

Use of personal digital assistants for data collection in a multi-site AIDS stigma study in rural south Nyanza, Kenya

*Onono MA¹, Carraher N², Cohen RC³, Bukusi EA^{1,4}, Turan JM³

1. Family Aids Care and Education Services (FACES), Kenya Medical Research Institute (KEMRI)
2. San Francisco Department of Public Health, San Francisco, California, United States of America
3. Department of Obstetrics, Gynaecology and Reproductive Sciences, University of California San Francisco, San Francisco, California, United States of America
4. Department of Obstetrics and Gynaecology, University of Nairobi, Nairobi, Kenya

Abstract

Objectives: To describe the development, cost effectiveness and implementation of a PDA based electronic system to collect, verify and manage data from a multi-site study on HIV/AIDS stigma and pregnancy in a rural, resource-poor area.

Methods: We worked within a large prevention of mother-to-child- transmission (PMTCT) program in nine rural health facilities to implement a PDA-based data collection system and to study the feasibility of its use in a multisite HIV research study in rural Kenya. The PDAs were programmed for collecting screening and eligibility data, and responses to structured interviews on HIV/AIDS stigma and violence in three local languages.

Results: Between November 2007 and December 2008, nine PDAs were used by Clinic and Community Health Assistants to enrol 1,270 participants on to the PMTCT program. Successes included: capacity-building of interviewers, low cost of implementation, quick turnaround time of data entry with good data quality, and convenience.

Conclusion: Our study demonstrated the feasibility of utilizing PDAs for data collection in a multi-site observational study on HIV/AIDS stigma conducted in remote rural health facilities in Kenya. However, appropriate and frequent data backup protocols need to be established and paper forms are still needed as backup tools in resource-poor settings.

Key words: Personal Digital Assistants, Developing countries, multisite studies, feasibility, cost effectiveness

African Health Sciences 2011; 11(3): 464 - 473

Introduction

Studies conducted in resource-poor settings face special challenges in data collection and management. These challenges include high illiteracy levels in the local communities, lack of skilled data collectors, and inadequate infrastructure including unreliable telecommunication networks¹⁻⁴. Good Clinical Data Management (GCDM) as defined by Good Clinical Practice- International Conference of Harmonization (GCP-ICH) may be burdensome, making data collection and management processes highly complicated^{5,6}. Even in developed countries, some clinical trial groups are unable to establish paper or electronic data systems that fully comply with the GCP-ICH guidelines⁷.

Paper records become difficult to manage in multi-site studies, especially when large sample sizes

are required or the number and complexity of procedures is large. The use of Personal Digital Assistants (PDAs) and other Electronic Data Capture Systems (EDCS) used for data management in research has increased in the last two decades⁸⁻¹². In both developed and developing countries PDAs have been used with varying levels of success and challenges¹³⁻²⁵. Benefits of EDCS include increased speed in data collection, transfer, and management, and improvement in quality control. Improved compliance with GCDM, with specific reference to standards for recording, handling, and storing study information and patient confidentiality has been demonstrated²⁶. While the use of EDCS is effective and there is literature to support its benefits, the failure rate of EDCS adoption is as high as 70%²⁷. Major obstacles in the adoption of EDCS include prohibitive costs of acquiring the system⁷, lack of skilled personnel, and poor information and communication infrastructures¹.

In this paper, we describe the development and implementation of a PDA-based electronic data system to collect, verify and upload data in a multi-site study on HIV/AIDS stigma conducted in rural Kenya. This paper highlights the issues encountered

*Corresponding author

Maricianah Onono

Lumumba Health Centre

P. O Box 614-40100

Kisumu, Kenya.

Tel mobile +254 720 793 871/ 732 390 992

E mail: monono@kemri-uicsf.org

maricianah@yahoo.com

in deploying this technology in a resource-poor setting. The lessons learned could be applied to other research institutions that are interested in using PDAs for data collection and management in studies conducted in rural, resource-poor settings.

Methods

Study design

The current analysis is a case study of the development and implementation of a PDA based electronic system to collect, verify and manage data from a multi-site study on HIV/AIDS stigma and pregnancy conducted in rural south Nyanza, Kenya.

Specific aims of this case study were:

To assess the usability and acceptability of the PDA system; to compare the frequency of errors in data collected using PDA electronic forms versus paper forms; and to compare the cost effectiveness of using the PDA electronic system versus a paper based system.

Setting

The case study was carried out in the context of the Maternity in Migori and AIDS stigma (MAMAS) study, being conducted at nine health care facilities in Migori and Rongo districts of Nyanza province in Kenya. Migori and Rongo districts are located in rural south Nyanza along the Lake Victoria border of Kenya and Tanzania. The MAMAS study is an on-going study implemented by Family AIDS care and educational services (FACES), a comprehensive family-based HIV

prevention, care, and treatment program run collaboratively by Kenya Medical Research Institute (KEMRI) and the University of California, San Francisco (UCSF).

Study procedures

Pregnant women who come to any of nine FACES-supported Ministry of Health facilities for their first antenatal care (ANC) visit are screened and taken through the informed consent process. Those who give consent are interviewed before their first ANC visit to assess their perceptions of HIV/AIDS stigma and other variables. Information on women's uptake of HIV testing and their HIV status is then obtained from medical records following their ANC visit. After the initial interview, a sub-set of the women are selected for longitudinal follow-up during late pregnancy (36-40 weeks) and 4-8 weeks postpartum to collect data on health care access and health outcomes through follow-up interviews. The interviews are conducted at the health facility if the woman returns for a health care visit or in the community if the woman does not return to the health facility. PDAs are used by the interviewers to collect screening data, to record responses to questionnaires, and to collect medical chart data. The questionnaires were semi-structured and designed to collect quantitative data. Questions were mainly closed-ended, with a few open-ended questions. Table 1 show topics included in data collection at different points in time. The study was approved by human subjects' research committees of KEMRI and UCSF.

Table 1: Topics included in data collection at different points in time

	First ANC visit (T1)	Late pregnancy (T2)	Postpartum (T3)
HIV/AIDS stigma and discrimination			
Anticipated stigma scale*	X	X	X
Perceived community stigma scale*	X	X	X
Scale of enacted stigma and self-stigma for PLWHA*	X	X	
Other determinants of maternity service use			
Socio-demographics: age, education, wealth, parity, etc	X		
Date of last monthly period (LMP) / weeks of pregnancy	X		
Distance, access to transport, from home to health facility	X		
Perception of the quality of maternity care at health facility	X	X	X
Other determinants of HIV service use			
HIV knowledge scale*	X	X	X
Knowledge about maternal to child transmission of HIV	X	X	X
Knowledge about antiretroviral therapy (ART)	X	X	X
Male partner's HIV testing status and sero-status	X	X	X
Anticipated response of male partner to an HIV-positive test	X	X	X
Perception of own risk / infant's risk of acquiring HIV	X	X	X
Intentions (I), behavior (B), and satisfaction regarding maternity service use			
continued use of ANC (number of visits and provider)	I	B, R	B, R
place for delivery (home vs. HF)	I	I	B, R

	First ANC visit (T1)	Late pregnancy (T2)	Postpartum (T3)
attendant for delivery (skilled vs. non-skilled)	I	I	B, R
postpartum check-up for the woman (Y/N)	I	I	B, R
six-week postnatal immunization visit for the infant (Y/N)		I	B, R
pregnancy, maternal and/or infant complications		X, R	X, R
maternal and infant survival, if survived, birth date of baby			X, R
Acceptance of HIV testing			
acceptance of HIV counselling and testing	R	X, R	X, R
reasons for refusing counselling, testing, and/or results	R	X	X
HIV status			
positive, negative, unknown (refused testing or testing services not available)	R	X, R	X, R
Disclosure and future HIV testing			
disclosure of HIV status to male partner	I	B	B
disclosure of HIV status to family, others	I	B	B
subsequent HIV testing		X, R	X, R
Intentions (I), behavior (B), and satisfaction with HIV services (HIV-positive women)			
enrolment in HIV care for self		X, R	X, R
initiation of anti-retroviral therapy (ART)		X, R	X, R
nevirapine for PMTCT (if not on ART)		I	B, R
nevirapine for baby w/in 72 hours		I	B, R
mode of infant feeding		I	B, R
HIV testing of infant		I	B, R
Severity of HIV disease			
CD4 count if available		R	R

(I=information on intentions from the questionnaire, B=information on behaviour from the questionnaire (self-report), R=information from the woman's medical record, X=other information from the questionnaire)

Hardware and Software

Hardware used in the study included eleven PDAS (3 Dell Axim v50 and 8 Hewlett Packard iPAQ hx 2400), twenty 1-gigabyte secure digital (SD) memory cards, 2 SD card readers for uploading of data to laptop computers, and six extra batteries for the PDAs. The PDAs used in this research were originally purchased for other FACES studies conducted in urban Nyanza. These PDAs were chosen due to their low cost, small size, and extended battery life, characteristics which are important in a rural set up where portability, security and electricity supply can be important constraints. The PDAs used had colour screens which made the display visible and clear and allowed the interviewers' to view the screen in environments with dim or no lighting.

The study coordinator and data manager, based in Kenya, used laptop computers installed with Questionnaire Development System (QDS) Warehouse manager and internet connection to manage and send the data from Migori to Kisumu and from Kisumu to UCSF. Electronic data collection was implemented using the QDS, software developed by Nova Research Company. The QDS

software includes a Design Studio for programming surveys; a HAPI (Handheld Assisted Personal Interview) module for running surveys on PDAs; and a Warehouse Manager for storing, merging, and cleaning the data. It was chosen because an experienced programmer with knowledge of this software suite was available to work with the study team.

The PDAs were programmed for collecting screening and eligibility data, responses to interviews, and medical chart data. The screening survey used at all study sites was a brief English language questionnaire programmed to automatically screen the women for eligibility (e.g. age, weeks of pregnancy, ANC visit, and knowledge of current HIV status). First ANC visit, late pregnancy and postpartum questionnaires were programmed with identical data structures for each of the three study languages (English, Dholuo, and Swahili). In the field, interviewers could use the most appropriate language for each participant from a menu. The medical chart survey was designed to be able to collect data from

the woman's medical records at all three time points (in English).

All surveys went through extensive development and revisions to adjust for language translation issues in the Swahili and Dholuo versions, as well as to make them fit the study design before finalization. Typically, changes to QDS surveys require that the PDA is synchronized with a desktop computer using Microsoft ActiveSync software. In this case, interviewers were located in Kenya at sites without desktop PC computers, while QDS programming took place in California, USA, therefore a different approach was used. The QDS survey files and the QDS HAPI module were installed directly onto the SD cards. Revised survey files were sent by email attachment from UCSF to the local data manager in Kenya, who would collect the SD cards and replace the old QDS survey files with the new revised files.

Staff training

Training interviewers was a two-stage process that involved pre-study formal didactic sessions, followed by hands-on on-the-job training. The pre-study training on study methodology and design was conducted over a period of two days. The interviewers were instructed on how to use the PDA, this included instruction on: switching on the PDA and entering the password, using the stylus, loading or selecting a questionnaire, entering responses to a questionnaire, charging the battery, general PDA maintenance, and how to insert and remove SD memory cards (insertion and removal). In addition to this training, the study coordinator was also trained on uploading data from an SD cards to a laptop, renaming files, zipping, encrypting (to ensure data security) and emailing the folders to the assistant data manager in Kisumu and the principal investigator (PI) based in UCSF. The study coordinator conducted weekly site visits to evaluate progress of the interviewers and use of the PDAs throughout the course of the study.

Overview of the data management processes

Interviewers entered participants' responses directly into the PDAs. However, in cases when the PDA could not be used, paper questionnaires were used as a back-up and the data later entered into a PDA. The QDS HAPI module, survey files, and the interview data were stored directly on the SD cards in the PDAs. When a survey was administered data were automatically stored to the SD card, eliminating

the need to synchronize the PDA to a desktop computer while in the field. At the end of each week one of the interviewers from each site removed the SD card from the PDA and gave it to the study coordinator, and a fresh SD card with only the HAPI module and the survey files (no data) was then received and placed into the PDA. The SD cards were rotated in this manner on a weekly basis.

The interview files from the SD cards were uploaded to the study coordinator's laptop using an SD card reader, renamed, and backed up on the laptop computer. From there, data files were sent by email to the data manager in Kisumu as zipped and encrypted files.

Once data was received in Kisumu, the files were inspected by the data manager using the QDS warehouse manager software. Basic checks were conducted on the data to check for outliers, such as incorrect participant IDs, incorrect dates, and verification that for each first ANC interview there was a corresponding first ANC medical chart data file. After these basic checks, the QDS warehouse manager software was then used to merge the different interview files into single merged files based on the language of the interviews, resulting in a total of 11 combined files in each data batch. Individual site files (up to 99 files) and the merged files were zipped and encrypted using the Ultimate Zip 2007 software..

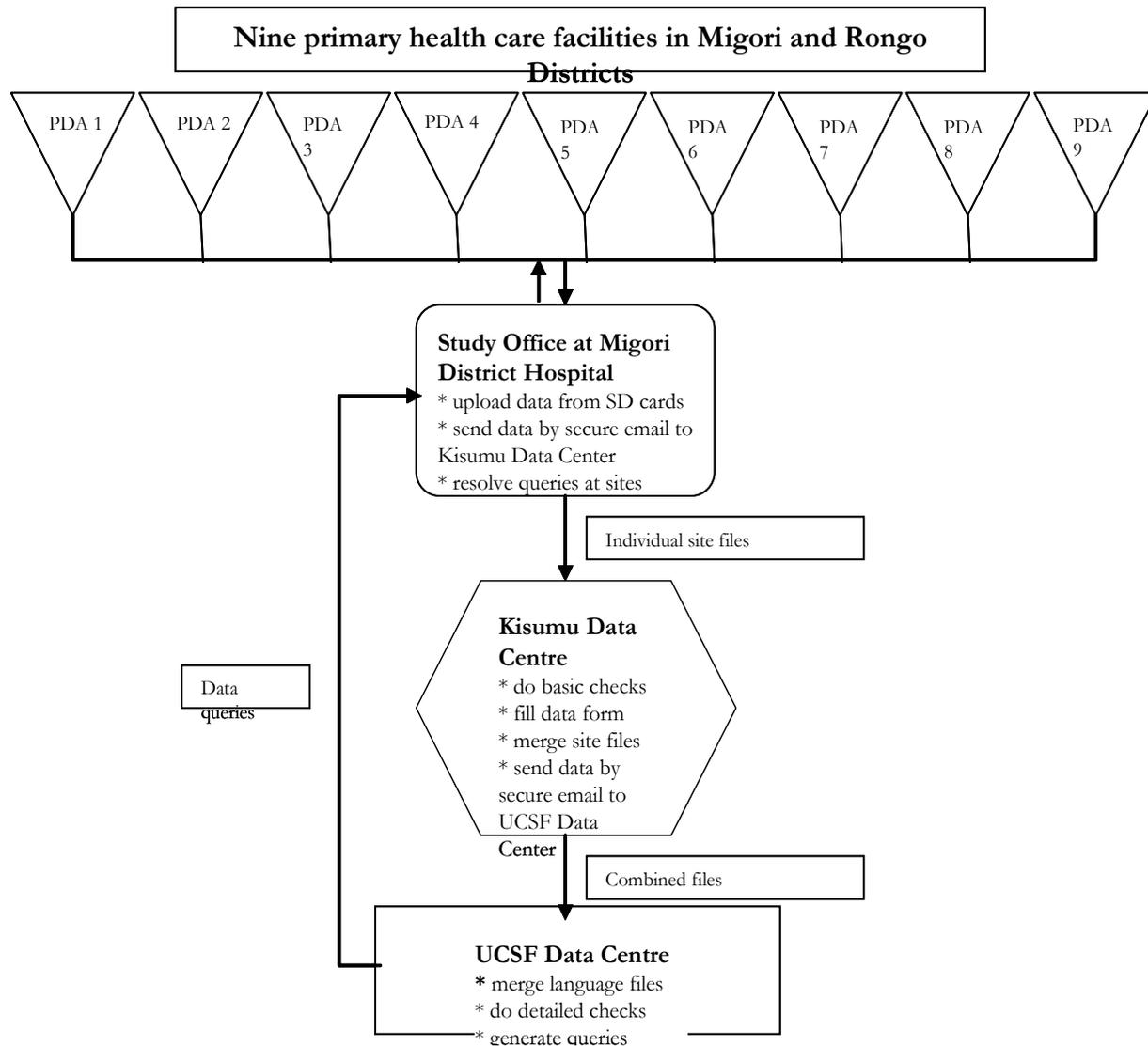
Overall supervision of the data management processes was provided by the principal investigator with guidance from the PDA Programmer based at UCSF in San Francisco via email and telephone calls, since at the beginning of the study there was a lack of in country capacity to carry out data management. As part of staff training, the local Kenyan data manager was trained to be able to conduct basic checks and cleaning and generated queries which were recorded in an Excel Data Management Form. The data manager in Kisumu supervised the data management process in Kenya. In order to improve data quality extensive data checks and additional cleaning were conducted after the data arrived at UCSF and queries were added to the preliminary Excel Data Management form. Queries were then sent via email to the study coordinator for resolution at the peripheral sites. This also served as a learning platform for the data manager. The excel data management form included data queries from the data manager in Kisumu and any additional queries from UCSF that he may have overlooked. . The nature of queries included mainly

incorrect entry of participant IDs resulting in duplicates in the dataset or inability to match with the medical chart and follow-up data, wrong interview or last monthly period (LMP) dates, missing information, or mislabelling of an interview, e.g. selecting the wrong

interview language. All queries were resolved by study closeout.

The data management system is presented schematically in Figure 1.

Figure 1: Data Management system for the PDA case Study



Data security

PDAs used in the study are all password-protected and the data stored on these devices was encrypted. The interviews in the PDA were programmed in such a way that on completion of an entry, the data were copied and automatically saved to the SD card. Interviewers were not able to retrieve or change an entry once it was copied and saved to the SD card and no record of the entry was retained on the PDA. In case of PDA loss, it was impossible to access and see the data on the SD card without the PDA

password and the QDS warehouse manager software. An electronic data audit trail was created in form of an excel spreadsheet that documented the name of the site, the date the SD card was uploaded, participant Ids for the interview data contained on the SD card, and the date of uploading data to the data manager and the PI in UCSF. Data were backed up on a regular basis at the Migori study office, at the Kisumu data centre, and on a secure password protected server at UCSF.

Results

Between November 2007 and December 2008, nine PDAs (one per site) were used to enrol 1,270 participants in Migori (7 sites) and Rongo (2 sites) districts in rural south Nyanza. The PDAs were used successfully to screen, conduct interviews, and capture information from the medical registers.

The PDA system usability and feasibility

Capacity-building

Thirty-four interviewers, with primary (5.9%) and high school education (94.1%) carried out the screening procedures, administered the surveys, and collected medical chart information with minimal technical supervision. Data from the first ANC visit interviews were recorded on the PDA for 1,270 participants in one of three languages (1,116 in Dholuo, 68 in Swahili and 86 in English). Following 2 days of didactic and hands on training, all of the interviewers became comfortable with using the PDA to enter information. Although the interviewers had limited computer skills before the pre-study training; they were able to master and successfully use the PDAs.

Convenience

The small size of the PDA made community follow-up interviews easier since it could be carried around in any weather and over long distances. It was also storage friendly, requiring minimal space. The device allowed for reduced paper use which was environmental friendly.

Quick turnaround time

Availability of an internet connection at the study office in Migori town allowed for online mentoring and troubleshooting of the technical problems and a rapid turnaround time for the data and queries, between the study office in Kenya and San Francisco, California. Resolution of data queries was generally completed within two weeks of data collection.

Security

One PDA (with two interviews on the SD card) was stolen from one of the study health facilities during the post election violence that hit Kenya between December 2007 and March 2008. The stolen PDA was not recovered, but because data on the SD card were encrypted and the PDA password protected, participant privacy was not compromised. Installation of locked mounted boxes for secure

storage of at each site prevented further loss of PDAs and promoted data security.

Poor infrastructure

Lack of electricity in 4 out of 9 study sites hampered use of PDAs when the battery power was insufficient, necessitating use of back-up paper interviews. This was our biggest challenge. Of the 5 facilities that had some electricity, three used the facility's solar charging system to charge the PDAs. Lack of computers and internet at all sites and irregular electricity supply and/or internet connection at the study office in Migori, caused delays in immediate transfer of the data collected. Data could not be backed up at most of the sites because of lack of internet access therefore data were only backed up once a week at the study office in Migori town. Paper surveys were used as an important backup in cases of lack of a charged battery, PDA malfunction, or corruption of interview files on the SD card. A paper study register with participant and interviewer IDs facilitated the query resolution process. Extra batteries delivered regularly to the 4 sites without electricity ensured that most data were collected electronically.

Operational difficulties with the PDA

During the early study period, some ineligible women were enrolled due to wrong calculation of gestational weeks of pregnancy by the PDA. The gestational age calculation was actually initially programmed into the PDA. The interviewers had to enter the day, month, and years separately. However we found that when the PDA lost power, it automatically reset the internal date to the date of PDA manufacture. If the PDA date was not subsequently manually reset to the correct date, the calculated gestational age would be wrong, resulting in the enrolment of women who were not eligible for the study. Thus, we decided to use pregnancy wheels together with the PDA as checking mechanism. Eighteen (1.41%) out of 1,274 enrolled were later determined ineligible for this reason. To overcome this challenge, the PDA screening tool was adjusted to allow manual entry of the current date and pregnancy wheels were used by interviewers to verify gestational age. This resulted in fewer errors and the number of women later judged to be ineligible because of gestational age miscalculations decreased from 16 between January and June 2008, compared to 2 between July and December 2008. We were unable to retrieve interview data for 22

enrolled participants (1.73%) from the SD cards used at two sites, despite troubleshooting by the data manager and the PDA programmer.

Cost effectiveness of using the PDA electronic system in comparison to the paper based system

The total cost of the PDA hardware and QDS software costs were \$6,471; an average of \$719 per PDA. Training for the use of this system consisted of a 2-day training with the users for approximately 4 hours each. Further, there was frequent phone and email contact between the developer and users during the entire development and implementation periods. The frequency of contact decreased during the study period. We have not included the costs of a computer with internet access here because we assume that a paper based system would also use a computer for data entry. We did not include the costs of the PDA-related supervision visits, because a paper-based system would also require the supervision visits to collect the completed forms, check errors, discern illegible handwriting, and provide fresh questionnaires in their appropriate languages. This paper-based survey supervisory work cannot be delegated due to the need to maintain data confidentiality and security. The PDA-related supervision was less time intensive because less qualified staff could be sent to deliver freshly charged batteries and exchange the SD card without fear of breach of confidentiality, as the data were encrypted and password protected.

The cost of doing the entire study with paper interviews was estimated at \$2,533.04. This included a 30% wastage factor based on the fact that women could speak each of the three languages to varying degrees and this was not easily predictable which language questionnaire would be required so sufficient numbers of the forms in all 3 languages had to be available. Additional costs of paper data collection included a filing cabinet at the study site plus binders for the paper forms which totaled to \$300.

PDA-based system reduced the number of staff required for the entire process by reducing the need for data entry clerks to enter data from the paper forms onto a computer. The average monthly salary of a data clerk working for FACES in Nyanza province, Kenya is \$280.00 or \$3,360 annually. A paper system for this study would have needed at least 2 data clerks – one for each district. Therefore this system provides important initial savings of \$6,720 annually for the entire team and a further ~\$3,841. Given that the PDAs will be reused in

subsequent studies in the programme the \$3,321 initial cost is an investment as opposed to the recurrent expenditure of printing of paper questionnaires over the years (table 2).

Table 2: Cost comparison of PDA system vs. paper system

PDA system	Price	Paper system	Price
Hardware	\$3,150	Printing cost	\$3,292.9
Software	\$3,321	Filing costs	\$300
		Data entry clerks	\$6,720
Total	\$6,471		\$1,0312.9

Frequency of errors entered with and without the electronic system

Due to the highly structured format of the questionnaires and the presence of automatic skip patterns programmed into the PDA surveys, data collected were “clean” with very few interviewer errors. Overall 15% of the surveys were collected on paper first and then later entered into a PDA. We found that the proportions of missing responses for selected key questionnaire items for interviews that were directly entered in the PDA were less than the same items initially collected on paper forms and later entered into the PDA . Although not all are statistically significant, there was a clear trend toward more “refuse to answer” responses with the paper forms on some key questions, such as:-

Women who give birth at health facilities are often treated harshly by health workers (3.2% PDA vs 6.1% paper, $p=.060$);

Traditional birth attendants have more flexible arrangements for payment than do health facilities 1.4% vs 5.6% ($p=.010$);

Childbirth goes faster with a traditional birth attendant 1.9% vs 7.6% ($p=.000$);

Can a person get HIV from mosquito bites? 2.5% vs 5.8%, $p=.034$;

Has your husband/partner/father of the baby been tested for HIV? 1.3% vs 4.2% ($p=.010$)

Errors were mainly encountered in instances when the interviewers had to key in a number or date using the PDA “keyboard”. Errors included wrong entry of dates, and incorrect entry of participant IDs resulting in duplicates in the dataset or inability to match with medical chart and follow-up data. These errors were detected during post entry data cleaning and the majority of errors were resolved on the study coordinator’s next visit to the site.

Data management complexity

There were 11 different data collection forms used at the sites in 3 languages. Each data “batch” potentially contained up to a total of 99 files from the 9 sites. These were merged by the data manager in Kisumu to a total of 11 files and then language versions for each questionnaire were merged at UCSF. The merging, checking, and querying of these data batches was a major time and resource commitment for this low-budget study but this would have been the case if this were a purely paper-based survey as well.

Discussion

With the increasing pressure on investigators and sponsor companies to bring public health solutions to the table faster while complying with all regulations, it is imperative that the research centres embrace technological change in a positive manner. In our case study we found that it is feasible to use PDAs to collect data in a multi-site observational HIV stigma study in a rural setting in Africa.

Barriers to implementation

In this study, as in much of the published PDA literature the barriers to full implementation included poor infrastructure, especially the lack of regular electricity supply and internet connection, and lack of local skills in PDA operation and programming. The single most important constraint was the lack of the opportunity to perform daily backup of the data. A study of PDA use in a pilot project in Bolivia also found that lack of an appropriate and adequate back-up protocol was the most significant limitation to PDA use in that setting²⁸.

Compliance with good clinical and data management practices

The GCP compliant software is now available for data management²⁶. Some of the primary requirements include use of validated systems, secure retention of electronic records allowing instant reconstruction of analyses, user-independent, computer-generated, time-stamped audit trails, system and data security, data integrity, and confidentiality through system access control²⁹. Rather than face non-compliance where this does not exist a hybrid system which encompasses a complementary mix of the two methods is desirable. In our case PDAs were used to manage certain aspects of the data management process and paper the other³⁰. In resource poor settings and in setups where EDCS is

new, a paper trail may be the only way to guard against lost data. In this study paper forms were of particular help in the 4 sites which had no or an erratic supply of electricity, so that these sites could still meet their enrolment targets. In rural sites where internet connection may be available, software that allows auto saving of the interviews as they are conducted, coupled with real time internet transmittal to the data centre, is desirable.

Benefits of electronic data capture systems

Data flow needs to be fast, accurate and confidential in addition to being complete and reliable³¹. This study is a successful example of a multi-site rural African research study being able to set up and facilitate good data flow in a region with minimal infrastructure. Because data were encrypted in the SD card, it remained confidential, thus protecting participant privacy. The PDA programme had self auditing features and did not allow an interviewer to move to the next question without entering a response, therefore 100% completion of interview questions was possible. The PDA allowed the interviews to be auto saved and thus data could not be altered. We were able to compare interviews entered directly into the PDA with those initially collected on paper and found that the rate of “refuse to answer” responses was lower with direct PDA data entry. One possible explanation may be that the women felt the information collected on the SD card was more secure and confidential than on paper, in terms of being able to be retrieved and connected to them, and hence were more open with the interviewer.

Study limitations

A limitation in our study is that we did not conduct a direct comparison of data entered in the PDA versus data only recorded on paper. However, both descriptive and randomized control studies conducted in other settings have determined that discrepancies between electronic and paper data collection are minimal^{10,32}. Another limitation of this study was that; although a significant amount of training for the study and PDA data collection was carried out in Kenya with the study coordinator, the data manager, and the interviewers; these training sessions were not formally evaluated.

Conclusion

Our study shows the feasibility of utilizing the PDA in data collection in a multi-site observational study in a rural setting in Africa. Despite interviewer

inexperience in using the PDA, we found they can be effectively trained, through didactic and hands-on methods, to collect high quality data with a relatively low error rate. The significance of the findings from this analysis is that though many studies have confirmed use of the PDA as a data collection tool, a complete overhaul of the need for paper, as recommended by many³³ as a primary source of data needs to be reconsidered, especially for low-resource rural settings in Africa with limited opportunity for back-up. Appropriate and frequent data backup protocols need to be established and paper forms still need to play an important role as a backup method in such resource-poor settings.

Acknowledgements

We would like to acknowledge the help of: the assistant data manager from FACES-Kisumu, who provided technical help in the data management process, the medical editor from UCSF who was involved in the formative drafting of the manuscript and revising and editing its structure critically, Dr. Nicole Kley and Ms. Katie Doolan for their assistance in MAMAS Study coordination, the Clinic and Community Health Assistants of the FACES programme in Migori and Rongo districts for their work in data collection and the Ministry of Health staff in the rural health centers in Migori and Rongo districts for their cooperation and support. This work was supported by grant from the U.S. National Institute of Mental Health.

References

1. Carrigan R, Milton R, Morrow D: Developing a pan-African resource network by adapting ICTs to meet site-specific needs: 2005 computerworld honors case study' [http://www.cwhonors.org/laureates/science/international%20aids.pdf] (Accessed on 15 January 2009)
2. Kauchali S, Rollins N, Van den Broeck J: Local beliefs about childhood diarrhoea: importance for health care and research. *J Trop Paediatr* 2003; 50: 82–89.
3. Good MJ. Local knowledge: research capacity building in international health. *Soc Sci Med* 1992; 35: 1359–67.
4. Jones RC, Sandoval JA, Gray JR, Ribo AI, Moreno SA, Armas LN, et al. Multidisciplinary approach to initiating HIV clinical trials in rural communities. In International Conference on AIDS 11-16th July 2004 : Bangkok, Thailand.
5. ICH Harmonised Tripartite Guideline: Guideline for Good Clinical Practice E6; 1996. [http://www.ich.org/LOB/media/MEDIA482.pdf] (Accessed 09 February 2009).
6. Grimes DA, Hubacher D, Nanda K, Schulz KF, Moher D , Altman DG. The Good Clinical Practice guideline: a bronze standard for clinical research. *Lancet* 2005; 366(9480): 172–174.
7. Fegan GW, Lang TA. Could an Open-Source Clinical Trial Data-Management System Be What We Have All Been Looking For? *PLoS Med* 2008; 5(3): e6
8. Diero L, Rotich J, Bii J, Mamlin B, Einterz R, Kalamai IZ, Tierney WM. A computer-based medical record system and personal digital assistants to assess and follow patients with respiratory tract infections visiting a rural Kenyan health centre. *Bio Medical Central Medical Informatics and Decision Making* 2006; 6: 21.
9. Dwolatzky B, Trengove E, Struthers H, McIntyre J, Martinson NA. Linking the global positioning system (GPS) to a personal digital assistant (PDA) to support tuberculosis control in South Africa: A pilot study. *International Journal of Health Geographics* 2006; 5: 34.
10. Missinou M, Olola CHO, Issifou S, Matsiegui P, Adegnika A, Borrmann S, et al. Short report: Piloting paperless data entry for clinical research in Africa. *American Journal of Tropical Medicine and Hygiene* 2005;72(3): 301–303.
11. Olola CH, Missinou MA, Issifou S, Anane-Sarpong E, Abubakar I, Gandi JN, Chagomerana M, Pinder M, Agbenyega T, Kremsner PG, Newton CR, Wypij D, Taylor TE; SMAC Network: Medical informatics in medical research—The Severe Malaria in African Children (SMAC) Network's experience. *Methods Information in Medicine* 2006; 45: 483–491.
12. Taylor T, Olola C, Valim C, Agbenyega T, Kremsner P, Kremsner P, Krishna S, Kwiatkowski D, Newton C, Missinou M, Pinder M, Wypij D: Standardized data collection for multi-center clinical studies of severe malaria in African children: Establishing the SMAC network. *Trans R Soc Trop Med Hyg* 2006; 100: 615–622.
13. Carroll AE, Saluja S, Tarczy-Hornoch P. Development of a personal digital assistant (PDA) based client/server NICU patient data and charting system. *Proc AMIA Symp* 2001:100-4.

14. Carroll AE, Saluja S, Tarczy-Hornoch P. The implementation of a personal digital assistant (PDA) based patient record and charting system: lessons learned. *Proc AMLA Symp* 2002;111-5.
15. Thomas SM, Overhage JM, Warvel J, McDonald CJ. A comparison of a printed patient summary document with its electronic equivalent: early results. *Proc AMLA Symp* 2001;701-5.
16. Caro Sr JJ, Caro I, Caro J, Wouters F, Juniper EF. Does electronic implementation of questionnaires used in asthma alter responses compared to paper implementation? *Qual Life Res* 2001; 10:683-91.
17. Jamison RN, Raymond SA, Levine JG, Slawsby EA, Nedeljkovic SS, Katz NP. Electronic diaries for monitoring chronic pain: 1-year validation study. *Pain* 2001; 91:277-85.
18. Hyland ME, Kenyon CA, Allen R, Howarth P. Diary keeping in asthma: comparison of written and electronic methods. *BMJ* 1993; 306:487-9.
19. Saleh KJ, Radosevich DM, Kassim RA, Moussa M, Dykes D, Bottolfson H, et al. Comparison of commonly used orthopaedic outcome measures using palm-top computers and paper surveys. *J Orthop Res* 2002;20:1146-51.
20. Lal SO, Smith FW, Davis JP, Castro HY, Smith DW, Chinkes DL, et al. Palm computer demonstrates a fast and accurate means of burn data collection. *J Burn Care Rehabil* 2000; 21:559-61. discussion 558.
21. Yon BA, Johnson RK, Harvey-Berino J, Gold BC. The use of a personal digital assistant for dietary self-monitoring does not improve the validity of self-reports of energy intake. *J Am Diet Assoc* 2006;106:1256-9.
22. Van Gerven JM, Schoemaker RC, Jacobs LD, Reints A, Ouwersloot- van der Meij MJ, et al. Self-medication of a single headache episode with ketoprofen, ibuprofen or placebo, home-monitored with an electronic patient diary. *Br J Clin Pharmacol* 1996;42:475-81.
23. Buck DS, Rochon D, Turley JP. Taking it to the streets: recording medical outreach data on personal digital assistants. *Comput Inform Nurs* 2005;23:250-5.
24. Kvien TK, Mowinckel P, Heiberg T, Dammann KL, Dale O, Aanerud GJ, et al. Performance of health status measures with a pen based personal digital assistant. *Ann Rheum Dis* 2005;64:1480-4.
25. Fletcher LA, Erickson DJ, Toomey TL, Wagenaar AC. Handheld computers. A feasible alternative to paper forms for field data collection. *Eval Rev* 2003; 27:165-78.
26. Avilés W, Ortega O, Kuan G, Coloma J, Harris E. Integration of Information Technologies in Clinical Studies in Nicaragua. *PLoS Med* 2007; 4(10): e291
27. Tyson G, Dietlin T. Realizing the Promise of Electronic Data Capture: A practical guide. Raleigh, NC: Campbell Alliance 2006 [http://www.campbellalliance.com/articles/Realizing_the_Promise_of_Clinical_EDC_A_Practical_Guide.pdf] (Accessed 24 February 2010)
28. Escandon, IN, Searing H, Goldberg R, Duran R, Arce JM. The use of PDAs to collect baseline survey data: Lessons learned from a pilot project in Bolivia. *Global Public Health*, 2008; 3(1): 93 - 104
29. FDA: 21 CFR Part 11, Electronic Records; Electronic Signatures, 2009 [http://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfcfr/CFRSearch.cfm?CFRPart=11&showFR=1] (Accessed on 21/02/10)
30. Society for Clinical Data Management: Good Clinical Data Management Practices 2009 [http://www.scdm.org/members] (Accessed on 19/02/10)
31. Van den Broeck J, Mackay M, Mpontshane N, Luabeya AKK, Chhagan M, Bennish ML. Maintaining data integrity in a rural clinical trial *Clinical Trials* 2007; 4: 572–582
32. Forster D, Behrens RH, Campbell H, Byass P: Evaluation of a computerized field data collection system for health surveys. *Bull World Health Organ* 1991; 69(1):107-11.
33. Shirima K, Mukasa O, Schellenberg JA, Manzi F, John D, Mushi A, et al. The use of personal digital assistants for data entry at the point of collection in a large household survey in southern Tanzania' Emerging themes in epidemiology 2007; 4(5) [http://www.ete-online.com/content/4/1/5] (Accessed 14/1/2009)