

## TITLE PAGE

JAMIA – REVISION 2, clean version

# Managing research and surveillance projects in real-time with a novel open-source eManagement tool designed for under-resourced countries

Andreas Steiner <sup>1 2 3 \*</sup>, Jerry Hella <sup>1 2 3 \*</sup>, Servan Grüninger <sup>1 2 3</sup>, Grace Mhalu <sup>1 2 3</sup>, Francis Mhimbira <sup>1 2 3</sup>, Colin I. Cercamondi <sup>4</sup>, Basra Doulla <sup>5 6</sup>, Nicolas Maire <sup>1 2 3</sup>, Lukas Fenner <sup>1 2 3 7</sup>

**1** Ifakara Health Institute, Tanzania

**2** Swiss Tropical and Public Health Institute, Basel, Switzerland

**3** University of Basel, Basel, Switzerland

**4** Laboratory of Human Nutrition, Swiss Federal Institute of Technology, Zürich, Switzerland

**5** Central Tuberculosis Reference Laboratory, Dar es Salaam, Tanzania

**6** National Tuberculosis and Leprosy Programme, Tanzania

**7** Institute of Social and Preventive Medicine, University of Bern, Switzerland

\* Contributed equally to this work

## Correspondence:

Dr. Lukas Fenner, MD MSc

Swiss Tropical and Public Health Institute, Basel, Switzerland;

University of Basel, Switzerland

E-Mail: [lukas.fenner@unibas.ch](mailto:lukas.fenner@unibas.ch)

## **Running head:**

Monitoring research projects in resource-constrained countries in real-time

## **Word count, inserts and supplementary information:**

Word counts: abstract 248 words, main text 3,271 words

Inserts: 6 figures, 18 references

Supplementary Information: 5 files

## ABSTRACT

**Objectives:** To develop a software tool to facilitate data entry and to monitor research projects in under-resourced countries in real-time.

**Methods:** The eManagement tool *odk\_planner* is written in the scripting languages PHP and Python. *odk\_planner* is lightweight and uses minimal internet resources. It was designed to be used with the open-source software Open Data Kit (ODK). The users can easily configure *odk\_planner* to meet their needs, and the online interface displays data collected from ODK forms in a graphically informative way. *odk\_planner* also allows users to upload pictures and laboratory results, and sends text messages automatically. User-defined access rights protect data and privacy.

**Results:** We present examples from four field applications in Tanzania successfully using the eManagement tool: 1. Clinical trial. 2. Longitudinal Tuberculosis (TB) Cohort Study with a complex visit schedule, where it was used to graphically display missing case report forms, upload digitalized X-rays, and send text message reminders to patients. 3. Intervention study to improve TB case detection, carried out at pharmacies: a tablet-based electronic referral system monitored referred patients, and sent automated messages to remind pharmacy clients to visit a TB Clinic. 4. TB retreatment case monitoring designed to improve drug resistance surveillance: clinicians at four public TB clinics and lab technicians at the TB reference laboratory used a smartphone-based application that tracked sputum samples, and collected clinical and laboratory data.

**Conclusions:** The user-friendly, open-source *odk\_planner* is a simple, but multi-functional, web-based eManagement tool with add-ons that helps researchers conduct studies in under-resourced countries.

**Key words:** Open Data Kit; open-source software; electronic data entry; data management tool; paperless; real-time monitoring

## INTRODUCTION

In under-resourced countries, paper forms are still frequently used to collect research data(1). Paper records are expensive to prepare, distribute, store and archive. Paper is easy to lose or destroy, and difficult to replicate. Now that affordable mobile electronic devices like tablet computers and smart phones are on the market, some of these drawbacks can be overcome using electronic data collection systems(2-6). These inexpensive electronic devices are especially useful in resource-constrained settings where lack of reliable infrastructure poses major obstacles to standard data collection, and where mobile phone use and coverage have surged in recent years(7-9).

Existing open-source software tools can serve as a base for developing data collection software to meet specific needs. For example, *Open Data Kit* (ODK)(10) is based on open interfaces and standards, speeds up data collection, and reduces opportunities for errors. OKD project has started in 2008 directed by a core team at the University of Washington, Seattle, and was then further developed by a growing community of developers and users(11). ODK's modular and customizable architecture attracts users(10, 11), and has sparked an active and growing ecosystem of service providers and extensions(11, 12). Briefly, the module *ODK Collect* allows users to collect data on Android-based smart phones or tablets; *ODK Aggregate* facilitates data storage and transfer, and provides access to the stored data. Study staff need no programming skills to create electronic questionnaires with the tool *XLSForm*. The forms are based on Excel templates, which are then transformed into *XForms* (the standard form supported by ODK)(11, 12).

ODK tools include utilities to collect and store data, but so far there are only limited options for viewing, monitoring or managing data. We thus decided to build on

the ODK software platform and develop a lightweight, user-friendly and open-source tool designed for under-resourced countries to make data management easier and allow real-time monitoring. In particular, we addressed the following key requirements: i) intuitive overview function of data in real-time; ii) user-friendly and easy to configure; iii) incorporates non-ODK data into the same data base; iv) automatically sends text messages and reports about data completeness by email; and v) implements rigorous data protection and preserves confidentiality.

## DEVELOPMENT AND FUNCTIONALITIES OF “odk\_planner”

### General description and key features

Figure 1 gives an overview of the modular structure and the main functionalities of the eManagement tool “odk\_planner”. The ODK Aggregate server receives data sent from Android-based mobile devices (tablets, smart phones), and stores it safely in a relational database. odk\_planner is run on the same server as ODK Aggregate, and accesses the same database. odk\_planner presents collected data in a clear and easily understandable format (“overview table”, Figure 2 and 3), and allows different data sets to be intuitively categorized (e.g. cases vs. controls). The admin user determines appearance and rules for reminder messages (see also, Figure 3; Figure S1, Supplementary Material). The admin user also defines fine-grained access rights for each user (e.g., investigator, fieldworker, and secretary) (File S2, Supplementary Material). odk\_planner has additional functions, including upload of external data like X-ray images (“odk\_pusher”, Figure 4), and automatically sending text messages (Figure 5, and File S3, Supplementary Material). Furthermore, odk\_planner integrates two other features: a user-friendly Quick Response (QR) barcode labeling system for patient files and sample containers, and an X-ray viewer that gives the user options for zooming, rotating, and modifying the contrast of the uploaded images (Figure 4). The source code and a detailed manual for odk\_planner are available online ([https://github.com/SwissTPH/odk\\_planner](https://github.com/SwissTPH/odk_planner) and <http://odk-planner.readthedocs.org>). The source code comes with a complete test suite that uses WebDriver to assess functionality of every module and facilitates further development of the odk\_planner by software developers.

### System prerequisites

The server component of odk\_planner was tested on Linux, OS X, and Windows and depends on a working ODK Aggregate installation (tested against v1.3.2; including the corresponding MySQL database), a web server (tested with Apache), and PHP (tested with version 5.5). No additional server infrastructure is needed for odk\_planner, but if server administrators want to create new instances of odk\_planner for testing purposes as well as to use odk\_pusher, they must install Python. Once set up, it can be accessed and used from any computer or mobile device with a web browser. The data can be sent to the server via cellular network or via Wi-Fi.

## System configuration

While the server administrator must install odk\_planner and create different instances, a single odk\_planner instance can be configured via a web interface. The configuration is based on a single file in .xls format ([File S1, Supplementary Material](#)). Earlier versions are stored as backups. The configuration file consists of the following categories:

- 1. Settings:** The *settings* sheet contains general information about the overall structure and appearance of the ODK database when it is accessed via the odk\_planner. Settings is where the title for the specific odk\_planner instance is defined, along with the variables used to link different forms together (like a primary key variable in an SQL database).
- 2. Users:** Specific user names, passwords and access rights for a variable number of study staff members are defined on the *users* sheet.
- 3. Overview:** Forms can be sorted into sub-groups by one or several specific characteristics on the *overview* sheet. For example, sub-headings can be defined for different study sites or patient categories.



**4. Colors:** The *colors* sheet provides the user with a wide array of functions to customize the app's appearance in the web browser. The user can define color and formatting styles to highlight specific forms so that progress in complex studies is clearly visible, based on the presence or absence of a form or based on the content of the forms.

**5. Text messages (SMS):** The *SMS* configuration sheet can be used to define an SMS schedule: rules defined by the researcher will trigger automated text messages. The sheet contains login credentials for an online SMS service provider (e.g., <https://bulksms.vsms.net/>) to send short messages to study participants or staff through the application program interface (API) of the provider ([Figure 5](#)). Text messages may have predefined content, or may display content that is extracted directly from the existing study database; for example, short messages can be combined with laboratory results ([File S3, Supplementary Material](#)). Please note that for confidentiality reasons messages should contain only anonymized and non-sensitive information.

**6. Reports:** The *reports* sheet can be configured to send out reports of specific or missing forms to a predefined email address and at a predefined interval.

**7. Cron:** The time-based job scheduler *Cron* contains the configurations for automated and regular execution of commands that send email reports and text messages.

## The web-based interface “overview table” and the application

### “odk\_pusher”

**overview table.** odk\_planner will recognize the same study form descriptions in .xls format that were used to generate the forms for ODK Aggregate using *XLSForm* ([Figure 2](#)). The online interface overview table enables the user to view all forms entered into the system and uploaded to ODK Aggregate ([Figure 3](#)). All forms

of a patient are shown in a row, and every form has its own column. The date when the corresponding form was entered is visible at each intersection of rows and columns. The coloring and formatting patterns in the *colors* sheet of the configuration file will show additional information about the progress of the study. For example, forms that have not been filled in within a pre-defined time period can be highlighted in red. When users click on a specific form, they will see corresponding data entries, pictures or X-ray images (depending on the access level, see above). Users can send short messages to study participants either manually, via the online user interface, or automatically if specified in the configuration file. Users can retrieve a list of missing or delayed forms either manually, by downloading the corresponding .csv file, or by automated email reminder, as defined in the configuration file. Specific data categories for different access levels can be defined in an additional “access” column (see also paragraph on data confidentiality below, and [File S2, Supplementary Material](#)). Only study staff with pre-defined access to sensitive information (such as HIV information) will be able to see these data points ([Figure 3 B](#)).

**odk\_pusher.** The odk\_pusher is a tool bundled with odk\_planner, programmed in Python 3. odk\_pusher provides a simple API, and comes with two Graphical User Interface (GUI) programs: “Xray uploader” and “MS-SQL uploader”. With odk\_pusher, users can upload any data, including images, in a semi-automated way. Xray uploader allows the user to wrap digitized X-ray pictures (or other image files) into an *XForm* and upload them to an ODK Aggregate instance ([Figure 4](#)). The Xray uploader integrates a configurable image-compression on-the-fly to save space and bandwidth. The MS-SQL uploader lets the user upload data from a MS-SQL database to an Aggregate instance. In both cases, the GUI and of the upload parameters can be configured via a JSON structured text file.

## **Confidentiality and security of data**

To meet Good Clinical Practice (GCP) guidelines, the database must only contain anonymous patient IDs. The patient log file that links patient ID with patient identity should be stored separately, offline, at a secured location accessible only to the Principal Investigator. Communication between the server and web browser uses industry standard SSL-encryption. The data on the server is not encrypted, and security relies on the fact that the server is inaccessible to unauthorized parties. ODK Aggregate is password protected, but offers no finer grained access control. However, odk\_planner can be configured with different levels of access rights ([File S2, Supplementary Material](#)). All odk\_planner users have an individual password, and users are only granted access to specific data ([Figure 3B](#)).

## FIELD APPLICATIONS OF THE “odk\_planner”

To demonstrate the functionality of odk\_planner in the field, we present below four studies from Tanzania that used or are using the newly developed tool ([Figure 6](#)). All studies were approved by the National Ethics Committee in Tanzania (National Institute for Medical Research).

### **“odk\_planner” in support of a clinical trial on iron absorption and utilization during tuberculosis treatment (TB-IRON)**

odk\_planner is being used in an ongoing investigator-initiated clinical trial about iron-associated anemia in TB patients (trial registration no. NCT02176772, <https://clinicaltrials.gov/ct2/show/NCT02176772>). The trial uses stable iron isotopes to characterize iron metabolism over six months of TB treatment. The trial has a complex visit schedule, including oral and intravenous administration of stable iron isotopes before, during and after TB treatment. Data is captured at the clinic, at the hematology and helminth laboratory (see schematic overview in [Figure 6, Panel A](#)). Automated SMS reminder messages are sent to study participants for follow-up visits. An overview of the study progress and detailed data (e.g., full blood counts) are visualized on odk\_planner, and regularly accessed by the international study team in Switzerland and Tanzania. Study staff (clinicians, coordinators, lab personnel, internal monitor) have different levels of access to the data. The study team and the internal monitor stated that reviewing the data and tracking the progress of the study on odk\_planner is useful and convenient.

### **“odk\_planner” in support of an observational study of tuberculosis cases and controls (TB-DAR)**

In 2013, we initiated a cohort study of adult tuberculosis (TB) patients in Dar es Salaam (Tanzania) which collects socio-economic and clinical data from patients and household controls, as well as clinical specimens (see schematic overview in [Figure 6, Panel B](#)). Clinical and laboratory data are entered directly at the hematology and helminth laboratories on touchscreen tablets ([File S4, Supplementary Material](#)). The follow-up schedule of study participants stipulates several follow-up visits. Study participants are automatically reminded of scheduled visits via the SMS reminder system ([File S5, Supplementary Material](#)). In Figure 3, red highlights indicate overdue data entry forms to alert study coordinators. Access levels are granted based on the needs of the study staff; e.g., auxiliary staff (nurse assistants, field workers) can only collect patient data, while medical staff involved in patient management (doctors, clinical officers) have full access to data, including HIV status. Chest X-ray images are digitized and automatically uploaded to the server using odk\_pusher ([Figure 4](#)). QR barcode labels are used for all study procedures to minimize the risk of mixing up patient files and biological specimens. The local study team was rapidly familiar with odk\_planner, and is using the tool in the daily work to identify incomplete data entries and to review laboratory results for patient care. SMS reminders were well accepted by study participants leading to an optimal patient retention (currently less than 20% loss to follow-up at 6-month).

#### **“odk\_planner” in support of an intervention study of intensified tuberculosis case finding at pharmacies with electronic real-time tracking (TB-PHARM)**

This ongoing intervention study aims to identify TB cases through pharmacies by referring pharmacy clients suspected of having TB to a TB clinic for diagnosis (see schematic overview in [Figure 6, Panel C](#)). Pharmacy clients receive a registration number on a card (with QR code), which they can bring to a nearby treatment center. Local pharmacists use tablets with internet bundles to enter the registration number.

At the TB treatment center, a study doctor registers referred clients. The study staff reported that odk\_planner was an intuitive tool for the study doctor, who could track the progress of client registration at each pharmacy and referrals at the treatment center. After a pre-specified time period, the automated SMS reminder system sent short messages to pharmacy clients with TB symptoms if they had not yet visited the health care center for diagnosis. When interviewed at the clinic, referred clients reported that they had read the SMS message and felt encouraged.

**“odk\_planner” in support of monitoring retreatment tuberculosis cases in real-time at the level of clinics and laboratories for improved drug surveillance (TB-R pilot study)**

According to national guidelines, all sputum samples from retreatment TB patients should be sent in to the Central Tuberculosis Reference Laboratory in Tanzania to be tested for drug susceptibility, but over 90% of retreatment sputum samples never reach the laboratory (13). This pilot study tested a real-time tracking system, choosing odk\_planner as a tool to identify crucial steps in the progress of sputum samples and laboratory results (see schematic overview in [Figure 6, Panel D](#)). We designed six forms to track every step in the process. First, ODK on Android smart phones (Samsung S6500 Galaxy Mini 2) was used to register TB retreatment cases at TB clinics. A second form was filled out when the sputum samples were sent to the laboratory. When the sputum arrived at the reference laboratory, the staff filled out a third form to confirm receipt. The laboratory results were reported with two additional ODK forms: the fourth form for rapid results (smear microscopy, GeneXpert MTB/RIF), and the fifth form for results that took longer (culture, drug susceptibility testing). All results were sent with an automated text message to the TB clinic coordinator with a unique patient identifier ([File S2, Supplementary Material](#)). Information was entered into the sixth and final ODK form when laboratory results

arrived at the TB clinic. The study teams at the clinics appreciated the prompt sending of laboratory results. At the reference laboratory, the staff could see when a sputum sample was sent to them, and could determine how samples and results were lost. The involved study staff from the routine settings reported no major difficulties with the system, even if they had no IT skills.

## DISCUSSION

Though modern communication technology and computer-like devices are now widely available in many under-resourced countries, internet access is often not continuous and applications need to be simple and robust. odk\_planner is a comprehensive, powerful, yet lightweight, and user-friendly data eManagement tool that extends the ODK software package, to monitor research projects in under-resourced countries in real-time. odk\_planner gives users full control, with maximum flexibility, offering a wide range options to customize the appearance and management of data. odk\_planner makes use of already existing server and database infrastructure. If ODK Aggregate is already in use, odk\_planner imposes no additional IT burden once it is installed within the server instance. The web-based tool needs only minimal internet resources. It can be used on any Android-based mobile phone device; odk\_planner needs only a small internet bandwidth available in presumably all regions and countries. No programming skills are required for its use.

The application odk\_pusher as part of odk\_planner integrates different data collection modalities in a straightforward way. The data will be shown exactly the same way as the data uploaded using ODK Collect. Because of its automated reminder system, odk\_planner is especially useful for interventional studies, for studies that demand participants and/or study staff to adhere to a specific time-schedule, or for studies in which temporal progress depends on the speed at which certain results can be communicated to participants and/or study staff. The SMS reminder system can send information about specific tasks that need to be completed, or provide people with information about study results or the progress of the study.



There are other data collection and management solutions available, such as OpenMRS(14), EpiCollect(15), the REDCap project(16), and the ODK-based LINKS system(17), but they offer only limited options for presenting and managing collected data, and additional features, such as automated text messages, are limited or not freely available. For example, the REDCap (Research Electronic Data Capture) project offers electronic data capture tools, and is widely used in high-income and also increasingly in low-income countries. REDCap is a mature and user-friendly system, but it does not support additional features such as X-ray image upload or text message reminders. It is also not possible for the user to initiate modifications or extensions.

The ODK-based LINKS system extends ODK, as odk\_planner does. LINKS is a cloud-based central database with a web interface to manage data collected by ODK(18). However, there are four main differences between LINKS and odk\_planner: 1) odk\_planner is open-source and can be modified and adapted by anyone; 2) odk\_planner can be hosted locally at an institution, without having to transmit the data into a cloud database; 3) odk\_planner requires no registration; 4) odk\_planner has extra features like sending automated text message, and integrating non-ODK data.

Although odk\_planner has been used only in a few projects so far, we demonstrated the usefulness and flexibility of odk\_planner in different study designs and settings. As with any other data management tool, odk\_planner must strike a balance between building a flexible system and limiting the complexity of the configuration process. Because odk\_planner extends ODK (which makes it attractive to the ODK community), this modular approach (set up ODK Aggregate, then install odk\_planner) makes installation no more complicated than a single executable system. Of note, there is currently no encryption at rest of data on the server. Finally,

we do not yet know how odk\_planner is evaluated by non-technical study staff outside of typical research settings, but our pilot study on TB retreatment patients indicates that it is also promising in a routine surveillance system.

Future developments and applications may include more automated data transfer tools (e.g., upload of full blood count and other laboratory results), alive view of data (e.g., figures for ratios, graphs for time series, maps for geographical data), and a virtual machine bundled with ODK Aggregate which further simplifies the use in a single download. We also hope the scientific and user community will improve and extend odk\_planner.

In conclusion, the field applications in different settings and with different designs clearly show that odk\_planner is an excellent eManagement tool for scientific studies conducted in under-resourced settings. User-friendly, with a variety of functions, it is a broadly applicable tool for managing clinical trials, observational, and intervention studies, and can also be used to improve routine medical care or accelerate collection and delivery of clinically relevant information in routine care.

## REFERENCES

- 1 Le Jeannic A, Quelen C, Alberti C, Durand-Zaleski I, CompaRec I. Comparison of two data collection processes in clinical studies: electronic and paper case report forms. *BMC Med Res Methodol.* 2014;**14**:7.
- 2 Zhang S, Wu Q, van Velthoven MH, et al. Smartphone versus pen-and-paper data collection of infant feeding practices in rural China. *J Med Internet Res.* 2012;**14**(5):e119.
- 3 Wilcox AB, Gallagher KD, Boden-Albala B, Bakken SR. Research data collection methods: from paper to tablet computers. *Med Care.* 2012 Jul;**50 Suppl**:S68-73.
- 4 King JD, Buolamwini J, Cromwell EA, et al. A novel electronic data collection system for large-scale surveys of neglected tropical diseases. *PLoS One.* 2013;**8**(9):e74570.
- 5 Giduthuri JG, Maire N, Joseph S, et al. Developing and validating a tablet version of an illness explanatory model interview for a public health survey in Pune, India. *PLoS One.* 2014;**9**(9):e107374.
- 6 Rajput ZA, Mbugua S, Amadi D, et al. Evaluation of an Android-based mHealth system for population surveillance in developing countries. *J Am Med Inform Assoc.* 2012 Jul-Aug;**19**(4):655-9.
- 7 Porter G, Hampshire K, Abane A, et al. Youth, mobility and mobile phones in Africa: findings from a three-country study. *Information Technology for Development.* 2012 2012/04/01;**18**(2):145-62.
- 8 Rotberg RI, Aker JC. Mobile Phones: Uplifting Weak and Failed States. *The Washington Quarterly.* 2012 2013/02/01;**36**(1):111-25.
- 9 Porter G. Mobile Phones, Livelihoods and the Poor in Sub-Saharan Africa: Review and Prospect. *Geography Compass.* 2012;**6**(5):241-59.

- 10 Anokwa Y, Hartung C, Brunette W, Borriello G, Lerer A. Open Source Data Collection in the Developing World. *Computer*. 2009;**42**(10):97-9.
- 11 ODK. <http://opendatakit.org/>.
- 12 Wiki of ODK developers. <https://github.com/opendatakit/opendatakit/wiki>.
- 13 Kilale AM, Ngowi BJ, Mfinanga GS, et al. Are sputum samples of retreatment tuberculosis reaching the reference laboratories? A 9-year audit in Tanzania. *Public Health Action*. 2013 2013;**3**:156-9.
- 14 Wolfe BA, Mamlin BW, Biondich PG, et al. The OpenMRS system: collaborating toward an open source EMR for developing countries. *AMIA Annu Symp Proc*. 2006:1146.
- 15 Aanensen DM, Huntley DM, Feil EJ, al-Own F, Spratt BG. EpiCollect: linking smartphones to web applications for epidemiology, ecology and community data collection. *PLoS One*. 2009;**4**(9):e6968.
- 16 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. *J Biomed Inform*. 2009 4/2009;**42**(2):377-81.
- 17 Pavluck A, Chu B, Mann Flueckiger R, Ottesen E. Electronic data capture tools for global health programs: evolution of LINKS, an Android-, web-based system. *PLoS Negl Trop Dis*. 2014 Apr;**8**(4):e2654.
- 18 LINKS. [www.linkssystems.org](http://www.linkssystems.org).

## **Acknowledgements**

We are indebted to the open-source community such as the core ODK team. We are also grateful to the IT unit at the Swiss Tropical and Public Health Institute for their support. We would also like to thank all study teams which pilot tested and substantially improved the data management system. We thank Kali Tal for her editorial support.

## **Sources of support**

The authors received no project funding specifically for this work. AS, SG, JH, FM and LF were partially supported by the Rudolf Geigy Foundation (Swiss Tropical and Public Health Institute, Basel, Switzerland). The content is solely the responsibility of the authors and does not necessarily reflect the official views of the funders. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

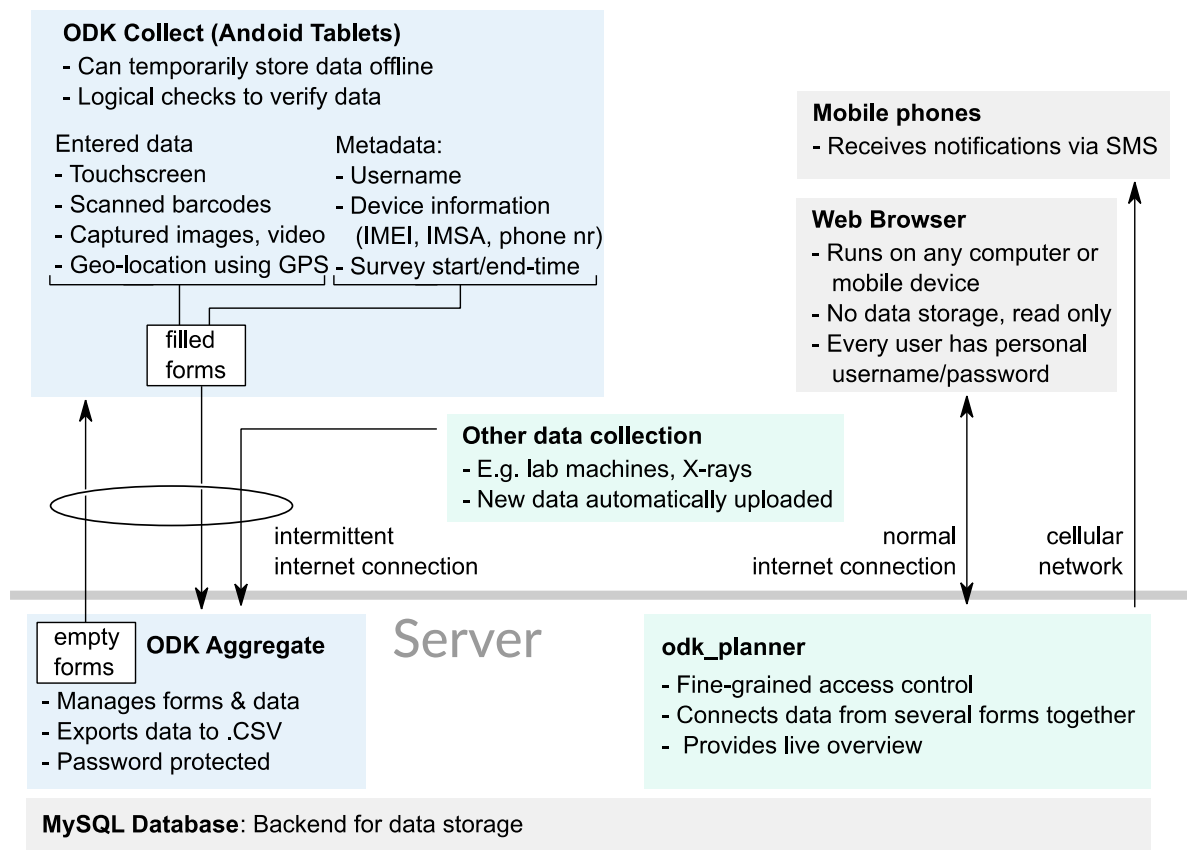
## **Authors' contributions**

Conceived and designed the study: AS, JH, SG, NM, LF. Analyzed the data: AS, JH, SG, GM, FM, BD, CIC. Contributed reagents/materials/analysis tools: JH, FM, GM, BD, CC. Wrote the manuscript: AS, JH, SG, GM, FM, BD, CIC, NM, LF. All authors approved the final version of the manuscript, and met the criteria of the ICMJE.

## **Competing interests**

All authors declare that they have no competing interests. AS developed the data management tool “odk\_planner” when he was employed by the Swiss Tropical and Public Health Institute, Switzerland, and the Ifakara Health Institute, Tanzania. The project was completed in December, 2014, and he joined the Google Inc., Zurich, Switzerland, in February 2015. At the time of development, AS had no affiliation with Google Inc., nor was he in contractual negotiations with Google Inc.

## FIGURES



**Figure 1. Modular structure of odk\_planner (green boxes) integrating an existing ODK setup (blue boxes) and commodity infrastructure (gray boxes).** odk\_planner is a read-only view of the data collected in the ODK database. It provides fine-grained access control and has added functionality, including automated notifications via SMS and tools that facilitate automated upload of data into ODK Aggregate from lab machines or X-ray scanners (odk\_pusher).

ODK data forms					
CRF1		CRF2		LRF1	
study_id	birthdate	study_id	temp	study_id	Hb
80001	1980-07-01	80001	38.3	80001-V01	14
80002	1981-07-01	80002	37.3	80002-V01	8
80003	1982-07-01	80003	37.4	80003-V01	9
80004	1983-07-01	80004	38.8	80004-V01	6
80005	1984-07-01	80005	37.6	80005-V01	13.1
80006	1985-07-01	-	-	80006-V01	15.5

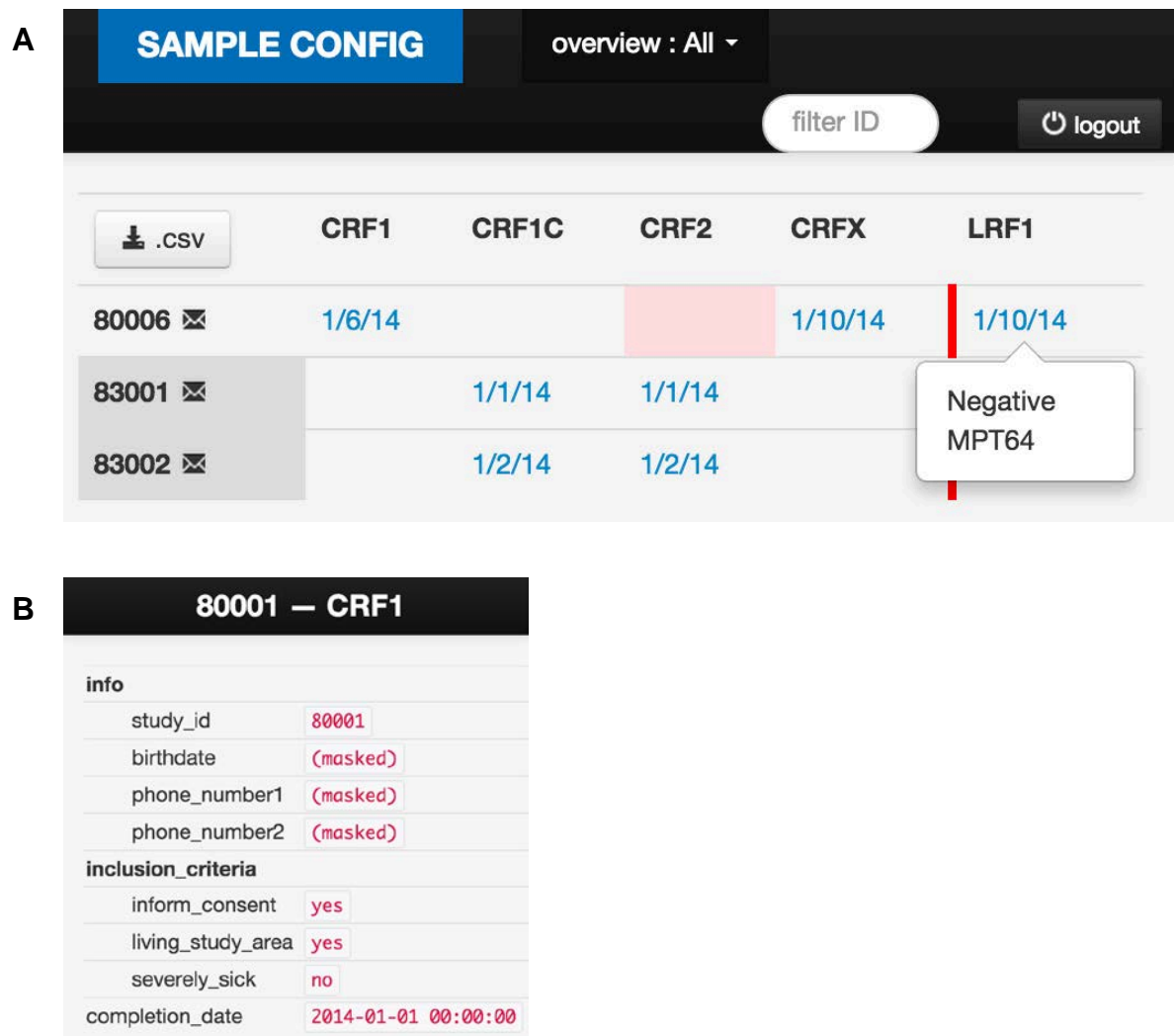
  

“odk_planner” overview				
study_id	CRF1	CRF2	LRF1	
80001	1/1/14	1/1/14	1/6/14	
80002	1/2/14	1/2/14	1/6/14	
80003	1/3/14	1/3/14	1/6/14	
80004	1/4/14	1/4/14	1/10/14	
80005	1/5/14	1/5/14	1/10/14	
80006	1/6/14		1/10/14	

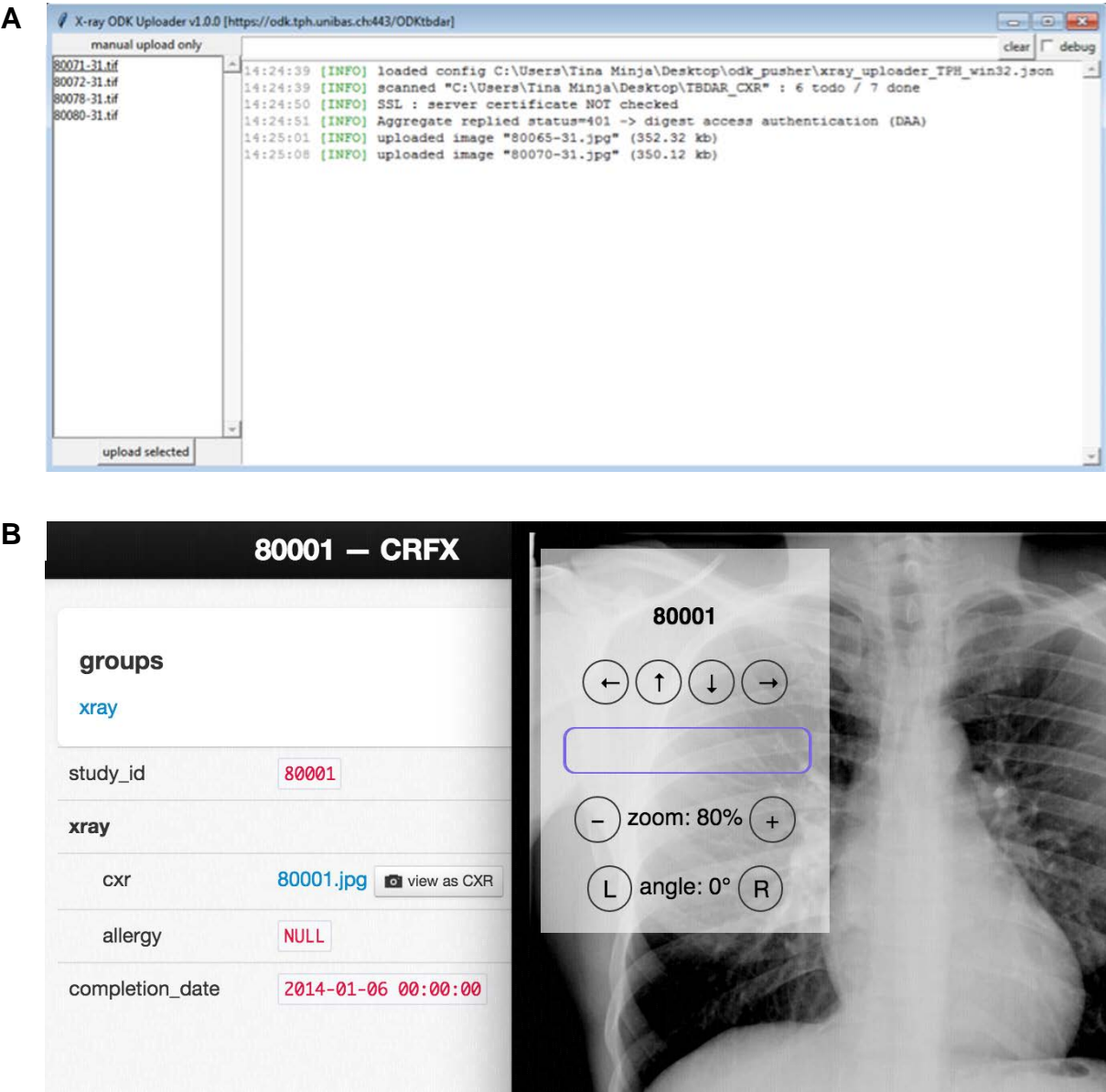
**Figure 2. Schematic overview of the functionality of odk\_planner’s “overview**

**table”.** **Top:** Three ODK forms from the same participants (linked by the unique identifier, in this example the “study\_id”). **Bottom:** odk\_planner generates a new tabular overview table that summarizes all data in the database: each column represents one form and all forms from a study participant are aligned on the same row. The date displayed (month/day/year) is the date the form data was collected. Missing forms are easily identifiable in this view (example in orange). The data can be displayed with definable graphical attributes (in this example, forms with elevated temperature are marked in red, low blood hemoglobin levels in blue, and cells of missing forms are highlighted in orange).



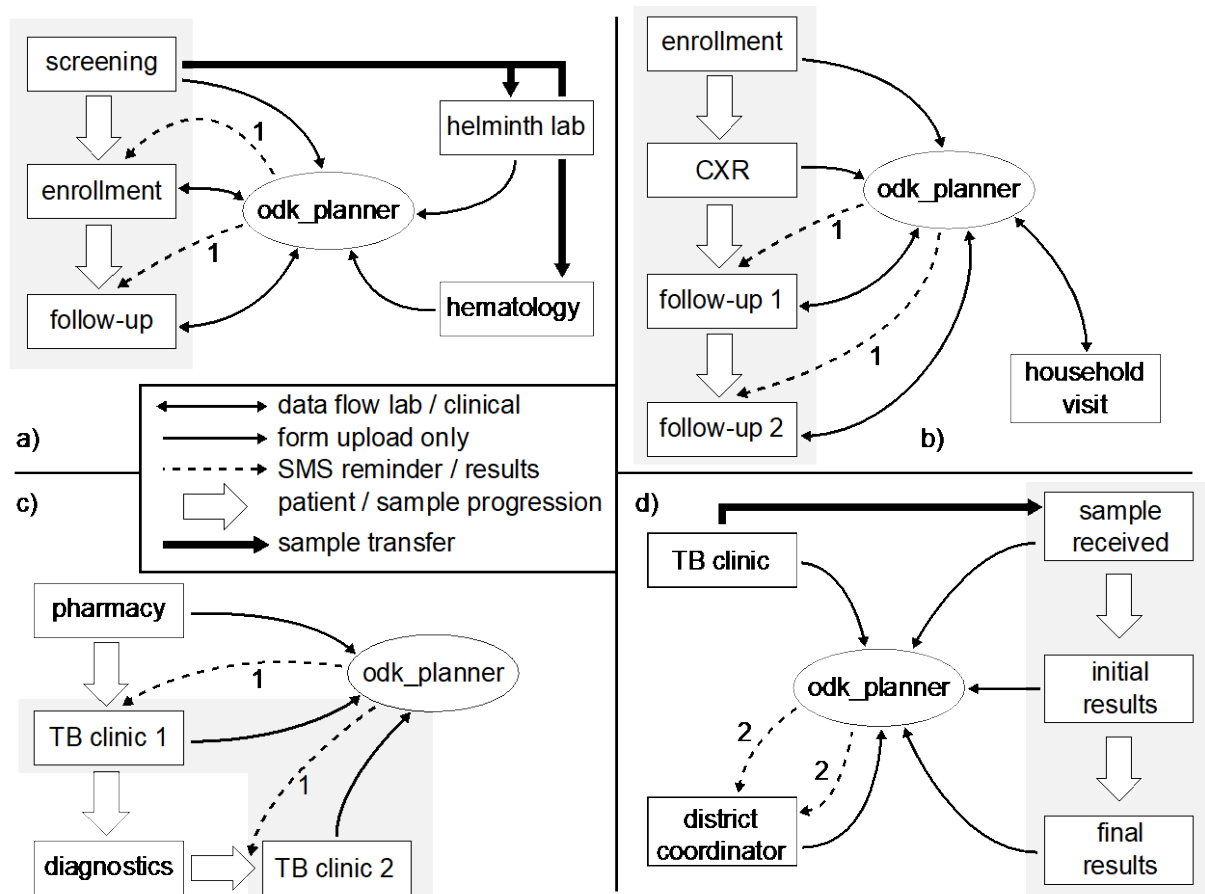


**Figure 3. Form overview and detail view of the odk\_planner. Panel A:** The screenshot shows the web-based interface (overview table) to display the data collected by ODK forms stored on the ODK Aggregate server. This example shows a study with five different forms (columns) and three participants (rows). The overview shows the date (month/day/year) every form was entered and highlights the missing forms in red (in this case patient 80006 is missing form CRF2). The overview also highlights a lab form because it contains unexpected data. A user-defined popover field (in this example “Negative MPT64”) gives more details on why a particular cell is highlighted. **Panel B:** Data detail view shown when the user clicks on a particular form in the overview. Note: some fields are masked because the user does not have access privileges to the fields “birthdate”, “phone\_number1”, and “phone\_number2”.



The screenshot displays the 'SAMPLE CONFIG' interface for sending SMS messages. At the top, there is a navigation bar with 'overview' and 'sms' tabs, and a 'logout' button. The main content area is titled 'send SMS to 80001'. It contains three input fields: 'Phone-Number' with a dropdown menu showing 'CRF1\INFO\_PHONE\_NUMBER1 : +255', 'Choose from Templates' with a dropdown menu showing 'DM', and 'Message' with a text area containing the message 'Please come back to the clinic for a blood glucose checkup.' Below the message field is a blue 'send' button.

**Figure 5. odk\_planner: short text messages.** The screenshot shows how SMS text messages are sent to study participants. Clicking on the SMS symbol (see [Figure 3, Panel A](#)), will automatically populate the interface with the participant's cell phone number. One of several templates can be selected and edited before sending. odk\_planner can also be configured to automatically send SMS, based on the content of different forms.



CXR, chest X-ray; 1, SMS reminder; 2, lab results sent via SMS.

**Figure 6. Field applications in Tanzania that successfully used odk\_planner. A:**

**Clinical trial (TB-IRON):** odk\_planner was used to collect data electronically, and to

coordinate between the Tuberculosis (TB) clinic, the Helminth and the Hematology

Laboratory. Patients were automatically notified of upcoming overdue visits by SMS

reminders. **B: Observational cohort study (TB-DAR):** odk\_planner was used to

monitor 20 different data entry forms at the TB clinic, hematology and helminth

laboratories. Patients were notified of upcoming follow-up visits by SMS reminders. **C:**

**Intervention study with an electronic referral system (TB-PHARM):** odk\_planner

helped to keep track of the patient flow. Clients were automatically reminded by a SMS

message if they did not show up at the clinic. **D: Surveillance of retreatment TB**

**cases (TB-R):** odk\_planner was used to collect data electronically, and to keep track of sputum samples sent from the local TB District Coordinator to the National TB Reference Laboratory. Laboratory results were sent via SMS messages.