*. Jan; 2005 93(1): 42–44. [PubMed: 15685273]
2. Holst R, Funk CJ. State of the Art of Expert Searching: Results of a Medical Library Association Survey. *Journal of the Medical Library Association*. Jan; 2005 93(1):45–52. [PubMed: 15685274]
3. Lasserre K. Expert Searching in Health Librarianship: A Literature Review to Identify International Issues and Australian Concerns. *Health Information and Libraries Journal*. Mar; 2012 29(1):3–15. [PubMed: 22335285]
4. Bredderman PJ. Paradox for Online Service Statistics and Reports. *Database*. Jun; 1991 14(3):31–38.
5. Vardell E, Loper K, Vaidhyanathan V. Capturing Every Patron Interaction: The Move from Paper Statistics to an Electronic System to Track the Whole Library. *Medical Reference Services Quarterly*. May; 2012 31(2):159–170. [PubMed: 22559179]
6. Gann LB, Pratt GF. Using Library Search Service Metrics to Demonstrate Library Value and Manage Workload. *Journal of the Medical Library Association*. Jul; 2013 101(3):227–229. [PubMed: 23930096]
7. REDCap. REDCap. Accessed November 5, 2013. <http://project-redcap.org/>
8. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research Electronic Data Capture (REDCap)—A Metadata-driven Methodology and Workflow Process for Providing Translational Research Informatics Support. *Journal of Biomedical Informatics*. Apr; 2009 42(2):377–381. [PubMed: 18929686]

The screenshot displays the 'HSCL - Database of Mediated Searches' interface. On the left is a sidebar with navigation links: 'My Projects', 'Project Home', 'Project Setup', 'Data Collection', 'Record Status Dashboard', 'Add / Edit Records', 'Search ID 1281', 'Patron Information', 'Search Question and Limits', 'Search Strategies and Databases', 'Stats And Deliverables', 'Lock all forms', 'Applications', 'Calendar', 'Data Export Tool', 'Data Import Tool', 'Data Comparison Tool', 'Logging', 'File Repository', 'User Rights', 'E-signature and Locking Mgmt', 'Graphical Data View & Stats', 'Data Quality', 'Report Builder', and 'Reports'. The main content area is titled 'Search Strategies and Databases' and shows 'Editing existing Search ID 1281'. It includes sections for 'Useful search terms', 'Databases Searched', 'PubMed Search Strategy', 'More PubMed Strategies?', 'Second PubMed Search Strategy', and 'Web of Science Search Strategy'. The 'Useful search terms' section contains a list of terms: 'Phytic acid[mh]', 'Phytic Acid[Substance Name]', 'Inositol hexaphosphate', 'IP6', 'phytic acids', 'Phytosterols', 'antioxidants', 'Alzheimer Disease[mh]', 'Alzheimer Disease', 'Alzheimer's Disease', 'treatment', 'therapy', 'therapeutics[mh]'. The 'Databases Searched' section lists various databases with checkboxes: PubMed, CINAHL, Web of Science, Cochrane Reviews, Cochrane Clinical Trials, ACP Journal Club, BIOSIS Previews, Psycinfo, CSA, CAB, and Other. The 'PubMed Search Strategy' section shows a complex search string: 'Phytic Acid[Mesh] OR Phytic Acid[Substance Name] OR "inositol hexaphosphate" OR IP6 OR "phytic acid" OR phytate[mh] AND therap[tiab] OR therapeutical[mh] OR treatment[tiab] OR clinical[tiab] AND English[tiab] AND human[tiab] NOT letter[tiab] OR news[tiab] OR comment[tiab]'. The 'More PubMed Strategies?' section shows a dropdown menu with 'One more' selected and a list of '5 more PubMed search strategies'. The 'Second PubMed Search Strategy' section shows a list of 'Two more' strategies. The 'Web of Science Search Strategy' section shows a search string: 'Phytic Acid OR "inositol hexaphosphate" OR "inositol hexakisphosphate" AND therap OR clinical'.

FIGURE 1.
Screenshot from data form 3: Search strategies and databases.

Counts/frequency: Medicine (464, 45.4%), Nursing (343, 33.5%), Dentistry (34, 3.3%), Pharmacy (50, 4.9%), Veterinary Medicine (50, 4.9%), Public Health and Health Professions (74, 7.2%), Other (17, 1.7%)

FIGURE 2.

Screenshot showing a REDCap-generated form representing the number of searches by college.

Searching Time: [Refresh Plot](#)

Total (N)	Missing	Unique	Min	Max	Mean	StDev	Percentile						
							.05	.10	.25	.50 Median	.75	.90	.95
1478	31 (2.1%)	76	0.10	60.00	2.06	2.93	0.25	0.50	0.75	1.50	2.50	4.00	6.00

Lowest values: .1, 0.15, .15, 0.16, .17

Highest values: 24, 25, 30, 45, 60

FIGURE 3.
Screenshot showing a REDCap-generated form representing the searching times.

Counts/frequency: Patient Care (413, 33.2%), Education (not patient/consumer) (260, 20.9%), Clinical Research (676, 54.3%), Bioscience Research (22, 1.8%), Patient Education (13, 1%), Medical Policy (24, 1.9%), Community Outreach (5, 0.4%), IACUC (150, 12%), IRB (5, 0.4%), Grant writing (30, 2.4%), Article Publishing (60, 4.8%), Systematic Review (8, 0.6%), Other (111, 8.9%)

FIGURE 4.

Screenshot showing a REDCap-generated form representing the number of searches by the purpose of the search.

Counts/frequency: PubMed (1124, 78.2%), CINAHL (310, 21.6%), Web of Science (237, 16.5%), Cochrane Reviews (15, 1%), Cochrane Clinical Trials (18, 1.3%), ACP Journal Club (3, 0.2%), BIOSIS Previews (84, 5.8%), PsychInfo (109, 7.6%), CSA (36, 2.5%), CAB (94, 6.5%), Other (355, 24.7%)

FIGURE 5.

Screenshot showing a REDCap-generated form representing the number of searches by database searched.

HSCL - Database of Mediated Searches

Data Quality

This module will allow you execute data quality rules upon your project data to check for discrepancies in your data. Listed below are some pre-defined data rules that you may utilize and run. You may also create your own rules or edit, delete, or reorder the rules you have already created. To find discrepancies for a given rule, simply click the Execute button next to it, or click the Execute All Rules button to fire all the rules at once. It will provide you with a total number of discrepancies found for each rule and will allow you to view the details of those discrepancies by clicking the View link next to each. [Read more detailed instructions.](#)

Data Quality Rules					Execute All Rules	Clear
	Rule #	Rule Name	Rule Logic (Show discrepancy only if...)	Total Discrepancies	Delete rule?	
	a	Missing values*	-	Execute		
	b	Missing values* (required fields only)	-	Execute		
	c	Field validation errors (incorrect data type)	-	Execute		
	d	Field validation errors (out of range)	-	Execute		
	e	Outliers for numerical fields (numbers, integers, sliders, calc fields)	-	Execute		
	f	Hidden fields that contain values**	-	Execute		
	g	Multiple choice fields with invalid values	-	Execute		
	<input type="text"/>		<input type="text"/>			
	<input type="button" value="Add"/>		<input type="text"/>			
	Enter descriptive name for new rule (e.g. Participants below age 18)		Enter logic for new rule (e.g. [age] < 18) How do I use special functions?			

FIGURE 6.
Screenshot showing the built-in data quality tools of REDCap.