

SHORT REPORT: PILOTING PAPERLESS DATA ENTRY FOR CLINICAL RESEARCH IN AFRICA

MICHEL A. MISSINOU, CHRISTOPHER H. O. OLOLA, SAADOU ISSIFOU, PIERRE-BLAISE MATSIEGUI, AYOLA A. ADEGNKA, STEFFEN BORRMANN, DAVID WYPIJ, TERRIE E. TAYLOR, AND PETER G. KREMSNER
Medical Research Unit, Albert Schweitzer Hospital, Lambaréné, Gabon; Department of Parasitology, Institute of Tropical Medicine, University of Tübingen, Tübingen, Germany; Kenya Medical Research Institute, Centre for Geographic Medicine Research, Coast, Kilifi, Kenya; Clinical Research Program, Children's Hospital, Boston, Massachusetts; Department of Pediatrics, Harvard Medical School, Boston, Massachusetts; Department of Biostatistics, Harvard School of Public Health, Boston, Massachusetts; Blantyre Malaria Project, Queen Elizabeth Central Hospital, Blantyre, Malawi; Department of Internal Medicine, College of Osteopathic Medicine, Michigan State University, East Lansing, Michigan

Abstract. Direct data entry, using handheld computers, may simplify and streamline data management, especially in remote settings. We compared the accuracy of data entry using the current standard practice (a paper-based case report form with double data entry) with that using a personal digital assistant (PDA) in a clinical study in rural Gabon. The rate of discrepant entries was 1.7%. Categorical data (presented in “pull down” menus on the PDA) were more commonly discrepant than were continuous “typed in” data (2.4% versus 1.2%; $P = 0.001$). Both systems functioned smoothly and no data were lost. The clinicians involved in this study preferred the handheld computers, and their use will be considered in future studies in an African clinical research network.

Quick, accurate data entry is required for clinical research studies. Patient data are commonly recorded on paper-based case report forms (CRFs). These are useful for data verification, allow tracking of corrections, and provide documentation. The current standard is to use a “double data entry” system. The data are entered from the CRF by two different individuals, the two databases are compared, and any discrepancies are addressed by double checking either the CRF or the source document. This approach can be very accurate but is also time-consuming, particularly when the variables are numerous and the study population is large. Direct data entry, using handheld computers or personal digital assistants (PDAs), is an alternative to the standard system. There are advantages (fewer individuals required, more rapid data entry) and disadvantages (lost data, no paper trail) associated with this approach. Handheld computers have been used in medical settings in developed countries to provide additional information at the bedside.^{1,2} Clinical studies in these countries have compared PDAs with paper-based CRFs, and have demonstrated equivalence in data accuracy.^{3,4} In these studies, patient assessments using PDA were as efficient and content-rich as paper assessments. In developing countries, handheld computers used in field studies provided improved data precision, decreased collection time, and fewer errors when compared with paper questionnaires.^{5,6}

We incorporated direct data entry into an ongoing clinical surveillance activity at the Albert Schweitzer Hospital in Lambaréné, Gabon so that we could compare it with the standard approach (data verification and double entry from a paper-based CRF).

The study site, the Medical Research Unit of the Albert Schweitzer Hospital in Lambaréné, is one of five sites in the Severe Malaria in African Children clinical network. Lambaréné is situated on the equator in a typical central African rain forest area in Gabon. The average temperature is approximately 27°C and rainfall falls throughout the year with a mean humidity of approximately 80%.

This study was reviewed and approved by the ethics committee of the International Foundation of the Albert Schweitzer Hospital (Lambaréné, Gabon), the ethics committee of

the University of Tübingen (Tübingen, Germany) the University Committee for Research Involving Human Subjects at Michigan State University (East Lansing, MI), and the Division of Microbiology and Infectious Diseases International Clinical Studies Review Committee at the National Institute of Allergy and Infectious Diseases, National Institutes of Health (Bethesda, MD). Informed consent was obtained from all participants' parents or legal guardians. All parasitemic children admitted to the hospital were eligible for inclusion, and 51 different variables (Table 1) were collected from each participant. Data were entered onto the standard CRF by the admitting clinician; simultaneously, a second clinician used a PDA (Palm Pilot m500 series; Palm Inc., Santa Clara, CA) programmed using Pendragon Forms software version 3.1 (Pendragon Software Corporation, Libertyville, IL) to enter independently the same data. Categorical variables were presented in “pull down” menus, continuous variables required direct written entry by the operator, and three values were pre-programmed to be automatic entries or calculations (Table 1). Four different clinicians (three MDs and one clinical officer) participated in the study, and none had any previous experience with direct data entry or with handheld computers. Eight hours of training was provided to familiarize them with the handheld computers and with the software. Upon enrollment, the clinicians entered the demographic, historical, and clinical examination data on the CRF and PDA independently. Laboratory results and clinical outcomes were later added to both the CRF and the PDA by the clinicians. The handheld computers were synchronized regularly (at least twice a day) with a desktop computer containing the parallel permanent database for the study. Data from the CRF were double-entered, verified, and checked prior to being added to the permanent database. For data security, both databases were backed-up daily on disks that were stored securely in sites different from where the original databases were located. The PDAs were also locked in secure locations. Only the investigators conducting the study had direct access to the databases, the backed-up data, and the PDAs. Discrepancies between the data set from the CRF and the data set from the PDAs were identified and analyzed independently

TABLE 1
Variables collected on the case report form and on the handheld computer, by type

Categorical variables ("pull down" menu) (n = 22)	Continuous variables ("typed in" values) (n = 26)	Automatic variables (no data entry required) (n = 3)
Informed consent obtained	Patient number	Date of admission
Ethnic group (code)	Weight	Time of admission
Sex	Temperature	Age (calculated from date of birth and date of admission)
Convulsions prior to admission	Date of birth	
Vomiting prior to admission	Blantyre Coma Score motor	
Site at which temperature was taken	Blantyre Coma Score verbal	
Deep breathing	Blantyre Coma Score eye movements	
Intercostal recessions	Respiratory rate	
Irregular breathing	Spleen (cm below costal margin)	
Is the child sucking?	Date of discharge or death	
Can the child sit?	Time of death	
Can the child stand?	Parasitemia	
Can the child walk unaided?	Hemoglobin	
Outcome	Hematocrit	
Blood transfusion given	Glucose	
Source of blood for parasitemia	Lactate	
Source of blood for hemoglobin	pH	
Source of blood for hematocrit	pCO ₂	
Source of blood for glucose	Base excess	
Source of blood for lactate	Pigmented mononuclear leukocytes/200	
Source of blood for pH/pCO ₂ /base excess	Pigmented monocytes/50	
Source of blood for pigment	Pigmented monocytes/ μ L	
	Pigmented polymorphonuclear leukocytes/200	
	Pigmented neutrophils/100	
	Pigmented neutrophils/ μ L	
	Pigmented parasitized red blood cells/200	

of the clinicians who collected the data. We performed a verbal survey to determine the users' satisfaction of using one approach compared with the other.

One hundred and four patients were included during January and February 2002. The two data entry systems functioned smoothly and no data were lost. There were no hardware problems due to tropical environments (heat, humidity). There were no instrument failures and the batteries worked well. A single person could accomplish all of the data entry using the PDA; at least three people (one clinician and two data entry clerks) were required for the paper-based system. The data collected with the PDA were immediately usable once synchronized to the study data base (at least two times a day), whereas the data from the CRF went through the standard double-entry and verification process prior to being added to the permanent database. The overall discrepancy rate was calculated using 48 different variables (the three automatically entered variables were excluded). Eighty-six discrepancies (discrepancy rate = 1.7%) were noted. These involved 61 of the 104 records and 25 of the 48 variables. Forty-two records contained 1 discrepancy, 15 records had 2 discrepancies, 2 records had 3 discrepancies, and 2 records had 4 discrepancies. The discrepancy rate for the categorical "pull down" variables (2.4%) was twice that of the continuous "written in" variables (1.2%) ($P = 0.001$, by two-sided Fisher's exact test). Thirteen of 22 categorical variables (59%) had discrepancies, compared with 12 of 26 (46%) continuous variables ($P = 0.40$, by two-sided Fisher's exact test).

The overall rate of discrepancies between the PDA and the CRF databases was low (1.7%). Discrepancies in the categorical "pull down" variables occurred more frequently than in the continuous variables, and may have reflected the lack of

experience of the users. Direct data entry onto the same record, but at different times, may also have been a source of error.

Fewer individuals were required to accomplish data entry with PDAs compared with traditional CRFs, but more effort was required to program the PDAs than to develop the CRFs. Because the PDA is an electronic system, the data collected using it were more readily available and usable than the data collected using the paper CRF. However, this comparison was not quantified. The CRFs were better for tracking additions and corrections to the data in this study, but the software for handheld computers has improved since the time of our study, and PDAs can now be programmed to track changes to the database. The clinicians enjoyed the process of using PDAs and found it to be efficient, but this is a simple, straightforward study. Little follow-up and checking were required. If these activities were more extensive, additional support staff would be needed.

It is worth noting that although paper-based CRFs with double-data entry are currently regarded as the gold standard, this approach is not infallible. Errors, especially errors in transcription, can still occur. A study design that included two CRF-based and two PDA-based databases, and which had the source documents available as the gold standard, would permit both within-system and between-system comparisons. It would also create a more appropriate context for assessing the importance of the rate of discrepancies between paper-based data management and direct data entry systems.

Based on our experience, the Severe Malaria in African Children clinical network would be willing to consider the use of handheld computers for direct data entry, especially for large, simple studies in which all of the data are collected at one time and in one place.

Received March 23, 2004. Accepted for publication August 11, 2004.

Acknowledgments: We thank the children and the parents for their patience. We acknowledge the immense collaborative efforts accorded by all Severe Malaria in African Children clinical network members.

Financial support: The study was supported by a grant from the National Institute of Allergy and Infectious Diseases (U19 AI45955).

Authors' addresses: Michel A. Missinou, Medical Research Unit, Albert Schweitzer Hospital, BP 118 Lambaréné, Gabon, Telephone: 241-58-10-99, Fax: 241-58-11-96 and Sektion Humanparasitologie, Institut für Tropenmedizin, Wilhelmstrasse 27, 72074 Tübingen, Germany, Telephone: 49-7071-29-82195, Fax: 49-7071-29-5189, E-mail: michel.missinou@uni-tuebingen.de. Christopher H. O. Olola, Kenya Medical Research Institute, Centre for Geographic Medicine Research, Coast, Box 230, Kilifi, Kenya, Telephone: 254-41-522-063, Fax: 254-41-522-390 or 254-20-271-1673, E-mail: colola@kilifi.mimcom.net. Saadou Issifou, Pierre-Blaise Matsiégui, and Akim A. Adégnika, Medical Research Unit, Albert Schweitzer Hospital, BP 118 Lambaréné, Gabon, Telephone: 241-58-10-99, Fax: 241-58-11-96, E-mails: issifou@lambarene.mimcom.net, matsiegp@yahoo.fr, and aadegnika@yahoo.fr. Steffen Borrmann, Sektion Humanparasitologie, Institut für Tropenmedizin, Wilhelmstrasse 27, 72074 Tübingen, Germany, Telephone: 49-7071-29-80234, Fax: 49-7071-29-5189, E-mail: steffen.borrmann@uni-tuebingen.de. David Wypij, Clinical Research Program, Children's Hospital, 300 Longwood Avenue, Boston, MA 02115, Telephone: 617-355-2641, Fax: 617-355-2312, E-mail: david.wypij@childrens.harvard.edu. Terrie E. Taylor, Department of Internal Medicine, B309-B West Fee Hall, Michigan State University, East Lansing, MI 48824-1315, Tele-

phone: 517-353-3211, Fax: 517-432-1062, E-mail: taylor@msu.edu. Peter G. Kremsner, Medical Research Unit, Albert Schweitzer Hospital, BP 118 Lambaréné, Gabon, Telephone: 241-58-10-99, Fax: 241-58-11-96 and Sektion Humanparasitologie, Institut für Tropenmedizin, Wilhelmstrasse 27, 72074 Tübingen, Germany, Telephone: 49-7071-29-87179, Fax: 49-7071-29-5189, E-mail: peter.kremsner@uni-tuebingen.de.

REFERENCES

1. Lewis JA, Sommers CO, 2003. Personal data assistants: using new technology to enhance nursing practice. *MCN Am J Matern Child Nurs* 28: 66-71.
2. Lapinsky SE, Weshler J, Mehta S, Varkul M, Hallett D, Steward TE, 2001. Handheld computers in critical care. *Crit Care* 5: 227-231.
3. Jamison RN, Raymond SA, Levine JG, Slawsby EA, Nedeljkovic SS, Katz NP, 2001. Electronic diaries for monitoring chronic pain: 1-year validation study. *Pain* 91: 277-285.
4. van den Kerkhof EG, Goldstein DH, Lane J, Rimmer MJ, van Dijk JP, 2003. Using a personal digital assistant enhances gathering of patient data on an acute pain management service: a pilot study. *Can J Anaesth* 50: 368-375.
5. Reitmaier P, Dupret A, Cutting WA, 1987. Better health data with a portable microcomputer at the periphery: an anthropometric survey in Cape Verde. *Bull World Health Organ* 65: 651-657.
6. Forster D, Behrens RH, Campbell H, Byass P, 1991. Evaluation of a computerized field data collection system for health surveys. *Bull World Health Organ* 69: 107-111.