Research Article **Open Access**

Developing a Web-Based, Patient-Centered Data Collection and Management Approach for a Multi-Center Lung Cancer Study

Monirul Islam KM[1*], Jiajun Wen[1], Xiaqing Jiang[1], June Ryan[3], Ann Fetrick RN[2] and Apar Ganti[4,5]

[1]University of Nebraska Medical Center (UNMC) College of Public Health (COPH), Department of Epidemiology, Omaha, Nebraska, USA
[2]UNMC, COPH, Department of Health Services Research, Administration, and Policy, Omaha, USA
[3]Nebraska Cancer Coalition, Omaha, Nebraska, USA
[4]Veterans Healthcare Administration, Nebraska-Western Iowa Health Care System, Omaha, USA
[5]UNMC, Division of Oncology and Hematology, Omaha, Nebraska, USA

Abstract

Research Electronic Data Capture (REDCap), a web-based approach, provides an easy and efficient way for trained individuals to enter, maintain, and use project data. It also offers the capacity to use surveys and common data collection forms in a way that saves researchers time and money. We found these strengths and the ability to store all study data in one place so as to reduce data safety and confidentiality concerns very supportive of our patient-centered outcomes research (PCOR), with emphasis on patient-centeredness. This article describes our development of a web-base, patient-centered data collection and management approach for our multi-center lung cancer PCOR.

Background

The Patient-Centered Outcomes Research Institute (PCORI) is funded through the Patient-Centered Outcomes Research Trust Fund. The mission of PCORI is to improve healthcare delivery and outcomes through informed healthcare decisions based on evidence coming from clinical effectiveness research [1]. This paper is based on one aspect of our on-going patient-centered multi-center lung cancer outcomes study funded by the PCORI.

The overarching objectives of our research are to determine whether patients' preferences about potential adverse chemotherapy treatment outcomes of advanced Non-Small Cell Lung Cancer (NSCLC) affect their definition of treatment success and/or their experiences with chemotherapy side effects; and, if so, to explore how to integrate NSCLC patient treatment preferences into clinical treatment planning. The current study is patient-centered in all aspects, but particularly in our conduct of Phases 1 and 2 that led us to develop a web-based patient-centered data collection and management approach for our multi-center lung cancer study.

The purposes of this article are, first, to outline the patient-centered methods we used to highlight the need for a web-based approach to data collection and management, and, next, to delineate the steps we took to develop the web-based approach whereby we learned that this approach clearly supports the patient-centeredness of our research, with the additional benefits of making data collection, management, analysis, and reporting more efficient and effective than other methods commonly used.

Methods

Conduct focus groups (Phase 1)

In Phase 1 we presented to four different focus groups our draft of the questions we wanted to ask patients in Phase 2 to address our research questions and contribute to the achievement of our overarching goals. The focus groups were composed of representatives of two classes: (1) advanced lung cancer patients (past or current patients, their family members, and other patient advocates) and/or (2) clinicians (oncologists, mid-level practitioners, nurses, and others involved in the care of advanced NSCLC patients).

Purposes of the focus groups included gathering information for refining the data collection tools/questionnaires to make them more patient friendly, collecting ideas for recruiting study participants, obtaining suggestions for dissemination of the study findings, and receiving input on recruiting and keeping patients engaged throughout the study. Focus group feedback resulted in significant revisions to the data collection tools and processes to make them more patient-oriented [2].

Three critically-important themes we heard from focus group participants included: (1) that patients would like personal interviews rather being asked to complete a printed or online questionnaire on their own; (2) that patients would like research team members to extract as much as possible of the information needed to answer study questions from the patient's medical record; and (3) that patients would like the research team to carry any technological burden of doing the study.

This led us to: (1) arrange for staff to conduct personal interviews, as the first choice offered to subjects; (2) cut in about half-from about 11 pages to about 5 pages the amount of data we would need to collect from the patient at any one time, with staff collecting as much data as possible from the medical record; and (3) develop a secure, web-based database for the four participating cancer center sites to submit data to

the coordinating center [3]. It is this latter point that we will focus on in more detail in the next section of this article.

Interview patients (Phase 2)

In Phase 2 we have obtained/are obtaining data regarding patient treatment success definitions and patient treatment experiences and preferences during three different times of chemotherapy treatment. Also, data concerning patient demographic and other characteristics, such as, co-morbidities, concomitant medications and types of chemotherapy used for treatment of advanced lung cancer are collected. This information from patients and/or their medical records reflects a continuum of patient-centered experiences and preferences of patients, before, during and after a number of chemotherapy cycles of treatment.

The research is ongoing, and is being conducted in four cancer centers in Nebraska and South Dakota, two of them mainly rural and two urban. Each participating center contributes its own research personnel with physicians and oncology nurses and other supporting team members. Because of the patient-centered results of Phase 1 (Focus Groups) and to improve the response rate and adherence to the study protocol, patients are interviewed by a study coordinator (an onsite nurse or experienced clinical trials staff person). Although no one has yet chosen to do so, there is the option for patients to complete a printed questionnaire on their own. Additionally, as a result of the focus group results, and in order to assure data quality and safety, provide consistency and promote project efficiency and especially to enhance timeliness of submission and receipt of the data by the coordinating center, we have developed a web-based data application to collect, store and manage the study data.

A Web-based approach

Compared to more traditional paper-based data collection systems, the advantages of systematic web-based data collection tools includes reduced time for data entry, lowered cost, ease of data entry, flexibility in format and the ability to capture additional response-set information [4,5]. Depending on the tools chosen and the format of the data collection tools/forms, the data-entry interface can be nearly identical to the paper-based questionnaire, which allows the data entry procedure to be easy for data-entry personnel.

The Research Electronic Data Capture (REDCap) is a secure, web-based application designed to support data capture for research studies [6]. Developed by Vanderbilt University, REDCap has been used for more than 92,000 projects worldwide. In addition to advantages of common web-based data collection tools REDCap has its own advantages including assigning users' rights and data access groups, longitudinal data collection tools, online designer, and a quality control module.

Based on the results of the focus groups, our review of various web-based applications, and, in consultation with our university's information technology team, we chose to develop our web-based, patient-centered data collection and management approach using the REDCap application. See the appendix for details that give more detail regarding the steps we took to do this.

Once developed, we did beta testing and our project staff experienced a short learning curve for using the web-based system. Therefore, we were able to go quickly to training staff at each of the four participating cancer center sites. Because of the distances between sites, coordinating center staff, and trainers, we conducted three of the four training sessions via video conferencing. The fourth site was

located across campus from the coordinating center; therefore, we were able to do this training in-person. The orientation and trainings by video-conferencing worked so well that we were in full production in a matter of weeks. Once in production, there were minimal adjustments required. The speed of start-up and testing and the few corrections needed once we were in production were taken as indications that we had made good decisions to: (a) utilize a web-based approach, and (b) create it using REDCap. Subsequent experience has continued to support these choices as patient-centered and, we have learned that they are also staff-friendly.

In our research, after the patient is recruited by the study coordinator and has signed the informed consent, the patient is interviewed by a research team member at their cancer center. The data of the interview is input into REDCap data collection forms, and the record for that patient is established. When it is time for the second and third interviews, staff at the cancer center repeats the process of interviewing the patient as for the first interview. Once the patient is dismissed, usually about 20 minutes per interview, research staff simply select the patient's record and continue working on it, which simplifies storing and tracking of data.

For the management of a multi-center study, the Users Rights and Permissions module provides an efficient way to assure data safety for a multi-center study [7]. In our study, we have personnel from four different centers in two states entering data. The Users Rights module of REDCap allows us to assign individual personnel the right of accessing or manipulating certain parts of the project, and it also allows limiting or expanding the privileges of access and/or revision to multiple individuals at one time. To avoid personnel from one site viewing data of other sites, we assign all users of one center to one data access group, so that the data entered by the users from the group will only be viewed within this user group. The users who are not assigned to any group have the access to all the data stored in the project, and they usually are the core study group and the data manager.

Discussion

For a patient-centered study, the experience and preference of patients and those who care for them, whether they are informal (such as, family members) or formal (such as, oncologists) caregivers must be considered throughout the study [8]. We conducted focus groups during the first phase of our study to receive feedback from patients, caregivers and advocates to refine our data collection forms and methods of data collection and management. The outcomes of this step include significant revisions to the patient data collection tools that we expect will result in high response rates and patient adherence to study protocols, and elimination of inappropriate and redundant questions.

REDCap provides an easy and efficient way for trained individuals to enter, maintain and use the project data. REDCap offers the capacity to use surveys and common data collection forms. For surveys, users can participate in the study just as they would any other secure, online survey. The survey link can be accessed on any public website on which the researchers post, or through emails that the researchers directly send to the participants. Although Phase 3 (Surveying Oncologists) is beyond the scope of this article, we would like to point out that the ability to for users to securely respond to a survey using the web will be a helpful feature for us in Phase 3. This may help improve the response rate and adherence of the participants [9].

The use of the web-based approach to data submission also saves the researchers time and money. They do not need to prepare hard-

copies or arrange for a place to store paper questionnaires safely. All study data, as well as, the files of the study can be stored in the REDCap system thus greatly reducing data safety or confidentiality concerns. Additionally, the researchers do not need to set up time schedules to distribute the questionnaires or study forms and spend effort to keep track of them, because the REDCap can automatically record and track the study data.

Conclusion

The REDCap platform provides a tested, secure, and efficient way to manage the study database for a multi-center study. We have been pleased with our first-time use of this web-based data collection and management approach for our four-cancer center study. And, especially appreciate that it has supported our research as patient-centered and has added the benefits of also being research team member centered.

Appendix

The basic function of REDCap is to create and manage a data-based project. The first step in using REDCap is to create a project following the directions under the "Create New Project" tab. Each new project must have a project name and purpose. Once these components are selected, one may practice or develop a real project that can be put into production; the purpose can be revised at any time during the project development. REDCap also allows one to develop a new project based on the existing project template or create an empty project and develop it innovatively.

Project setup

REDCap includes two basic project formats: classic data collection forms and surveys. The classic data collection forms are used mostly among researchers who directly enter the REDCap platform to record study data and perform quality checks; the surveys provide a friendly interface for participants to enter by simply clicking on the public link from any computer. In our research, the four on-site coordinators enter the data from the paper-based questionnaires answered by patients, REDCap also supports longitudinal data collection when the study includes forms that will be utilized multiple times throughout the research. This function is included on the project set up page and events can be assigned to trigger the use of a certain form at a certain point in the study. It is important r to decide whether to use the longitudinal data collection function before developing the data collection forms to avoid duplication of effort. Our research goals include determining whether patients' preferences toward treatment side effects and the definition of treatment success change over time, requiring repeated measurement of the same variables, so we selected the longitudinal data collection function in our study.

The Online Designer is the REDCap module where forms can be designed and developed.

The basic unit of a study form is an "instrument" which can be a survey form or a data collection form. The instrument can be used to start a new section of a data collection form, and these new sections can be arranged in order and assigned events later to enable longitudinal data collection. In our study, all data-entry is done by -site study coordinators at four collaborating centers who are able to access REDCap simply by inserting usernames and passwords assigned by the REDCap administrator. All instruments in our study are data collection forms.

The data collection fields are built within each instrument. One may choose to add a new single field or a matrix of multiple choice fields of the same format. The fields can easily be used to input demographic other study data. The type of a field can be text box, notes box, calculated field, multiple choice (drop-down list or radio buttons), checkboxes, yes-no question, true-false question slider scale, file upload tool or descriptive text. These types cover the most types of questions in a general study instrument. Our patient-centered study used descriptive text to guide the data-entry personnel; we built the fields in the same pattern as on the paper-based questionnaires. In other words, the computer form and the paper forms were identical in format as well as content.

REDCap is able to make any specific field a required field or. If the required field is left blank when the form is submitted, a warning will pop up to alert the data-entry person to fill the blank. REDCap also enables tagging the identifier variables for each field and if the identifiers are indicated, these variables appear in red color in the Data Export Tool. The researcher can select the de-identification option prior to exporting the data to avoid violating the HIPP compliance.

Project testing

The last and perhaps most important step before putting your project into production is to test the project thoroughly and make sure every field and logic branch works correctly. Since there is no automatic testing procedure in REDCap, the only way to do the testing is to enter test/sample data into each field under the developing status. To test the data collection forms, use the "Add/Edit Records" to add new records, and input the test data. After the data has been entered, the forms can be viewed under the "Record Status Dashboard," from which all forms of records are marked completed, unverified or in-completed in the grid. The REDCap allows multiple users to test the same project at one time, and the users can view all records or only those in the group they have been assigned. To test the survey, it is as simple as sharing the public survey link via email or other approaches. The users can access to the survey interface through a simple click to the link, entering the data and testing the logic brunches.

Move to production

To move the project to production status, REDCap will send a request to the REDCap administrator, who will help move the project into production status. Before this, one should make sure all identifiers have been tagged to avoid violating the HIPPA compliance. Also, all testing data and calendar events will be erased, so you do not store any real data in the testing procedure.

Quality control and database management

Quality control and database management are conducted by the core study group and the database manager regularly. Each time an interview is completed, the record is locked by the data manager; the data is may not be modified until it is unlocked under the permission of the database manager. Our study included data required by the Scientific Review Committee (SRC) of the University of Nebraska. The twelve identifier variables required by the SRC are recorded and sent to the SRC as soon as a new patient record is established. At the same time, the data manager is responsible for checking the data quality. The data quality check includes calculating missing values, checking inappropriate format, typographical and other data entry errors. The database manager contacts the data entry person to compare the data entered from hardcopy forms and corrects any data entry errors if necessary. The database manager generates a monthly report regarding the data quality of each center and distributes the reports to the on-site

coordinators. Each center is only able to view the data quality report for that center to ensure integrity. The report serves as feedback on the work that the center has done in the previous month, and provides advice and suggestions for future work to maintain or improve the data quality.

The whole database is downloaded by the data manager using the REDCap data export tool at the end of every Tuesday and Friday when there is any change of data or any new data added to the database. The data is exported in the format of an Excel document and then uploaded to the secured shared drive accessed only by the project team members.

References

1. Patient-Centered Outcomes Research Institute (PCORI) (2013) Mission and Vision. Retrieved March 24, 2014.

2. Rabiee F (2004) Focus-group interview and data analysis. Proc Nutr Soc 63: 655-660.

3. de Carvalho EC, Batilana AP, Simkins J, Martins H, Shah J, et al. (2010) Application description and policy model in collaborative environment for sharing of information on epidemiological and clinical research data sets. PLoS One 5: e9314.

4. Granello DH, Wheaton JE (2004) Online data collection: Strategies for research. Journal of Counseling and Development 82: 387-393.

5. Krishnankutty B, Bellary S, Kumar NB, Moodahadu LS (2012) Data management in clinical research: An overview. Indian J Pharmacol 44: 168-172.

6. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, et al. (2009) Research electronic data capture (REDCap)-A metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 42: 377-381.

7. Das S, Zijdenbos AP, Harlap J, Vins D, Evans AC (2012) LORIS: a web-based data management system for multi-center studies. Front Neuroinform 5: 37.

8. Sacristán JA (2013) Patient-centered medicine and patient-oriented research: improving health outcomes for individual patients. BMC Med Inform Decis Mak 13: 6.

9. Obeid JS, McGraw CA, Minor BL, Conde JG, Pawluk R, et al. (2013) Procurement of shared data instruments for Research Electronic Data Capture (REDCap). J Biomed Inform 46: 259-265.