MedRed: A Health-care Data Acquisition Service for Research Purposes

Jean-Paul Calbimonte[1], Fabien Dubosson[1], Roger Hilfiker[2], Alexandre Cotting[1], and Michael Schumacher[1]

[1]Institute of Information Systems, e-Health Unit
[2]Institute of Health Sciences
University of Applied Sciences and Arts Western Switzerland,
HES-SO Valais-Wallis, Sierre, Switzerland
{name.surname}@hevs.ch

Abstract. Research in the health-care domain requires the collection of important and exhaustive datasets, in order to validate a scientific hypothesis, or to assess the effectiveness of a treatment, technology, medicine, or procedure. The data acquisition phase for this type of work requires an often under-estimated amount of time and effort, while needing to keep high quality standards for the entire process. Many of the tasks associated with data acquisition are often carried out manually, resulting in error-prone procedures, hand-transcription, inaccuracy, and time delays to produce a resulting usable dataset. This paper presents MedRed (Medical Research Data Acquisition Platform), a platform and a service designed to facilitate the data acquisition process for researchers in the health-care domain, using the REDCap software for data capture. This service is available in a first stage, for all scientists of the HES-SO (University of Applied Sciences and Arts Western Switzerland) schools in Switzerland, and partially supported by the SwissUniversities CUS-P2 program.

Keywords: clinical data acquisition, data capture, data life-cycle management

1 Introduction

In the medical domain, research protocols often require the participation of patients and volunteers in scientific studies, which are conducted in order to assess, evaluate and validate a certain hypothesis. Although these studies have important differences and are heterogeneous, depending on the topic, and the nature of the investigation, they usually follow similar data acquisition patterns. Concretely, a large amount of data is obtained in these studies through data acquisition instruments, such as surveys, forms and questionnaires, gathering both quantitative and qualitative data. Examples of these include demographics, symptoms, perceptual information, cognitive data, feedback, emotions, descriptions of experience, etc.

The implementation and usage of these instruments is challenging, considering the many different types of inputs depending on the type of study, and the need to adapt the instruments to the target population, including usage of language, usability, and interaction. Moreover, cross-cutting issues, such as data privacy protection, safe

storage, and data quality assessment need to be handled, which are oftentimes not the expertise of clinical researchers. Furthermore, nowadays these instruments need to be implemented supporting different devices, including smart-phones, tablets, laptops, etc., further increasing the load of work on the researchers.

Although in certain Swiss medical research centers, these issues are addressed by a dedicated infrastructure (e.g. Clinical Trial Units), there is a need and an opportunity for providing support on clinical data acquisition for researchers in research and education institutes. The University of Applied Sciences and Arts Western Switzerland (HES-SO), is an institution dedicated to applied research in different areas, including health-care. Being the 2nd largest higher education center in Switzerland (in terms of student population), and having several affiliated schools and institutes working on different health research topics, it lends itself as a potential important player in clinical data acquisition at a national level.

This paper presents and describes MedRed@HES-SO (or simply, MedRed[1]), a Medical Research Data Acquisition Platform, available for all researchers of the HES-SO affiliated institutes and schools, with a vision of national outreach. MedRed is based on the usage of the REDCap (Research Electronic Data Capture) system, a widely used software for data acquisition in the medical domain, with proven support in web and mobile environments. Te MedRed project and its scope is discussed in length in the remainder of the paper, starting with its goals and approach in Section 2, the implementation in Section 3, its relationship with previous works in 4, and the sustainability and outreach aspects in Section 5.

2 MedRed: Platform Approach

In the MedRed project, the general objective is to establish a research data acquisition and storage platform based on the REDCap software, aiming to support current and future health research projects at the HES-SO (in the cantons of Geneva, Vaud, Valais, Fribourg, Neuchâtel and Jura of Switzerland).

As a pilot project, in this first stage, MedRed focuses on the deployment of the platform in selected pilot projects led by the Haute École de Santé at HES-SO Valais.

The project goals can be summarized as follows:

- To setup, install and deploy a data management solution for surveys and questionnaires for research medical data, centralized in the HES-SO data center, and based on the REDCap platform.
- To design, adapt and implement data acquisition instruments (e.g. surveys and questionnaires) for selected pilot projects led by the Haute École de Santé at HES-SO Valais.
- To provide validation and evaluation of the effectiveness and usability of the platform in the selected pilot projects.
- To implement ad-hoc training modules for HES-SO researchers.
- To analyze and formulate the needs of a middle and long-term strategy for a research lifecycle data service facility (methodological support for data science, support of

[1] MedRed: `http://w3id.org/medred/project`

writing of protocols for ethical committees, software for data acquisition, support in statistical evaluation).

The project is designed to cover the management of research data mainly from health science studies in an end-to-end fashion: starting from the design of input questionnaires and surveys, the data acquisition process, the storage and persistence of data, and its posterior availability for analysis. This holistic data management approach includes cross-cutting issues such as data security and privacy, authentication, validation and data integration. In order to attain the expected goals of the project, we propose adopting an ICT platform designed specifically for this purpose, namely REDCap (Research Electronic Data Capture).

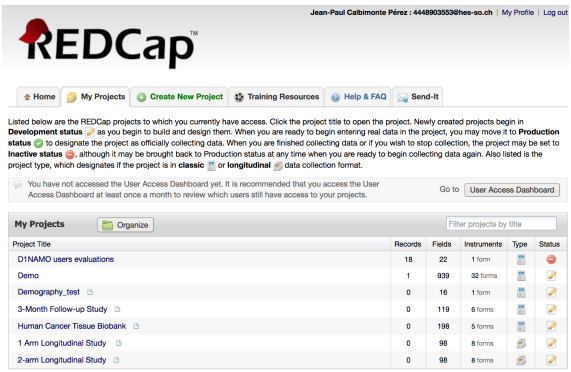

Fig. 1. REDCap instance deployed at HES-SO for MedRed. REDCap users can create their own projects, define their instruments, and collect the data using the provided functionalities.

REDCap is a web-based electronic data capture (EDC) solution for designing clinical research databases and questionnaires. It has been built by a dedicated team at Vanderbilt University, and has been widely used as a platform for numerous projects in hundreds of institutions, including hospitals, universities, medical research centers, national research programs, etc. The successful adoption of REDCap in such a large and diverse set of clinical projects provides substantial evidence that it can be used to efficiently manage the complete data management process.

The adoption of such a platform entails several challenges that we address in this project. We can summarize them as follows.

Paperless research data input: Currently, most of the clinical data acquisition activities at the HES-SO Institute of Health are performed using paper-based questionnaires and forms. This leads to several issues related to missing data, mistakes made while filling the forms, interdependence of questions, which are not respected, among others. These issues have as a consequence, a reduced accuracy and quality of the data. Moreover, when it comes the time to export the data and process it, manual processing is needed, which leads to further potential mistakes.

Computer-supported acquisition design: Many clinical data acquisition protocols exist in standardized forms, and could be adapted to an electronic format. In these cases, it is necessary to preserve the characteristics of the questionnaires and guarantee that it follows the same goals. For new protocols and new studies, it would be expected that the electronic approach would facilitate the task of designing the data input instruments.

Mobile-device electronic acquisition: In many use cases the usage of mobile phone and tablets will be instrumental to facilitate the data acquisition. However, this would change the current workflow of the personnel in charge of this task. While the mobile applications are designed to make this easier, it is important to avoid perturbing the work and procedures during the clinical data acquisition.

Distributed data acquisition: The usage of electronic data capture will allow automatic collection and storage of data, in a distributed fashion, and coming from a series of smartphones and tablets. While this is advantageous, permitting simultaneous data capture, it is also needed to verify that the platform is capable of maintaining performance and availability levels even if the system is under considerable load.

Support for instrument designers and administrators: Using the proposed new set of tools, we will change radically the way in which the instrument designers work. It is important to guarantee that they are able to craft the forms and questionnaires that they need, avoiding mismatch and design mistakes that can have negative consequences during the acquisition process. Also, it is vital to train administrators in the authorization management, to determine who has access to functionalities such as, importing/exporting data, access to identifiers, access to instruments, etc.

Training for acquisition assistants: Similarly, it is important to provide the necessary training for the people who will be in charge of the data acquisition. These persons not always have the technical expertise to master the electronic data capture tools. Therefore, it is key to provide them with pedagogical material and easy-to-follow instructions, so to avoid mistakes and other issues.

Deployment of an inter-institutional service: The usage of these tools is thought to be applied to all institutes affiliated within HES-SO, in the long term. This will also be challenging from the administrative and managerial point of view, as it will deal with data that belongs to different parties, although under the same institutional umbrella. Coordination and support at different levels should be considered. While in this project we limit to studies carried out by the Institute of Health, it is important to take into account these aspects from the beginning.

Data security and privacy management: The data that is collected in this type of service can be very sensible and should be subject to strict security and privacy regulations. These must be ensured not only during the usage of the application, through the setting of the appropriate authorization rules, but also for the storage of the data, which would also require secure backups and potentially replication.

3 Implementation

The MedRed project is structured in such a way that it addresses the challenges raised in the previous section, and is composed of the following four main pillars: (i) Infrastructure & Technical Support, (ii) Methodological support, (iii) Evaluation and Validation, and (iv) Data lifecycle management. The general coordination of the MedRed service is managed by HES-SO Valais-Wallis in Sierre, providing support for all technical, methodological, validation and data lifecycle management aspects. In addition, the technical infrastructure, i.e. the deployment of the REDCap instance, is managed by the HES-SO central IT services at Fribourg (see Figure 2). Both HES-SO users of the platform as well as external partners can access the MedRed infrastructure through the same interface, secured with a SWITCH AAI authentication system. The MedRed REDCap instance is available at the following URL: `http://redcap-hes.so.ch`.

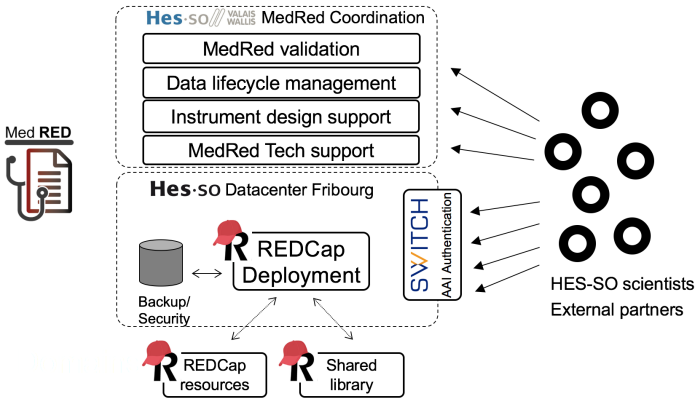

Fig. 2. MedRed project organization: All HES-SO researchers as well as allowed external partners can access the REDCap infrastructure through the SWITCH authentication. MedRed also coordinates the different support services described in this paper.

3.1 Infrastructure and Technical Support

The first pillar of MedRed is to provide the technical infrastructure and support for the data acquisition platform. This translates first into the acquisition of the REDCap license, its installation and deployment, and its maintenance. The HES-SO Valais-Wallis Institute of Information Systems (IIG) leads this effort, with the support of the central HES-SO It services in Fribourg, where the services is hosted, including backup and monitoring operations. This configuration enables the MedRed REDCap installation to be available from the start for all researchers associated to any of the schools affiliated to HES-SO. The REDCap instance has been running since fall 2016, and its authentication system has been integrated with the SWITCH AAI mechanism provided for all shared HES-SO infrastructures.

Nevertheless, the deployment of the platform is only the first step, and since the start of the project, MedRed has consolidated a support service for technical issues, so that researchers who use the platform can find suitable solutions, work-arounds, and recommendations in case of need. This also includes support for using the platform with mobile devices, and the REDCap app, which facilitates the work on the field for administrators and end-users.

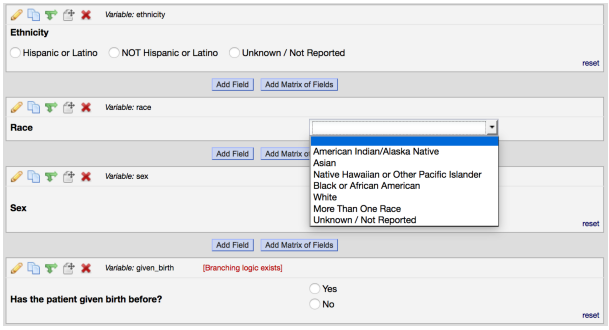

Fig. 3. Composition and derivation in MedRed. Left: a sample instrument including its organization in sections and items. Right: a study may incorporate instruments created previously (e.g. Hip survey) or create new instruments reusing items from others (e.g. the eating disorder questionnaire).

3.2 Methodological support

One of the key innovations of MedRed is the inclusion of dedicated methodological support and assistance, so that researchers who use the platform can create, manage and conduct their studies following best practices and guidelines for data acquisition. In a first step, this includes the design and adaptation of training material for data acquisition. This is mainly directed to health professionals who will use the MedRed infrastructure for conducting their studies. A first training phase has already been performed for the first pilot projects that started along with MedRed, but will continue through the lifetime of the project. The training ensures that the study administrators are first able to design their own instruments, use the REDCap designer (see Figure 3) to formulate the questionnaires, associate them to variables, make use of longitudinal, randomization, privacy preservation features, etc., so that the quality of the instruments is optimal. On a second stance, the training is tailored for the users of the instruments, those who will actually perform the data acquisition in the field, with patients and volunteers. The HES-SO Valais Institute of Health leads the conduction of the methodological support, which also includes aspects related to the posterior use of the data, e.g. considering statistical data analysis, adherence to norms, and reuse of existing or similar instruments, etc.

In order to guarantee that at least an initial number of real studies are implemented using the platform, a few projects have been selected for piloting, including:

- Swiss CHEF Trial: Comparison of Home-based Exercise programmes for Fall prevention and Quality of life in older adults (Randomized Controlled Trial).
- Validation of a decision tool for the prediction of non-return to work after occupational rehabilitation (Cohort Study).
- Non-Invasive Ventilation in Swiss Hospitals (Survey)
- Self-reported tool for the prediction of falls in community dwelling older adults (Cohort Study with monthly assessment during 12 months).

Additional studies have also been added, and others are under discussion for inclusion in the pilot phase.

3.3 Evaluation and Validation

In health-related studies, it is crucial to maintain high quality standards throughout all steps of the data acquisition process. For this reason, it was important to include a dedicated task for the evaluation and validation of ongoing studies and their instruments, as well as evaluation of the datasets collected.

For the pilot phase of MedRed, the main goal of this task is to assess the effectiveness of the usage of the platform for the chosen pilot projects, and to assess the quality levels of the data that was obtained. This includes the assessment of the usability of the tools, and identification of pitfalls and potential issues that would need to be addressed in future studies. Concerning the acquired data evaluation, this include finding inconsistencies, missing data, violation to input rules, quality of answers, etc.

3.4 Research Data Lifecycle Management

Data acquisition is only one part of the larger life-cycle of a scientific process. For this reason, MedRed incorporates tasks related to the task of defining a strategy for a research data lifecycle management, potentially reaching in the long term for all of the affiliated HES-SO institutes. This includes the definition of guidelines and foundations for a potential service for research data management, including a medium and long term vision, support and maintenance plan, feasibility analysis and inter-relationship with other existing services.

Such a strategy will allow in the future to integrate the MedRed platform with Open Data services, link with Open Access catalogs, create potential Open Innovation opportunities, etc.

4 Related Work and Services

Up to now there has not been an initiative for establishing a comprehensive research data management solution in the scope of the HES-SO. Currently, each research group handles the acquisition, storage and analysis process of research data independently, through ad-hoc combinations of paper-based forms, scanned documents, excel files, etc.

However, the introduction of the proposed platform into the workflow of the different research groups of the HES-SO Institute of Health requires a careful integration into

their everyday practices, in order to avoid intrusiveness, and in such a way that existing protocols and norms are respected.

While in the past, manually filled forms were the norm for acquiring data in this context, nowadays the use of Electronic Data Capture (EDC) solutions have shown to improve the efficiency of the process, while maintaining quality and accuracy standards [1, 6]. In particular, EDC helps increasing the efficiency of the entire data management process, reducing and/or eliminating data transcription and transmission times, providing data validation and input enforcement, or helping scheduling the site visits [2, 3]. Furthermore, EDC is expected to provide faster access to data, which can help to perform live-analytics and decision making over the existing acquired datasets. Due to these benefits, clinical research organizations, pharmaceutical companies, and university hospitals, among others, make use of EDC and related clinical data management systems such as OpenClinica, REDCap, TrialDB, InForm, Medidata Rave, Datatrak, to name a few [5].

Given the large number of clinical studies that are performed worldwide, and their complexity, it has become a need to share their results, as well as their structure and metadata. This would make it possible to perform tasks such as: validating existing protocols, reusing and refining clinical research instruments, extending previous studies, performing surveys and systematic analytics of clinical trials, etc. Significant efforts have been made to agree on standards for clinical studies, and the ODM (Operational Data Model) [4] proposed by CDISC[2] has been adopted by several regulating bodies and also EDC software tools such as REDCap.

Although there exist other data management systems for electronic data capture, the proposed REDCap solution has the advantage of being provided at essentially no cost for academic purposes, even if it is not an open-sourced software. Furthermore, it has an active and well-established community of users, as well as a large number of partner institutions, including several universities and institutes in Switzerland (e.g. HUG, CHUV, Lausanne, CTU Bern, etc.) Among these institutions, none belong to a Swiss University of Applied Sciences. Thus, there is a clear need to establish in Switzerland such a platform to support for data research in professional health sciences (e.g. nursing, physiotherapy, ergotherapy, etc.)

5 Sustainability, Outreach & Discussion

We have presented the MedRed project for data acquisition in health-care studies, implemented for the entire network of HES-SO researchers, institutes and schools. This initiative has the ambition to constitute a first-of-a-kind service that supports scientists in the health domain, not only with technological but also methodological tools that will help them facilitating common tasks in their research workflows, while raising the quality of the research outcomes.

We have outlined in this paper the different advantages of such approach, but we have also analyzed the costs of maintaining such an infrastructure. While in the current pilot phase, there is direct support form external funds (SwissUniversities CUS-P2), afterwards the platform will continue working as a service available to the whole HES-SO.

[2] CDISC (Clinical Data Interchange Standards Consortium): `http://cdisc.org`

The maintenance and support will be assured through a medium-term contract with the Rectorate of HES-SO, and a business plan that relies on mandates for the participating schools, similar to the functioning of a Clinical Trials Unit.

Finally, we think that this initiative would pave the way for a future creation of a dedicated center for data lifecycle management for the HES-SO, which would cover not only the acquisition and storage but other additional services, such as data policy regulations, ethical committee submission support, data reuse best practices, etc., which would help unifying the research data strategies, which are essential for today's requirements for project fund-raising and for data quality assessment.

References

1. P. Bruland, B. Breil, F. Fritz, and M. Dugas. Interoperability in clinical research: from metadata registries to semantically annotated cdisc odm. *Stud Health Technol Inform*, 180:564–8, 2012.
2. K. El Emam, E. Jonker, M. Sampson, K. Krleža-Jerić, and A. Neisa. The use of electronic data capture tools in clinical trials. *J. medical Internet research*, 11(1):e8, 2009.
3. P. A. Harris, R. Taylor, R. Thielke, J. Payne, N. Gonzalez, and J. G. Conde. Research electronic data capture (redcap)—a metadata-driven methodology and workflow process for providing translational research informatics support. *J. of biomedical informatics*, 42(2):377–381, 2009.
4. W. Kuchinke, J. Aerts, S. Semler, C. Ohmann, et al. Cdisc standard-based electronic archiving of clinical trials. *Methods of information in medicine*, 48(5):408–413, 2009.
5. J. Shah and et al. Electronic data capture for registries and clinical trials in orthopaedic surgery: open source versus commercial systems. *Clinical Orthopaedics and Related Research®*, 468(10):2664–2671, 2010.
6. T. Souza, R. Kush, and J. P. Evans. Global clinical data interchange standards are here! *Drug discovery today*, 12(3):174–181, 2007.