OMEVAC – Open Mobile Electronic Vaccine Trials, an interdisciplinary project to improve quality of vaccine trials in low-resource settings

Jørn KLUNGSØYR1, Thorkild TYLLESKÄR1, Bruce MACLEOD2, Paul BAGYENDA3, Weiqin CHEN1, Peter WAKHOLI1

1 University of Bergen, P.O. Box, 7804, Bergen, Norway, Tel: +4755588560, Email: jorm.klungsoyr@cih.uib.no
2 University of Southern Maine, P.O. Box 9300, Portland, Maine 04104-9300 Tel: +1207780449, Email: macleod@usm.maine.edu
3 Makerere University, P.O. Box 7062, Kampala, Uganda Tel: +256-414-540628, Email: bagyenda@cit.mak.ac.ug

Abstract: Emerging international standards and regulations will in few years require complete electronic systems for management of vaccine trials. Clinical trials conducted in low-income countries need to have the same level of quality and reliability as comparable studies conducted in high-income countries. This will require data collection and management systems specifically designed and developed for these settings. Research data in low-resource settings are currently mostly collected on paper forms, a process which is susceptible to errors and inefficiency. The lack of control and compliancy to study protocol is a great challenge.

To solve this and related problems we will replace the paper based process with a completely digitized mobile system for conducting clinical trials based on EpiHandy and R. Researcher and field workers will use handheld computers and directly enter the collected information. This will drastically reduce the logistical challenges related to paper handling and digitization.

1. Relevance

We aim to develop a complete secure electronic system for data collection and management in vaccine trials, that handles data from source to publication using handheld computers and complying with international standards and regulations of clinical trials. It will be available at no cost including all generated software source codes and documents. We aim at improving the quality and cost-effectiveness of vaccine trial sites and epidemiological studies through transfer of technology and usage of new and emerging mobile technologies for data collection and management.

This is an interdisciplinary action research project that combines health and information technology expertise with field studies in Africa. The partners of this application have already developed several different systems: 1) EpiHandy, a generic electronic open source and free mobile data collection and management system, that is being used for regular studies in low-income countries, 2) R, a widely used and renowned open source statistical analysis tool, and 3) an online data publishing system.

The consortium consists of Norwegian and international organizations, universities, software and hardware development groups and includes TDR/WHO (TDR is a UN-
sponsored global programme for Research and Training in Tropical Diseases), HandheldsForHealth.org (a Linux company in Bangalore India), Makerere University in Uganda, INDEPTH Network / Malaria Clinical Trial Alliance, Promise Consortium and two faculties at University of Bergen, viz. Centre for International Health and InfoMedia.

Our consortium has extensive experience in conducting clinical trials, collecting data using handheld computers for regular studies, managing, publishing and protecting research data and developing software suitable for the context and challenges found in low-income countries.

1.1 Background and status of knowledge

There is a global trend and push to replace paper based systems with electronic systems to improve cost-effectiveness. An everyday example is the upcoming requirements making all airline tickets fully electronic in only a few years. This is also a trend affecting the field of vaccine trials and medical research.

Emerging international standards and regulations will in few years require complete electronic systems for management of a clinical trial [1]. Clinical trials conducted in low-income countries need to have the same level of quality and reliability as comparable studies conducted in high-income countries and must be conducted in accordance to international regulation. This will require secure, valid, accurate, comparable, efficient and cost-effective methods of data collection and management specifically designed and developed for settings with often poorly developed infrastructure and technical facilities.

With the rising costs and probability of failure in vaccine development, innovators often focus on candidates with potential high market return. Developing products targeted at important public health needs and less common diseases prevalent in low-income countries is becoming an increasing challenge [1]. FDA (Food and Drug Administration, US) have identified a number of obstacles in transforming the recent revolution and discoveries in biomedical science into more effective, more affordable and safe medical products. These obstacles along the critical path (see figure 1) to product development include efforts needed to streamline clinical trials and include development and compliancy to international standards and regulation and complete electronic submission of study data in the final approval processes [2]. FDA is in many ways a de facto definer of standards in conduct of clinical trials, and their standards are often rapidly adopted in most other countries.

![Figure 1: Critical path to product development](image)

FDA states that the development of new medical products “relies on the tools of the last century to evaluate this centuries' advances” [1]. This shows the need for new and innovative tools to reduce the potential stop points along the critical pathway from basic science to more timely, affordable and predictable access to new vaccines. The costs of the clinical trial phases (including I-III) are large, calling for new tools to reduce costs and improve effectiveness in data collection, management and distribution.

This project is in line with the goal of critical path research which “is to develop new, publicly available scientific and technical tools – including … clinical trial endpoints – that make the development process itself more efficient and effective and more likely to result in...”
safe products that benefit patients” [1]. Furthermore FDA states that “If we do not work together to find fundamentally faster, more predictable, and less costly ways to turn good biomedical ideas into safe and effective treatments, the hoped-for benefits of the biomedical century may not come to pass, or may not be affordable.” [1]

Research data in low-resource settings are currently collected on paper forms during the interview/investigation process. The papers are collected and transported to a central site where data are manually entered into a database and later analyzed. This process is susceptible to error and inefficiency. First there is no validation of data at point of collection; they are only checked for consistency when digitized. Errors found impose a resource intensive and logistically challenging process/loop of sending forms back, identifying the source, validating the data, sending the forms and finally digitizing the corrected data (see figure 2). There is often a long delay between data collection and availability of datasets ready for analysis.

Compliance to study protocol and ensuring correct linking between study subjects and the data collected at every time point is extremely important. For example, the clinical trial HIVNET 012 which tested Neveriapine, an antiretroviral drug to prevent mother-to-child transmission of HIV, was not approved as some paper forms were missing, not because the drug was judged dangerous [3, 4]. In low-resource settings official and usable identification papers like ID cards and social security numbers are lacking, this imposes a potential for mix-up of records as in the case of the HIVNET 012 study and could seriously influence the statistically analysed outcome of a study. Multi-site/country studies require 100% compliance and compatibility between protocols at the different sites in order to be able to compare and conduct pooled analyses. With more or less compatible and standardized systems this quickly becomes a great obstacle for such undertakings. It is surprising how many research projects use self-made and simple systems, e.g. based on Microsoft Access, to handle large quantities and values of data, which is an extremely risky approach.

An obstacle to cross site and country collaboration is lack of controlled and efficient access to anonymized analytical datasets and source documents for internal inspections and review process during publication. There is a lack of unambiguous verification of data sources and originality of data for a particular dataset. There are no good publicly available means of tracking a dataset related to a published article that also can clearly show versions and meta data on where, when, by whom, from whom, in what context and for what purpose it was collected.

There is a need for strict control of data sources, from point of collection and all the way to final publication, including audit trails and certificates of authenticity and origin (when, where, from, by whom etc.). The situations (or cases) call for secure, valid, accurate, comparable, efficient and cost-effective methods of data collection and management in order to produce high quality and reliable data that can be used for vaccine trials and thereby virtually any type of field research.

Recent advances in technology and software presents an opportunity to develop a system capable of solving many of the outlined obstacles and thereby preparing and enabling research institutions in low-income countries without the extreme costs related to commercial and inappropriate solutions. A number of studies in high-income countries show the improved quality, timeliness, effectiveness and usability of handhelds in clinical settings [5-7], in addition there are studies from low-income countries showing positive results of handheld computer usage in general [8, 9].

There is now a window of opportunity to prepare and enable research institutions in low-income countries for the upcoming revolution in the way clinical trials are run.

1.2 Relevance to society

This project is an action research oriented project focusing on developing new and secure methods for field data collection and fully electronic conduct of clinical trials in low-income
countries. The contributions to new knowledge will be the development of methods and tools, but also the long term improvement in quality and reliability of research conducted in low-resource settings. This can potentially make new treatments and medicines available to low-income countries that otherwise would not have been possible to fully investigate.

1.3 Environmental perspectives

The project will produce a system than can drastically reduce usage of paper, transportation and storage costs and waste problems of large volume printing in research studies. This project will as far as possible use electronic means of communication and collaboration to reduce the need of face to face meetings requiring long distance flights.

![Figure 2: Data collection process](image1)

![Figure 3: Simplified process by mobile data collection](image2)

2. Approaches, hypotheses and choice of method

**Go for Action:** This project is a multi disciplinary project that actively seeks to change the way a vaccine trial is conducted and thereby fits in the realm of Action Research (AR). AR is suitable for introduction and development of new technologies for improving processes within an organization [10]. “Action Research aims to contribute both to the practical concerns of people in an immediate problematic situation and to the goals of social science by joint collaboration within a mutually acceptable ethical framework” [11]. The actual software development methods will be based on a mixture of existing methods with primary focus on Agile methods and Rapid Prototyping to create an iterative development which will ensure a better end product through numerous possibilities of testing and input from end users.

**Go Empower:** This project aims at enabling research institutions in low-income countries to conduct high quality clinical trials and research by supplying a free and open source system for complete electronic data collection and management using handheld computers for on the spot validation and utilization of existing tools for distribution, sharing and validation of datasets in the publication phase. Enabling and strengthening research institutions in low-income countries will open for new constellations and possibilities of collaboration on more equal terms than what has been possible up to now.

Proceedings of M4D 2008, Karlstad University, Sweden – 39
Go Mobile and Electronic: Our ultimate goal is to eliminate the paper based process with a completely digitized mobile system for conducting clinical trials, and thereby virtually any type of field research (figure 3 showing simplification). This project will use mobile technology to improve the quality and reliability of clinical trials and increase the speed, efficiency and cost-effectiveness of information collection and thereby also the applicability of clinical trials in low-resource settings. Researcher and field workers will use handheld computers and directly enter the collected information on electronic forms. This will drastically reduce the logistical challenges related to paper handling and digitization. The system will also support alternative methods of data collection including paper and web based forms. This could look like a contradiction, but a system of this magnitude must be capable of handling the different scenarios that are likely to exist. This will allow research sites to run different studies on the same system using different modes of data collection.

Go for International Standards: This project is based on emerging international and accepted standards for and ongoing efforts in development of data collection and management systems and collection. This interdisciplinary research project seeks to develop specifications and requirements for a system based on the EU, FDA and GCP (Good Clinical Practice) regulations, for electronic data collection and management in clinical trials and existing standards developed by HL7 (Health Level Seven – ANSI/ISO organization) and CDISC (Clinical Data Interchange Standards Consortium). The later was established by ICH (The International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use). HL7 and CDISC have developed a number of standards that have now been merged into a joint initiative called BRIDG (Biomedical Research Integrated Domain Group). These standards are the foundation for future electronic submission of clinical trials. Another important standard related to data documentation is the Data Documentation Initiative (DDI) which contains data documentation for a specific dataset.

Conforming to standards is critical to ensure widespread acceptance of the system. The software requirements for electronic data in clinical trials are particularly rigorous. Key elements are: audit trail, electronic signatures, security and storage of data, review and inspections, standard operating procedures, validation and system documentation [14-17]. There is a need for a thorough review of these regulations including publications and their implications. Such a review will lay the foundation for the further development.

Go Open Source: EpiHandy is a free and open source data collection and management system (see figures and www.epihandy.com) that has been developed by this team over the last years. EpiHandy is a generic system for designing and collecting data on electronic multilingual forms. It uses handheld computers with validation of data input at the point of collection and GIS for tracking of data origin. EpiHandy is being used for ongoing studies in the Promise Consortium and several other multi country/centre studies in low-income countries. This project is not starting from scratch, but focuses on further developing EpiHandy to comply with relevant regulations and standards related to electronic data collection and management in clinical trials. The system will also be made to run on a multitude of platforms and operating systems by using established cross-platform development tools and frameworks.

Go for Security: Ensuring unambiguous and secured linkage between study subjects, protocol events and case report forms (eCRF) is a crucial part of this project. We will implement a quadruple redundant linking scheme to protect against loss of this linkage (see figure 4). All subjects, protocol items and case report forms will embed the linkage between these entities.

Go Analyse: By adapting and linking to R[12], a renowned and free open source software environment for statistical computing and graphics, we will integrate powerful data analysis tools and functions into OMEVAC. Several subprojects of R will be of particular interest including R package for Epidemiological Data and Graphics[13]. R is developed by
the R Foundation and supported by a high number of renowned institutions. The combination of secure mobile data collection and integrated analysis functionality will be a powerful tool for the trials sites.

**Go Publish:** This project will not only secure the data collection and management at site level, but also handle it all the way through to publication and secure storage over longer periods of time. By alleviating and adapting online publishing systems for data sharing and distribution, a complete and secure chain can be established to increase the trustworthiness and reliability of data collected in low-resource settings. We will setup a server that will enable free publishing, archiving and sharing of data from research projects focusing on important global health problems in low-income countries.

**Go user-fee free:** There exists commercial systems that can manage some parts of what is required, however these are prohibitively expensive, require extensive training and maintenance by professional system developers and are designed for settings with very good infrastructure and thereby not suitable to solve the problems described earlier.

Even if the project is aiming to solve the challenges related to data management in low-resource settings it will also be usable for similar studies in high resource settings as it solves many of the problems of conducting clinical trials in general.

The return on the investments in this project will be in the form of savings at the different sites eventually implementing the completed system.

### 3. System Development

The foundation of this system will be the existing large code base of EpiHandy. We will further design and develop EpiHandy to be compliant with regulations on data security and integrity in health research as described earlier. The development process will be carried out using Agile and Rapid Prototyping methodologies, which are iterative software development methods with each iteration including design, development, and testing (black box testing, functional testing and user tests). This work package will involve a large number of developers from many of the project partners.

#### 3.1 Study Protocol Management Module (SPMM)

To ensure a well conducted clinical trial, a study protocol has to be carefully designed. The protocol acts as a guide for stakeholders, participants, institutional review boards and ethics committees which approve whether the trial can be initiated. The protocol design process is one of the most crucial steps in a clinical trial as defines the why, how, when and who’s who of the study. Paper based studies are prone to protocol inconsistency and might lead to errors not easily detected until late in the study and thus cause major drawbacks in both trial schedule and the validity and reliability of outcome of the results. In extreme cases, a poorly defined protocol may impose health risks to the involved subjects.

SPMM will enable research groups to collaborate on developing a study protocol that once complete can be immediately utilized in any data collection system complying to the

---

*Proceedings of M4D 2008, Karlstad University, Sweden – 41*
Trial Design Model (TDM) developed by CDISC (Clinical Data Interchange Standards Consortium). SPMM will include functions for versioning, commenting and approval processes and amendments. The system will as for the other parts not require any programming skills and should be intuitive and simple to use, but contain powerful functionalities to cater for virtually all types of study protocols.

3.2 Secure Person Identification Module (SPIM)

In low-resource settings official and usable identification papers like ID cards and social security numbers are lacking, this imposes a potential for mix-up of records. Wrong linkage by i.e. typing in the wrong ID number can be disastrous. The first step needed before starting collecting data for a case report form is to uniquely and positively find the identity of the study subject for whom data is to be collected. Through SPIM we will actively investigate potential solutions to these problems. This includes investigating usage of biometrics, imaging techniques, GIS (Geographic Information Systems) and study subject context data such as family relations, address, gender, age and more. Different algorithms and matching patterns will be developed to positively identify a study subject before collecting any piece of information. We will explore the ethical and security related issues to avoid potential misuse of sensitive information and implement techniques to prevent identity disclosure.

3.3 Data analysis by linking to R

R is a free open source language and environment for statistical computing and graphics. R provides a wide variety of statistical (linear and nonlinear modelling, classical statistical tests, time-series analysis, classification, clustering, etc.) and graphical techniques, and is highly extensible [12]. We will create a unique linkage between EpiHandy and R. This will enable OMEVAC to utilize the powerful statistical tools and functions available in R.

3.4 Data publishing and sharing Module (DPSM)

This project will go beyond the data collection and management at site level and create a unique linkage between EpiHandy and an online data repository software. It will include functionality for publishing and sharing datasets over internet, including display of aggregated and micro level data with a number of statistical and analytical options. We will be able to establish a complete electronic chain of data management from point of collection all the way through to publishing of datasets for internal and external usage. Data access permissions are easily managed and every dataset will be uniquely identified to facilitate version control and attribution of data sources.

This feature will be an important capability to enable verification of originality of a dataset and thereby the ability to discover constructed data claiming to be from a certain study [18]. It will also make it possible to give reviewers restricted access to the datasets used in a particular paper.

3.5 System Validation and Field Testing

Technical validation of the software is a prerequisite for using the completed system in a real clinical trial. Technical validation includes code review, regression testing, testing on existing data and scenario testing. This is required in order to prove the validity, consistency, stability and security of the system prior to real life testing and deployment.

We will conduct formative evaluations of the OMEVAC system through real life field test in research sites of the PROMISE Consortium and INDEPTH/MICTA. The validation will include side by side comparison and testing with both the current paper based systems and a commercial system. The Health Technology Assessment (HTA) framework [19] will be used as a foundation for the evaluation of OMEVAC. HTA includes both qualitative and quantitative approaches and focuses on the users, technology, economy and organization of the system.
4. Project partners

Centre for International Health, University of Bergen (www.cih.uib.no) has developed a number of different systems for data collection and management in low-resource settings. This includes EpiHandy a generic survey and forms tool (www.epihandy.com), WHO Anthro 2005 a tool for child growth monitoring developed for and published under the logo of World Health Organization (WHO) (www.who.int/childgrowth/software) CIH has established and renowned master and PhD program in global health research where students can get training.

TDR, UNICEF/UNDP/World Bank/WHO - Special Programme for Research and Training in Tropical Diseases (www.who.int/tdr). TDR has extensive experience in research and clinical trials and are heavily involved in research, training and supporting research and governmental institutions in low-income countries

Promise Consortium (www.promiseresearch.org) is a consortium running clinical trials in 4 countries in Africa. The experience and expertise of the Promise Consortium act as domain experts/advisors and one of the sites will be used as the primary testing and development site.

INDEPTH Network (www.indepth-network.org) is a network of more than 35 demographic surveillance sites in 19 countries in Africa, Asia, Central America and Oceania, the network is headed from Ghana. MCTA – Malaria Clinical Trial Alliance (www.indepth-network.org/mcta/mctaindex.htm), a project of INDEPTH, is actively involved in developing new vaccines against malaria. Clinical trials within an INDEPTH site has the potential giving more insight than a regular clinical trial as many of these sites have historic data at the individual level up to 30 years back.

InfoMedia, University of Bergen (Department of Information Science and Media Studies, UoB – www.infomedia.uib.no) is an academic institution at UoB. InfoMedia has been actively involved in the development and evaluation of.

Department of Computing, University of Maine, USA (http://www.usm.maine.edu/ ). Ass. Prof. Bruce MacLeod has been developing systems for data collection and management in low-resource settings in the past 2 decades including HRS and MobileHRS.

Faculty of Computing and IT (FCIT), Makerere University, Uganda (www.cit.ug) has been actively involved in the development of EpiHandy and several spin-off solutions. The Department of Software Development & Innovations (DSDI - www.cit.ac.ug/disd ) is an establishment of the Faculty that will drive FCIT’s engagement in applied software research and the application of ICT.

HandheldsForHealth.org / Encore Software (www.handheldsforhealth.org / www.ncoretech.com ) is a company based in India which focuses on developing hardware and software solutions to solve common problems related to data collection and management in health services in low-income countries.

References


44 – Proceedings of M4D 2008, Karlstad University, Sweden