Can Health Workers capture data using a generic mobile phone with sufficient accuracy for Capture at Source to be used for Clinical Research Purposes?

MICHAEL WORKMAN

WRKMIC001

Minor Dissertation presented in partial fulfilment of the requirements for the degree of

Master of Science (Information Technology)

Faculty of Science

University of Cape Town

May 2014

Supervisors: Professor Gary Marsden and Dr Anthony Hawkridge

I dedicate this dissertation to my family Bumble, James & Jessica

DECLARATION

I know the meaning of plagiarism and declare that all of the work in the dissertation, save for that which is properly acknowledged, is my own.

WRKMIC001

Can Health Workers capture data using a generic mobile phone with sufficient accuracy for Capture at Source to be used for Clinical Research Purposes?

Michael Workman

31 October 2013

ABSTRACT

Objective:

To determine the accuracy, measured by error rate, with which Clinical Research Workers (CRWs), with minimal experience in data entry, could capture data on a feature phone during an interview using two different mobile phone applications, compared to the accuracy with which they could record data on paper Case Report Forms (CRFs).

Design:

A comparative study was performed where 10 participating CRWs performed 90 mock interviews using either paper CRFs or one of two mobile phone applications. The phone applications were a commonly used open source application and an application custom built for this study that followed a simplified, less flexible user interface paradigm. The answers to the interview questions were randomly generated and provided to the interviewees in sealed envelopes prior to the scheduling of the mock interview.

Error rates of the captured data were calculated relative to the randomly generated expected answers.

Results and Conclusion:

The study aimed to show that error rates of clinical research data captured using a mobile phone application would not be inferior to data recorded on paper CRFs.

For the custom application, this desired result was not found unequivocally. An error in judgment when designing the custom phone application resulted in dates being captured in a manner unfamiliar to the study participants, leading to high error rates for this type of data. If this error is condoned by excluding the date type from the results for the custom application, the custom application is shown to be non-inferior, at the 95% confidence level, to standard paper forms when capturing data for clinical research.

Analysis of the results for the open source application showed that using this application for data capture was inferior to paper CRFs.

Secondary analysis showed that error rates for data captured on the custom mobile phone application by non-computer literate users were significantly lower at the 95% confidence level than the error rates for data recorded by the same users on paper and for data captured by computer literate users using the custom application. This result confirms that even non-computer literate users can capture data accurately using a feature phone with a simplified user interface.

ACKNOWLEDGEMENTS

I would like to thank the following people:

My supervisors, Prof. Gary Marsden and Dr Tony Hawkridge, for their encouragement, guidance and constant support.

The staff at SATVI: Dr Hassan Mohammed for permitting his staff to participate in this study, Dr Hennie Geldenhuys for preparing the English and Afrikaans versions of the Case Report Form (CRF), Ms Tembakazi Daki for the Xhosa version of the CRF, Ms Marwou de Kock for all arrangements at the Worcester site, and Clinical Research Workers Mss Lorraine Coetzee, Christal Ferus, Xoliswa Fortewu, Diann Gempies, Gail Jacobs, Gloria Khomba, Nambitha Nqakala, Rose Ockhuis, Elizabeth Swarts and Nokuzola Williams who were willing and eager participants in capturing the data.

Prof Tim Dunne for checking my statistical assumptions and giving me some guidance on the structure.

Ms Dagmar Tilmer of CellLife who provided support for the implementation of openXdata.

Mr Shai Almog and Mr Eric Coolman of the CodenameOne User Forums for assistance with CodenameOne and in particular with the predictive text function.

Finally, my family for their continual encouragement and support, my wife Lesley for the support with statistics and especially my son James who showed me what could be achieved when programming a feature phone.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

AIDS	Acquired immunodeficiency syndrome
API	Application programming interfaces
CDMS	Clinical data management system
CFR 21	Code of Federal Regulations Title 21 Part 11
CHW	Community health worker
CLDC	Connected limited device configuration
CRF	Case report form
CRW	Clinical research worker
CRW	Clinical research worker
eCRF	Electronic case report form
EDC	Electronic data collection
EDCTP	European and Developing Countries Clinical Trials Partnership
FDA	Food and Drug Administration, United States of America
GCP	Guideline for Good Clinical Practice E6(R1)
HIPAA	The Health Insurance Portability and Accountability Act of 19
HIV	Human immunodeficiency virus
HIV/AIDS	Human immunodeficiency virus/acquired immunodeficiency syndrome
IDC	Industrial Development Corporation
J2ME	Java Micro Edition version 2
LWUIT	Lightweight User Interface Toolkit
MIDP	mobile information device profile
NIH	National Institutes of Health (United States of America)
PC	Personal computer
PDA	Personal digital assistant
SATVI	South African Tuberculosis Vaccine Research Initiative
SATVI	South African Tuberculosis Vaccine Initiative
TB	Tuberculosis
VM	Virtual machine
WHO	World Health Organisation

1. INTRODUCTION

This document covers the research conducted, findings made and conclusions drawn while attempting to answer the question:

> *Can trained clinical research workers, with little or no experience in data capture, use generic mobile phones (feature phones) to capture data with sufficient accuracy for 'capture at source' to be accepted by the clinical research domain?*

The research study answers this question by comparing the accuracy of collecting data by Clinical Research Workers (CRWs) using a feature phone, with either an open-source application or a custom-built application, with the accuracy of conventional paper-based data collection at a clinical research facility near Cape Town.

1.1 Importance of this research

The burden of disease affects developing countries disproportionately, especially the three primary poverty-related diseases, acquired immunodeficiency syndrome (AIDS), malaria and tuberculosis (TB). Africa in particular has a growing burden of these diseases. For instance, 90% of malaria deaths occur in sub-Saharan Africa (WHO 2013a), 95% of TB cases and deaths occur in Africa and the rest of the developing world (WHO 2013b) and almost 70% of all people living with human immunodeficiency virus/acquired immunodeficiency syndrome (HIV/AIDS) live in sub-Saharan Africa (WHO 2012b).

Clinical research is an important part of addressing this far-reaching health problem. According to the (Academy of Science of South Africa 2009), "[c]linical research in a developing country … contributes to health care at all levels by identifying the causes of problems, facilitating diagnosis, improving the efficiency and effectiveness of care".

There are sometimes different strains of these diseases in Africa or other developing regions than in the developed world. For example, human immunodeficiency syndrome (HIV) sub-type B is the most researched strain of this disease and is dominant in the Americas and Europe, while in Africa, the non-B sub-types are dominant (Santos & Soares 2010). Drug resistant strains of these diseases are also appearing, for example, artemisinin-resistant malaria has been reported in remote areas of Cambodia (WHO 2012a). Thus, the results of clinical research conducted in a certain region cannot be assumed to be constant and applied in another region. The research has to be performed on the target populations themselves, which are often located in the remote regions of the developing world.

A means of easily recording research data in the remote regions will be of significant benefit to the clinical research community, locally in South Africa and in the developing regions, as it could lead to better understanding of the burden of disease, provide better access to health care and recommended treatments and reduce costs and time to get new drugs to market.

The Food and Drug Administration of the United States (FDA) has recommended the use of electronic data collection (EDC) at source as a solution, specifying the following requirements for EDC (FDA 2013b):

"Capturing source data electronically and transmitting it to the eCRF should:

- Eliminate unnecessary duplication of data;
- Reduce the possibility for transcription errors;
- Encourage entering source data during a subject's visit, where appropriate;
- Eliminate transcription of source data prior to entry into an eCRF;
- Facilitate remote monitoring of data;
- Promote real-time access for data review; and
- Facilitate the collection of accurate and complete data."

1.2 Clinical Research

The report "Revitalising Clinical Research in South Africa" (Academy of Science of South Africa 2009) describes clinical research as "... research primarily conducted with human participants … during which investigators examine the mechanisms, causation, detection, progression and reversal of human disease". The report categorises clinical research as epidemiological (causes, distribution and control of disease), pharmaceutical (drug development) and translational (bridging the gap between the basic science and clinical application).

Within the epidemiological category, there are various types of studies, and one taxonomy used to define the types of studies is that by Grimes & Schulz (2002), and the algorithm used by Grimes is illustrated in Figure 1. In an experimental study and in a cohort or cross-sectional study, data is obtained during a consultation or interview with the study participant when (s)he visits a participating medical facility or when a research worker visits the participant. For case-control studies, the aim is to determine the potential cause of or influence on a particular medical condition and this data is obtained from historic patient records.

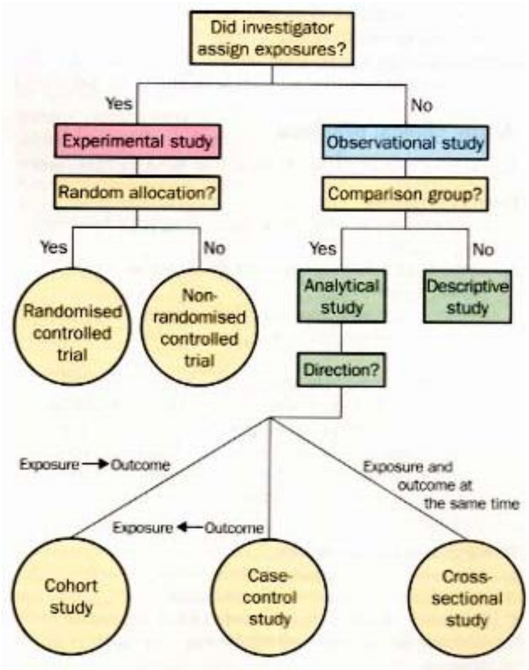

Figure 1 Algorithm for classification of types of clinical research

Pharmaceutical trials are conducted in four phases. Phases 1 and 2 are small-scale trials and focus on dose determination, safety and efficacy. Only once these trials are successful, the large scale Phase 3 trial is performed, which ascertains whether the drug is suitable for its intended purpose. These Phase 3 trials often require a medium-term (six months to two years) follow-up by research workers. Finally, once regulatory approval has been obtained, Phase 4 studies are conducted to follow up on the efficacy of the drug as it is dispensed to patients and to identify any long-term adverse events. As this efficacy must be assessed over a period of time, the participants must be followed up on, often by research workers visiting and interviewing the participants.

Translational Research studies the mechanism of disease and how it could be controlled. There is no human participation and hence participant data is not collected by research workers. Accordingly translational research falls outside the scope of this study.

1.2.1 Data collection for clinical research

All the above study and trial types require high quality data to be captured and subsequently analysed. This data has historically been recorded by research staff on Case Report Forms (CRFs) and then captured centrally on computer using a Clinical Data Management System (CDMS) (Lu & Su 2010), in custom-built databases (typically Microsoft Access) or even on spreadsheets.

In the meantime, so-called EDC has been introduced to the CDMS systems (Lu & Su 2010) where the aim is that the clinicians and research workers could capture the data directly during the interview and avoid using the paper CRFs. However, this aim new system has proven to only be successful to a certain extent: Research workers can capture the data at the site and thus avoid transporting them to the central office, but hard copy CRFs are still completed (Scurtu 2013).

Where research workers visit patients, they have to take the correct CRFs with them, complete them and then return them to site for data capture. This transporting of CRFs requires rigorous document control to ensure no pages are lost or misfiled.

A means of capturing the data directly using a mobile device at the point of interview would simplify the data collection process for the clinical research workers out in the field. It would also ease the management and logistical tasks related to the secure handling of large numbers of paper forms.

1.2.2 Data Quality

For clinical research, the quality of the data required is critical to any outcomes that may be determined. Invalid outcomes, whether caused by poor data or due to other reasons, can cause major health issues. An example of an invalid outcome is the fraudulent linkage of autism to the MMR vaccine administered during early childhood (Godlee, Smith & Marcovitch 2011). This caused many parents to avoid vaccinating their children, which lead to an upsurge in childhood diseases. It is for this reason that the FDA mandates very strict standards regarding clinical trial data and its provenance (FDA 2013a). To this end, there is a generally accepted target of 99.9% accuracy for key outcomes when comparing the CRFs to the captured data (Kleppinger & Ball 2010; Nahm, Pieper & Cunningham 2008), although this measure does not address issues where the patient may provide the incorrect answer in the first place, nor where the interviewer records the incorrect value on the CRF (Nahm, Pieper & Cunningham 2008).

On the normal paper CRFs, there are often a number of fields where the clinical research worker records the answer in words, or even sentences. This is known as free text and is naturally unstructured. Free text cannot be analysed due to multiple spellings, misspellings, abbreviations, terminology differences and even multiple languages. Therefore, after being captured, the data is 'cleaned' and categorised for analysis. However, even after cleaning and categorisation, the data may contain errors as misspellings and may not provide sufficient information for correct categorisation(Hogan & Wagner 1996).

From the above, it can be seen that it is critical to a clinical research study that data for analysis is captured correctly and completely. Ensuring this completeness of data usually requires comprehensive data recording, collection, cleaning, storage and management procedures to be rigorously enforced.

1.2.3 Statutory Regulation

There are various national and international laws and regulations regarding data and electronic data management systems. The ones that generally affect South African clinical research are:

1.2.3.1 South African Laws and Regulations

- The Electronic Communications and Transactions Act 2012 (Republic of South Africa 2012) covers the security and protection of data from unauthorised use; and
- The Protection of Personal Information Bill (Republic of South Africa 2009) was approved the National Council of Provinces on 20th August 2013 and has been referred to the President for his assent (Parliamentary Communication Services 2013). This act covers data privacy requirements, specifically mentioning health information related to the individual.

1.2.3.2 International Laws and Regulations

"Much of the clinical and epidemiological research now being performed in South Africa is funded by foreign, non-private foundations and government bodies…" (Academy of Science of South Africa 2009). This funding model often requires the study data to be submitted to the funder, which implies that the regulatory requirements applicable to the funder have to be met. Common sources of funding are bodies such as the National Institutes of Health (NIH) in the United States of America and the European and Developing Countries Clinical Trials Partnership (EDCTP) in the European Union.

The following laws and regulations are applicable in the United States.

- The Code of Federal Regulations Title 21 Part 11 Electronic Records; Electronic Signatures (FDA 2013a) (CFR 21), which covers electronic records and electronic signatures and what is required to ensure they are reliable and non-repudiable;
- The Guideline for Good Clinical Practice E6(R1) (International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use 1996) (ICH-GCP), which covers patient confidentiality; and
- The Health Insurance Portability and Accountability Act of 1996 (United States of America House of Representatives & Senate 1996) (HIPAA), which covers protection of personal health information.

The various countries that comprise the European Union have enacted laws covering data used in clinical research as required by Directive 2001/20/EC of the European

Parliament and of the Council of the European Union (The European Parliament & The Council of the European Union 2001) This directive specifies the requirements for, amongst other things, patient confidentiality as well as security and integrity of data.

The FDA has issued the "Guidance for Industry Electronic Source Data in Clinical Investigations" (FDA 2013b), which proposes to regulate mobile applications classified as 'medical devices', such as heart rate monitors. While mobile data collection applications, such as those that are the subject of this study, are used for medical research, they are not classified as medical devices in terms of the guidance and therefore no additional regulatory requirements stem from this guidance.

The above legal and regulatory requirements are not directly applicable to this study, but their requirements were considered to ensure that further work in this field can be compliant.

1.3 Clinical Research Database

Data for clinical research studies that comply with the above laws and regulations must be stored in a validated Clinical Research Database (CRDB) (FDA 2013a). A widely adopted open source CRDB that is compliant, is OpenClinica (Akaza Research 2012), which was used in a comparative study of data capture accuracy on various types of computers, including hand-held devices (Walther *et al.* 2011). Where data is captured using a mobile device, the software on that device is considered to form part of the CRDB application and hence must conform to the same laws and regulations. Further, the requisite audit trail in the CRDB must include the part of the data capture process that was performed on the mobile device.

1.4 Clinical Research Workers

Typically, data is collected by Clinical Research Workers (CRWs) who are lay workers (i.e. not professional) who should match the social demographics of the participants in the research studies and are generally recruited from the communities where the research participants live. They would have to have obtained a school-leaving qualification which implies they would be literate and numerate, but not necessarily have passed a high-school exit examination such as the National Senior Certificate. They would be fluent in the language(s) spoken in the community, although that language might not be their home language. Where English is the medium of communication within the research study, they would also be fluent in English. They would not necessarily be computer literate. They would be trained in Good Clinical Practice (GCP), which is an international ethical and scientific quality standard for designing, conducting and recording of research involving humans. Apart from GCP they need not have any specialist training.

The lack of computer literacy is in accordance with Ziphorah, who reports that, in South Africa, people in remote areas are unlikely to be computer literate and computer literacy in the main centres is low (Ziphorah 2011).

CRWs are a class of Community Health Workers (CHW) who perform similar roles to the CRWs in terms of health monitoring and education for the community, and who may, in addition, provide basic interventional health services such as assisting with childbirth. In South Africa, the training for CHWs is governed by the Health and Welfare Sector Education and Training Authority, which provides certification programs for CHWs (SAQA 2012).[1] Other countries also make use of CHWs, such as Ethiopia's Health Extension Workers (WHO 2008).

This profile of CRWs indicates that as they are not necessarily computer literate, the use of mobile/portable computers is unlikely to be a good solution for data capture by them, and non-computer devices such as mobile phones must be considered.

Due to the possibility that the CRWs would possibly have lower educational qualifications, it would also be desirable if the collection of data could avoid the CRW having to record textual information, particularly containing long or complex medical terminology.

[1] SAQA: South African Qualifications Authority.

1.5 Mobile phones for clinical research

For various reasons outlined below, mobile phones have the potential to enhance the collection and capture of data for clinical research in remote areas in Africa and other regions of the developing world.

1.5.1 Challenges of the developing world for Clinical Research

Africa and the rest of the developing world bring their own challenges to effective clinical research, especially in the remote areas, and these challenges can easily impact on the quality of the captured data:

- Transport infrastructure is poor with badly maintained roads between the major centres and often only dirt tracks and footpaths to the outlying villages (Iarossi 2013);
- High-speed communications networks required for Internet access are almost non-existent outside the major cities (Richardson 2006); and
- Electric power for computing and network equipment is usually only available in major centres and is unreliable at best (Iarossi 2013).

All the above challenges have the potential to hamper the traditional collection and capture of the research data using paper forms and/or computers on high-speed networks, and possibly impacting its accuracy.

1.5.2 Mobile phones in the developing world

Mobile phone technology has been a success in Africa where, apart from a few countries/regions, more than half the populace usually has access to mobile phones (Delmas 2013). This access is not limited to urban areas as mobile phone usage is widespread in the rural areas of South Africa as well (Skuse & Cousins 2007). Mobile phones have not only provided modern communications in the remote areas, but have changed the way that people live. "Across Africa, mobile technology is becoming a cornerstone for industries like healthcare and agriculture. For millions of people, it is making banking truly accessible for the first time" (Hoeller & McHenry 2011).

1.5.3 Smartphones vs. Feature Phones

There is no generally agreed definition for smartphones and feature phones, but at this point, it is sufficient to define them as:

- Smartphones cost more than feature phones and use a more powerful operating system such as Apple's iOS, Google's Android, the Blackberry OS, or Microsoft Windows.
- Feature phones use manufacturer specific operating systems such as Nokia's Symbian 60.

There is much interest in the utilisation of smartphones in various contexts and for various applications. Forecasts suggest that smartphones will outnumber generic feature phones in the very near future, but while these forecasts may be true on a global scale, projections for Africa for 2016 show that 70% of the phones used will still be feature phones (Portio Research 2013).

The Industrial Development Corporation (IDC) states that "the smartphone market is growing rapidly across the region … in poorer countries and where mobile operators do not subsidize phone purchases on usage contracts feature phones are still the majority of sales in units sold" (IDC 2013). Gartner also demonstrates that in the third quarter 2012, feature phone shipments to Africa were far higher than smartphone shipments.

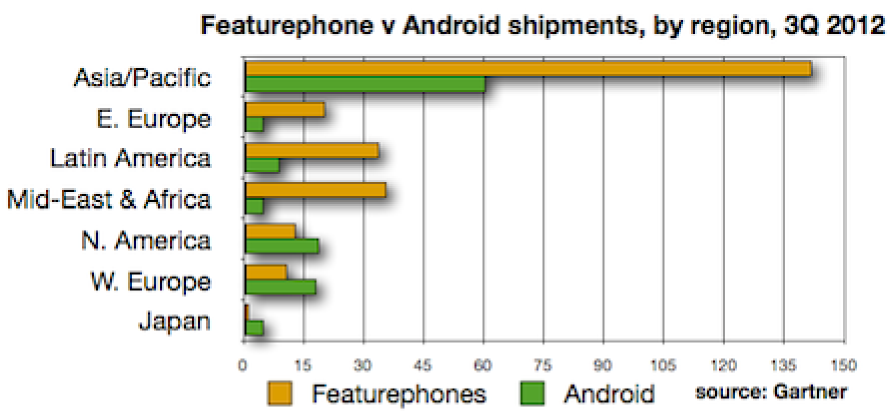

Figure 2 Mobile phone shipments in the third quarter of 2012

1.5.4 Mobile phones for data capture

Given the challenges identified in Section 1.5.1 and the ubiquity of the feature phone in Africa (Sections 1.5.2 and 1.5.3), using a feature phone for data capture may be beneficial to clinical research in Africa. However, the lack of high-speed networks, mentioned above, could inhibit such use and great care will need to be taken in the design of such an application to enable it to operate in conditions of slow and intermittent network connectivity.

1.6 Research problem statement and hypothesis

Clinical research in Africa faces significant challenges, especially in remote areas. In order to enhance clinical trials and research in remote areas, the use of a feature phone as a data capture device with a specialised user interface design could

- Simplify the logistics (no need to transport paper forms);
- Enable data collection and capture in remote areas (not dependent on a high-speed network);
- Use a familiar device and avoid data capture on computers, laptops, tablets and smartphones (no requirement for scarce computer literate workers); and
- Enable non-computer literate CRWs to capture data at the source (eliminates the need for paper and its transport to major centres).

The above benefits depend on the ability of the typical clinical research worker to capture data with sufficient accuracy on a mobile device with which they are familiar, such as a feature phone.

1.6.1 Hypothesis

A data capture application for a generic feature phone with a simple user interface design that avoids free text entry and that supports multiple user languages should contribute to the accurate capture of clinical research data by a typical research worker who is not necessarily computer literate.

Arguments supporting this hypothesis are:

- Although the average clinical research worker may not be computer literate, they are probably 'mobile phone literate'. A mobile phone application that uses the familiar interface paradigm of the mobile phone will enhance their ability to use the application with minimal errors. It is conjectured that provision for clinical research workers to use their own mobile phone will enhance this ability even further. Supporting this conjecture, is the extensive use of the mobile phone for mobile banking (Hoeller & McHenry 2011) as banking is a multi-step transaction where a user error (e.g. entering the wrong account number) can leave the user poorer than before.

- Many current mobile applications impose a significant learning curve on the users and some use user interface devices that are considered inappropriate for the small display and cramped keypad of the feature phone handset. A simple consistent user interface will allow the user to focus on the interview and the recording of the answer and minimise the distraction of using the handheld device. This simplified user interface should lead to reduced error rates in the capture of data.
- A common area of error in clinical research data, whether captured electronically or written on a paper form, is free text items. Some of the causes are misspellings, abbreviations and different languages, as mentioned before. Constraint of the problem space and implementation of user interface devices to exclude free text should enhance accuracy.
- In a multi-language environment, interviews in the home language of the participants could reduce scope for error caused by misunderstanding and/or 'on the fly' translation.

1.7 Scope of Study

The aim of this study was to investigate the ability of CRWs to capture data accurately on feature phones. Therefore, the scope was limited to implementing the necessary data capture applications with a mechanism to record the captured data and measure error rates. The scope excluded any technical and procedural requirements that would normally be necessary for an actual clinical research study.

The scope was limited to clinical research interviews/questionnaires with the answers being captured by CRWs at the Worcester site of the South African

Tuberculosis Vaccine Research Initiative (SATVI). However, the results and conclusions of this study should be generalisable to other health workers, other clinical research studies and other regions in the Africa, because the profile of the CRW is very similar to that of other CHWs (Section 1.4) and because clinical research is conducted in terms of internationally recognised regulations (Section 1.2.3).

1.8 Organisation of this dissertation

This dissertation is organised in the following structure:

- In Chapter 2 the prior research that has been done on this topic is reviewed and aspects of usability that affect the feature phone that will be used for this research are discussed;
- In Chapter 3 the methodology for this research study is specified and the characteristics of the participants, as these characteristics relate to the study, are given;
- In Chapters 4, 5 and 6 the implementation of the mechanisms to perform this research study are covered:
    - In Chapter 4, transaction analysis is covered to identify the user tasks required for clinical research data capture, and following on from this analysis, the approach to implementing these tasks on a feature phone are discussed;
    - In Chapter 5 a set of open source applications used in clinical research studies are reviewed against the outcomes of Chapter 4, followed by a selection of one of these applications for this research study. After that, the implementation of the selected open source application is described; and
    - The design, development and implementation of the custom application are documented in Chapter 6;
- In Chapter 7 the results of the research are documented, these results are analysed in Chapter 8; and
- Finally, the conclusions are documented in Chapter 9.

2. PRIOR RESEARCH AND WORK

2.1 Prior Research

In order to establish the status of high quality data and electronic data capture for clinical research using mobile devices, a literature search was conducted for clinical research data capture, as well as data capture using mobile devices in general.

To provide an informed view of the usability of feature phones, a search was performed for mobile interface design. As the focus of this research is the use of feature phones, publications specifically relating to smartphones and personal digital assistants (PDAs) were ignored as the user interface of these devices is significantly different to that of the feature phone (they have a larger display, a QWERTY keyboard and/or a touch-screen) and most of these publications were iPhone or Android focused.

The search showed that research has been done on the use of mobile devices for capturing data for surveys and other research, but the results are not conclusive, especially when applied to clinical research in the developing world.

Many of these studies have been qualitative and focused on speed, cost, ease of use and 'trackability' of the interviewer. For example, Tomlinson investigated the feasibility of and the extent to which community health workers could be trained to collect data using mobile phones in a large baseline survey (Tomlinson *et al.* 2009). He identifies that further research is needed into data accuracy and other aspects and says, "Rigorous controlled trials comparing data accuracy, readability, reliability and validity checks comparing paper based approaches, PDA's (*sic.*) and mobile phones are needed." Ogunmefun compared data capture on a mobile phone with data recorded on paper forms for community surveys and measured the perceived value in supporting management of the community care giver programme (Ogunmefun, Mothibe & Friedman 2010).

Very little research has been done regarding the issue of data accuracy when using mobile devices. That which has been done, has typically used PDAs and not mobile phones and has recorded negative findings:

- Haller compared speed, accuracy and user satisfaction for data capture using a laptop and a PDA against paper forms. He found a 45% increase in errors and a seven-fold increase in missing data when using a hand-held device compared to a laptop (Haller *et al.* 2009);
- Shelby-James compared the accuracy of data captured on PDAs with the same data recorded by the same participant on paper. She found error rates using PDAs to be 6.75% (Shelby-James *et al.* 2007); and
- Walther compared data capture accuracy across three types of EDC devices, namely a netbook, a PDA and a tablet personal computer (PC). She found accuracy to be acceptable when using a netbook or tablet, but not when using a PDA (Walther *et al.* 2011). She identified the small screen, small fonts and use of a stylus with a web-based applications on a PDA as being issues that probably led to poor accuracy when using this device.

Studies have been conducted on mobile health (mHealth), namely the use of mobile phones in a support role for healthcare, but here too the results are inconclusive and not applicable to the developing world. For instance, Tatalović (2013) found that none of the 42 studies he reviewed were conducted in the developing world, while Free *et al.* (2013) found that only three out of 75 studies were conducted in low or middle-income settings and that "many had methodological problems likely to affect the accuracy of their findings".

Regarding measurement of data quality or accuracy for clinical research, in a definition of standard metrics for EDC systems (EDM Forum 2004) the quality metrics are defined according to:

- The number of unlocks after database lock for final analysis (in order to correct data);
- The number of help desk calls per 1000 fields; and
- The number of manual queries raised by a sponsor per 1000 fields.

Despite being termed 'metrics of quality', none of these metrics measure the accuracy or correctness of the data.

Summarising the above, it appears there has been no research into accurate data capture using a feature phone for clinical or other research. There is also no research into data quality measurement in clinical research where correctness is the measure of quality.

Clearly more work needs to be done in this area and a study that measures accuracy of data when captured on mobile phones will possibly stimulate the uptake of such use for clinical research in remote areas.

2.2 The Feature Phone

There is no universal definition of a feature phone. For instance, Nokia defines a feature phone as one without web access (Nokia 2012), while PC Magazine states that feature phones usually have web access (PC Magazine Encyclopaedia [n.d.]). The general understanding is that it is a mid-priced device falling between the basic phone and the smartphone, and with capabilities between the two. However, the features that used to be only found on smartphones are now available on the mid-range feature phones as well.

For the purposes of this study, a feature phone will be defined as one with a small screen (less than 2.5" diagonal) and a 12-key keypad (i.e. physically smaller than a smartphone), which can access the Internet and install applications (a functionality that is not available on a basic phone).

2.2.1 Application architectures

Each handset manufacturer has its own application architecture, some more than one. (Nokia Symbian 40 and Symbian 60). This fragmentation has made it difficult for third parties to develop applications for these handsets. The relatively recent introduction of Apple's iPhone and the release of the Android operating system for mobile devices have led to even more fragmentation of the architecture space.

The standardised architectures noted below have been developed and implemented by most manufacturers on their feature phone platforms.

2.2.1.1 CDLC and MIDP

Connected Limited Device Configuration (CLDC) and Mobile Information Device Profile (MIDP) are specifications published for the use of Java on devices with very limited resources, such as mobile phones and PDAs. CLDC defines the low-level set of Java Virtual Machine (VM) features that must be present on all embedded devices. MIDP sits on top of CLDC and defines the set of application programming interfaces (APIs) that are required for the use of Java on mobile phones.

2.2.1.2 J2ME

Sun Microsystems standardised Java Micro Edition (J2ME) in 2000 and feature phone manufacturers have included a J2ME runtime environment in their application architectures, which has allowed third parties to deliver applications on the feature phones across all manufacturers.

2.2.1.3 LWUIT

The Lightweight User Interface Toolkit (LWUIT) was developed by Sun Microsystems, and later taken over by Oracle. This toolkit was investigated for use as the application architecture for this study. While investigating LWUIT, CodenameOne (discussed in the next section) was found and as this platform seemed to offer better customisability than LWUIT, further investigation into LWUIT was stopped.

2.2.1.4 CodenameOne

CodenameOne (http://www.codenameone.com/) is a startup created by ex-Sun staff who have forked the LWUIT code base and extended it. Their aim is to create a cross-platform tool which will cover Android, Apple iOS, Windows Phone 7 and 8, Blackberry RIM and J2ME platforms. While the cross-platform capability is not required for this study, CodenameOne provides the capability at a sufficiently low level to create the desired user interface.

Other options such as J2ME Polish were considered, but not taken further as a satisfactory user interface could be implemented using CodenameOne.

2.2.2 Display

The display size of the feature phone for this study has been limited to a maximum of 2.5" (Section 2.4), but was often smaller. This restriction implies that to show a reasonable amount of information of the display, the fonts have to be relatively small. Care must be taken to ensure the fonts are still large enough to read, as Walther referred to problems encountered due to small font sizes used on the PDAs (Walther *et al.* 2011).

2.2.3 Keypad

The feature phone's keypad has a set of 12 keys arranged in four rows of three keys each, as defined by the ITU (ITU-E.161 2001) and shown in Figure 3. Above these keys are other keys used for various functions, including the 'soft keys', which are programmable function keys and are needed when using and interacting with the phone handset. A key requirement for this study is to ensure the users are able to collect data accurately, even when using this very limited keyboard.

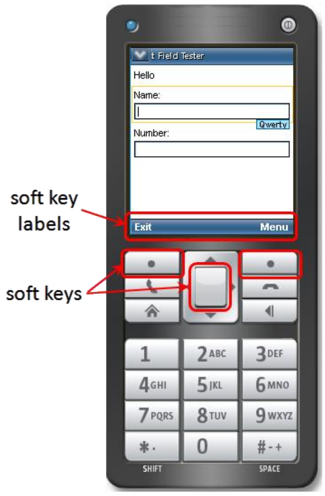

Figure 3 Feature phone with soft keys

2.2.4 Predictive text

Predictive text is a technique used whereby the application tries to predict the desired textual output based on the press of specific key combinations and a list of words which are mapped to the key combinations that represent each word (Dunlop & Crossan 2000). One common predictive text technology is Text on 9 keys (T9), which is often used on feature phones where the alphabet is overlaid on nine of the 12 keys. Predictive text offers an effective alternative to free text provided the list of words is pre-defined (refer to Section 1.2.2).

2.2.5 J2ME soft keys

Soft keys on a mobile phone are a form of programmable function keys typically found on computer keyboards. On the mobile phone's limited keypad, there are at least two soft keys (left and right). Often there is a third centre key with four cursor movement keys, or navigation/directional keys.

The soft keys are implemented by the manufacturer to suit the keypad and operating system of each distinct phone model. For phones with physical directional keys, keys are mapped to the J2ME events 'fire', 'up', 'down', 'left' and 'right'. For phones without the physical directional keys, the numeric keys are usually mapped to the J2ME events as 5 (fire), 2 (up), 8 (down), 4 (left) and 6 (right). Labels are usually displayed at the bottom of the display just above the left, centre and right soft keys that indicate the action associated with each key.

However, when implementing the J2ME standard, the manufacturers are inconsistent with the handling of the soft keys. In many cases, additional commands are added that are not defined by the developer. Figure 4 shows the Nokia Symbian 40 and Symbian 60 operating systems manipulating the soft keys: Initially, only the 'OK' and 'Exit' keys are set. When the text field gains focus, the 'OK' command is moved from the left to the centre soft key and an 'Options' sub-menu is added to the left key on a Symbian 60 device. On Symbian 40, an 'Options' sub-menu replaces the 'OK' command on the left soft key and the 'OK' command is included in the sub-menu 'Options'.

Initial softkey state | Textfield has focus - Symbian 60 | Textfield has focus - Symbian 40

Figure 4 Symbian 60 and Symbian 40 soft keys

Further, J2ME defines the soft keys with a soft key type, the 'negative' types being 'BACK', 'CANCEL', 'EXIT' and 'STOP' which are all very similar. Depending on which type you define, Nokia arranges the order on the left and right soft keys (and sometimes allocates the 'STOP' type to the phone's 'Power Off' key), irrespective of the order the programmer specifies (Nokia 2013). Samsung and Sony Ericsson implement the soft keys in a different order and allow the programmer more control over placement.

This idiosyncratic behaviour of the various handset platforms must be overcome when designing an application aimed at multiple handset types, especially as one of the usability requirements is predictability, which must be applied throughout the application (Ballard 2005).

As the native J2ME implementations would appear to be problematic for the user, CodenameOne was selected for the development of the custom application as it provided sufficient control over the layout and functionality of soft keys, as well as control over the placement of all display widgets in a manner similar to the Java Swing environment.

2.3 Open Source Application Review

A number of software houses have taken advantage of mobile technology and released mobile data collection applications for various models of mobile phones. Over 35 such applications were identified through an Internet search.

Of these applications, three were reviewed in detail by defining a CRF for each and performing test interviews using a Nokia 2750 handset (Symbian 60). The applications selected for review were Mobenzi Researcher by Clyral, mforms by openXdata and epiSurveyor by DataDyne.

2.3.1 Mobenzi Researcher

An abbreviated CRF was configured for the application covering all the data types defined in Section 3.1.2. Initially, the application could not handle repeating groups, but a special update was implemented to cater for the test repeating group. Two test interviews for each application were completed successfully, but various problems that could hamper usability were identified:

- The soft keys have a sub-menu of commands on the left-hand soft key. Although the primary action was on the right-hand soft key (usually 'Select'), all other possible actions were in the sub-menu, which implies that at least two clicks are required for these actions;
- When the answer required text to be captured, the question form would display the question and a simple text box for capturing the answer. However, as soon as the first character was pressed, a new form would be displayed for the text to be entered on the entire form (a J2ME text area) and this first character (the one that triggered this new form) was dropped and not shown on the form anymore. For example, to capture the interview number, the cursor would be positioned in the text field on the form. When the first digit (e.g. '2') was pressed, this event would cause a new blank form to open and the '2' would be discarded, forcing the user to enter '2' again;
- Repeating groups were enabled by a change of application, but in such a way that one first had to capture the number of repetitions required. In this case, the clinical research worker would have to ask the patient "How many symptoms do you have today?" before actually being able to capture the symptoms. In some circumstances, this restriction could be acceptable (for example, "How many children do you have?") but not in the case of recording signs and symptoms.
- Dates are captured with a date-picker widget. The use of this widget is acceptable when the date to be selected is in the current month, because the user then needs only to click the directional keys until the correct day is reached. However, capturing the date of birth of an adult would lead to a large number of clicks to reach the correct year.

2.3.2 openXdata with mforms

The openXdata application was set up on an Amazon Web Services (AWS) cloud service and a full CRF was configured for openXdata mforms. Issues encountered with the usability of openXdata mforms were the following:

- The soft keys are inconsistent and the labels are not clear. There are often two 'positive' commands 'Select' (meaning 'open for editing' or 'take this action') and 'OK' (meaning 'move on to next form or question'). Where there are only two commands, 'Select' would be the left-hand soft key. Where there are three commands, 'Select' would be the centre soft key and 'OK' would be the left-hand soft key. Where there are more than three commands, 'Options' would be a sub-menu for the left-hand soft key and 'Select' would usually (but not always) be the centre soft key.
- The 'home page' for openXdata is inconsistent:
    - If there are one or more unsent results, then the user is shown a form which lists the saved results as 'Data 1', 'Data 2', etc., without an identifying title or description. The user can then select a result to send to the server. If the user wishes to start a new interview, the 'New' command under the 'Options' sub-menu has to be selected.
    - If there are no unsent results, the user is positioned at question 1 of a new interview and the soft keys are set accordingly, even if the user does not desire to start a new interview.
- Repeating Groups are handled as a 'sub-questionnaire' with a 'title' on the main list of questions. When the user selects the title, (s)he is presented with a form that displays '{empty}', (s)he then has to know to take select the 'New' command (centre soft key) to create a new instance in the repeat cycle. After completion of the sub-questionnaire, the results are displayed as a comma-delimited string, and the commands 'New' (new repeat row) and 'OK' (all repeats complete; move to next question) are in the 'Options' sub-menu. The centre soft key is 'Select' (open highlighted answers for editing).

2.3.3 epiSurveyor

An abbreviated CRF, similar to that used with Mobenzi Researcher, was configured for the application covering all the data types defined in Section 3.1.2 below. Two test interviews were completed successfully, but significant problems were identified:

- epiSurveyor takes the Options sub-menu to the extreme – all tasks and actions apart from the 'Back' action are included in the sub-menu on the left-hand soft key, as shown in Figure 3.
- The right-hand soft key is for the 'Back' command (go back one form) in most instances, but it performs the 'Backspace' function during text entry. The centre soft key is not used.

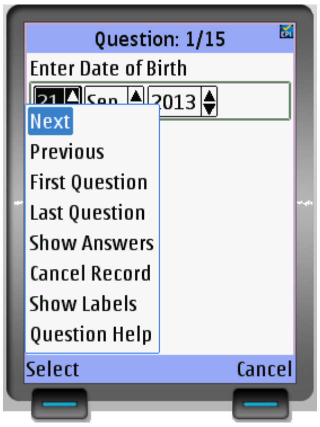

Figure 6 EpiSurveyor Options menu

Figure 6 EpiSurveyor date spinner widget

- Dates are captured with a spinner widget (shown in Figure 6). This widget is just as problematic as the Mobenzi Researcher's date picker as one can only press the up and down directional keys to change the date.

To summarise, the three applications reviewed all had usability issues, but none were bad enough to prevent them from being included in the selection of applications for the purpose of this study. A brief re-evaluation of these three applications that fell into the open source category did not identify one that appeared to be significantly better for the purposes of this study. Accordingly, openXdata was selected for this study as it is supported locally and thus assistance that might be needed could be obtained more easily.

2.4 Usability principles and guidelines for mobile phones

Usability consultant and author Jakob Nielsen states that, when browsing the Internet, the mobile experience is "miserable" (Nielsen 2011a) and lists various reasons for making this statement. Although his research was directed at website usability when using a mobile device, he identified two aspects that relate to the device itself: small screen and awkward input. The usability design must take into account these constraining characteristics of the mobile handset.

2.4.1 Dix's principles of usability

Dix (Dix *et al.* 2004) defines principles to support usability of information communications technology (ICT) under three conceptual headings:

- **Learnability:** Learnability covers the ease with which new users can start using a new application;
- **Flexibility:** Flexibility allows users to achieve the same outcome in different ways; and
- **Robustness:** Robustness, in this context, relates to the support for the user to complete the current task correctly.

These principles apply to ICT devices in general, including mobile phones. However, the drawbacks of the mobile phone identified by Nielsen above inhibit much of the application of these principles to the feature phone; for example, Nielsen states "… you should thus limit the navigation options, because you can't show full contextual information on every screen …"(Nielsen 2011b).

2.4.2 Shneiderman's 'Golden Rules'

Shneiderman defined 'Eight Golden Rules of Interface Design' (Shneiderman & Plaisant 2005) which is a set of design guidelines which generally conform to the principles to support usability defined by Dix. Shneiderman's rules are:

1. **Strive for consistency.**

   This rule is documented as a principle by Dix and supports the abstract principle of learnability. It is the principle most commonly applied, but unless there is a recognised standard, the application of this rule is only within a limited context.

For instance:

- Apple's user experience is consistent for the iPad and iPhone, and Apple demands the same consistency for all applications in its AppStore.
- Android as supplied by Google is consistent within each version of this operating system but has a different look and feel compared to Apple. Further, different versions of this operating system behave differently and handset vendors modify the user interface to suit their own paradigm. Applications in the Google Play application store need not support the same look and feel.
- Blackberry (RIM) has its own consistent user interface preferences which differs from those of the two operating systems mentioned above.
- Feature phone behaviour is inconsistent across the various manufacturers, and some do not fully comply with industry standards, such as J2ME.

2. **Enable frequent users to use shortcuts.**

Shortcuts enhance productivity for experienced users and fall under Dix's principle of substitutivity, which falls under the abstract principle of flexibility. They are not necessarily intuitive (e.g. inserting a non-breaking hyphen in Microsoft Word by pressing Ctrl + Shift + -) and thus conflict with the principle of learnability. With the cramped keypad of the feature phone, chorded input required for most shortcuts is nigh impossible. Therefore, this rule will be ignored, unless there are very specific needs such as the shortcuts that are provided for in the implementation of predictive text (Section 4.3.5).

3. **Offer informative feedback.**

Dix's principle of observability is met by this rule. It can be applied by, for example, moving to a new form, when the previous action was successful. This example meets with Shneiderman's principle of 'modest response to minor actions'. Further, the display of a dialog informing a user of the result of a significant task meets Dix's requirement for substantial response to major actions.

4. **Design dialog to yield closure.**

This rule is related to the previous rule of informative feedback. The principle of task conformance would include this rule. Initiation of an interview task from a menu, followed by a sequence of forms as the user moved through the interview, followed by a dialog when the interview was completed, would meet the requirements of this rule.

5. **Offer simple error handling.**

   In his description of this rule, Shneiderman extends it to include error prevention as much as error handling by saying, "As much as possible, design the system so the user cannot make a serious error." This extension aligns with Dix's principle of recoverability which he separates into 'forward recoverability' and 'backward recoverability'.

   - Backward recoverability is the preferred mechanism, as it allows the user to simply move one step back to review the situation and decide on the correct way forward.
   - Forward recovery is undertaken where the action cannot be undone (e.g. a database commit action). In this case, the user has to assess what is an 'unnatural' situation and then decide how to reverse its effect by applying a reversing transaction. This behaviour is not desirable, but is sometimes the only way to correct an error.

6. **Permit easy reversal of actions.**

   This rule aligns with Dix's 'forward recoverability' as noted in Point 5 above.

7. **Support internal locus of control.**

   This rule falls under Dix's conceptual principle of flexibility. When using a large screen device such as PC or tablet, flexibility of use can be supported as multiple applications can be viewed simultaneously, or at least shown to be active. However, in the context of a feature phone, this flexibility is not easy to achieve. Invariably, only a single application can be seen at any one time, although it can be interrupted by other phone-driven events. Any other running applications are difficult to find.

8. **Reduce short-term memory load.**

   This rule falls under Dix's learnability principle. There is a limited amount of processing that can be done in human short-term memory, so the user interface design should seek to minimise what must be remembered and processed. The feature phone inherently enforces this rule due to its miniscule display with a single form, so multiple windows, complex layouts etc. are not possible. The only case where this rule can be applied is ensuring that scrolling lists are not too long.

2.4.3 Ballard's usability for J2ME

Ballard (2005) defines specific guidelines for J2ME handsets covering predictability, simplicity, responsiveness, efficiency and optimisation. Her three key guidelines for the design of the user interface itself are:

1. **Predictability:** "Predictability is accomplished having the features match the user tasks ...." This guideline enables the user to readily understand what he is required to do and to know what will happen next.
2. **Simplicity:** Given the small display, cramped keypad and lack of pointing device, mobile applications should avoid complex interfaces and too many features.
3. **Efficiency:** "[U]sers are generally looking to accomplish a single task, and may be paying for connection ...." Therefore, the interface should be designed so the user can operate it as efficiently as possible, while still following the guidelines of simplicity and predictability.

As Shneiderman's eight golden rules are the most comprehensive, implement most of Dix's principles and can be applied to the feature phone with some limitations, they were used for the user interface design. Ballard's rules for J2ME devices were also taken into consideration as there were situations where Shneiderman's rules are not applicable to the small devices.

This chapter has described the problem space of this study and identified the 'building blocks' for the applications that were used to investigate solutions for the problem investigated in this study. The next chapter (Chapter 3) defines the research methodology and design for this investigation and later chapters cover the requirements for the custom application(Chapter 4), the custom user interface design (Section 4.3) and its implementation (Chapter (5) as well as the implementation of openXdata (Chapter 6).

3. RESEARCH METHODOLOGY AND DESIGN

The aim of this study was to investigate and evaluate the question posed in Chapter 1:

> *Can trained clinical research workers, with little or no experience in data capture, use generic mobile phones (feature phones) to capture data with sufficient accuracy for 'capture at source' to be accepted by the clinical research domain?*

A comparative study was performed comparing the accuracy of data recorded on paper CRFs (which was used as the control) to the accuracy of data captured using mobile phone applications during a mock interview. The dependent variable of interest was the error rate which was determined by comparing the answers captured or recorded during the mock interviews to the known correct answers.

This form of study was selected as it is commonly used in clinical research when evaluating the benefit of new methods compared to an existing method (the control), such as was done when evaluating a different method of vaccination of children against TB (Hawkridge *et al.* 2008). This methodology was previously adopted when comparing captured electronic data using various mobile computing devices with paper forms (Walther *et al.* 2011).

As well as performing a quantitative analysis, a qualitative analysis of the ease-of-use of the two applications was conducted. This was done to help identify any shortcomings in the usability of both applications and influence further developments of MyApp, the custom application that was developed for this study.

3.1 Study Design

As noted above, there are a number of open source applications for mobile phones that are currently used for data capture in various forms of surveys and other research applications. One of these open source applications, namely openXdata, was used as a basis for this study. As these applications were considered to have usability issues (Section 2.3) a custom application(MyApp) with a simplified interface was also developed and assessed.

The CRF was structured to include the types of questions which could affect accuracy in different ways, depending on the medium used to record/capture the answer. For example, when recording data on a paper form, a question about the gender of the participant may be followed by a question regarding child bearing potential (CBP), which is to be completed only if the participant is female. The CRW must remember to skip this question if the participant is male. On the other hand, it would not be possible for the CRW to capture an answer regarding CBP for male participants when using a data capture application, as the CBP question would be skipped by the application.

The mock interviews were conducted by a set of participants interviewing each other and recording the answers on a CRF or one of the two mobile phone applications provided. This process is described in Section 3.1.3 below.

The CRFs were returned for central data capture, while the results of the mobile phone interviews were uploaded to a central website. All answers were downloaded from the website and stored in a Microsoft Access database and then analysed using the STATA statistical package.

The accuracy was measured by comparing the recorded/captured result to the known expected answers which had been previously generated in a random manner.

3.1.1 Sample Size Calculation

The error rate of each method of recording was assumed to be normally distributed and a sample size calculation for a test of proportions was used (Fox, Hunn & Mathers 2007).

Assuming an error rate of 5% (Walther *et al.* 2011), in order to obtain 80% power to detect a significant difference at the 3% level between the three methods, 1 063 questions were required for each method.

This number was adjusted to 90 interviews of 36 questions each, providing 3 420 questions in total.

3.1.2 Question Types and the Case Report Form (CRF)

The English and Afrikaans versions of the CRF (Appendix A) were drawn up by Dr Hennie Geldenhuys, a senior researcher at the SATVI site, as a combination of questions that the CRWs or research nurses could be expected to complete in fulfilling their normal duties. This familiarity was expected to reduce the learning load and eliminate the possibility that the participants would not understand the questions, which could introduce an error. The types of questions identified that were felt to be important in the measurement of accuracy were:

- **Simple selection**: Where the user makes a choice between two or more options, e.g.

  Blood sample taken for:
  - Safety bloods and immunology
  - Safety bloods only
  - Immunology only
  - No sample
- **Selection causing skips**: This is where the user makes a choice which modifies the selection of the following question, e.g. if "No sample" was selected for the above questions, the user would be prompted to select the reason for this:

  Reason not taken:
  - Misbleed
  - Participant refused
  - Not required
  - Lab closed

  Any other selection (answer) would skip the question about the reason for not taking the sample.
- **Selection from a large data set**: This is similar to the simple selection above, but where the number of options is too large to scroll through as a single list, e.g. "Residential location" in the CRF has 112 options from which to choose one.

- **Continuous variables**: These are variables from a continuous range of values. Values may be bounded by a lower and upper limit, e.g. a temperature of 36.9°C (36.5 >= t >= 38.1).
- **Continuous variables with special formatting**: These are variables from a continuous range which are captured with special formatting, e.g. a date as 2011/03/05 or blood pressure as 120/90. Values may be bounded by a lower and upper limit, as above.
- **Repeating groups**: This is a group of questions where the group may be repeated more than once. The questions in the group are of the types identified above. For example, "Medical condition" consists of zero or more repetitions of data relating to each symptom:

Sign or symptom	Ongoing?	Date of diagnosis	Date of resolution

3.1.3 Mock Interview

Mock interviews were conducted by pairs of participants, one in the role of interviewer and the other in the role of interviewee. A matrix (a Graeco-Latin square – see Sections 3.3.1 and 3.3.2) defined the interviewer, the interviewee, the interview language and the method of data collection for each mock interview.

An answer sheet with randomly generated answers (Section 3.3.3) was provided to the interviewee for each mock interview. The answers were read by the interviewee from this sheet, in response to the questions by the interviewer, who then captured the answers on the phone or recorded them on the CRF.

3.1.4 Open Source Application

The three mobile applications for surveys and clinical research that were briefly reviewed (Section 2.3) all displayed behaviour which could possibly affect the ease of use, and hence have an impact on the accuracy with which data could be recorded.

It should be noted that all three applications used the basic J2ME platform and widget set as implemented on the various feature phones, which might lead to usability shortcomings.

There was no obvious advantage of any of the applications considered for this study. Therefore, openXdata was chosen for use in this study, primarily because of the support available in Cape Town provided by Cell-Life, who are key developers in its community of users organisations.

3.1.5 Custom Application

As there is potential for user interface problems with the open source application to capture data (Section 2), a custom user interface was designed taking into account the factors affecting usability of the feature phone by the target population (Section 2.3).

3.1.6 Measures of Accuracy

The values captured were compared to the generated answers and classified into one of three categories:

1. OK – the values were the same, taking into account formatting differences between the three options. For example, dates in the custom application were represented as a text string "2011/05/27", on the paper CRF as "27 May 2011" and in openXdata as a date value of "27/05/2011";
2. MISSING – the value had not been captured; and
3. DIFF – the values captured were essentially different (excluding formatting differences).

For values that were recorded as free text on the paper CRF or on openXdata, the level of correctness can be difficult to assess. In this case, the comparison was done by removing all spaces and punctuation and converting to uppercase before automated comparison. Where there was a difference between the English and Afrikaans versions of the text, the comparison was done against each language in turn. After the automated comparison was completed, all remaining errors were inspected to see if it was possible to deduce the correct answer (e.g. in the case of a misspelt word) and the saved value was adjusted.

3.1.7 Qualitative Measures

A simple end-of-study survey of 15 questions on a semantic differential scale of 1 (bad) to 5 (good) was conducted to measure qualitative factors regarding the ease of use of the two applications. The questions covered the overall ease of capturing data on the mobile phone handset using either application, as well as the ease of capturing data for each of the data types. (See Appendix B for the survey form.)

3.1.8 Participant Attributes

The following possible participant attributes that might have an impact on the accuracy of data collection were identified and allowed for in the study design.

- **Level of education**: Would a higher level of education tend to enhance the accuracy of data collection? The level of education of participants was measured in terms of secondary qualifications. Further training, such as courses in GCP, was not considered, as all trained health workers should obtain these qualifications.
- **Computer literacy**: Would a participant who is computer literate make fewer mistakes than one who is not? Participants were considered computer literate if they answered 'Yes' to any of the following three questions:
    - Do you use a PC at work?
    - Do you surf the Internet at home?
    - Do you use email?
- **Home language interview**: If the interviewer and/or the patient were not sufficiently fluent in the language in which the questions were asked, more mistakes might be made. If the interviewer and patient both spoke the same language at home and the interview form was in the same language, the interview was considered to be conducted in the home language.

The purpose of collecting this information was to show that there was no bias in the selection of participants which may influence possible generalisation of the results.

3.1.9 Confounding

It was considered that there may be a causal relationship between low education levels and computer literacy. However, for the purposes of this study, such a causal relationship is not relevant to answering the question regarding the participants' ability to capture data accurately. Therefore, a relationship between these two aspects was not tested.

3.1.10 Ethics and other approval

Ethics approval was received from the Science Faculty Ethics in Research Committee at the University of Cape Town (UCT) on 20 June 2012, reference SFREC 020_2012.
Permission to Access UCT Staff was received from the Executive Director, Human Resources, on 26 June 2012.

3.1.11 Data collection

The initial plan was for the participants to perform the mock interviews when they had a break in their work schedule and that results would be uploaded as soon as possible after the interview, depending on the availability of a network connection. However, this plan was difficult to manage as the participants' movements depended on their own work needs and they could not often co-ordinate their schedules to accommodate the interviews. There was also a very strong drive to finalise data for a major clinical research study in the third quarter of 2012 and therefore the time for this study could not be spared.

Accordingly, the participants were requested to meet on a Saturday morning to perform the mock interviews with each other. This activity was successfully completed on 16 March 2013. The participants were compensated for participating in this project on a Saturday morning.

3.1.12 Analysis

The study was designed to evaluate whether capturing data on a cellphone was at least as accurate as recording the same data on paper CRFs. The primary variable of interest was the error rate for answers recorded on the traditional paper CRF vs. the error rate for answers captured into either of the two phone applications. Responses that were expected, but missing, were classified as an error.

Simple descriptive statistics were used to characterise the study results. For the collected responses, the variable of interest (error rate) was summarised as proportions with 95% confidence intervals.

Although this document uses the error rate as the variable of interest, the analysis was performed using the 'proportion correct'. This change was made as it is easier to visualise the terms 'inferior', 'equivalent', 'not inferior' and 'superior' as being an increase in the proportion correct, thus making it easier to adjust for confidence limits and equivalence ranges and verify the correctness of the calculations. The initial estimate for the proportion correct was 95%, which is equivalent to an error rate of 5% (Walther *et al.* 2011).

Initially, it was planned to follow the model for analysis used by Hawkridge *et al.* (2008) in an equivalence study where the analysis compared the outcome at the 95% confidence level for the new procedure to the outcome for the current procedure, plus or minus an allowable equivalence range. Figure 7 shows how the new procedure would be judged equivalent to the old.

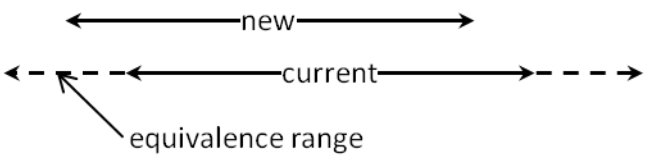

Figure 7 Equivalence limits

However, subsequent research showed that 'non-inferiority' testing could be more suitable. According to Allen and Seaman (2007) "[t]esting for the non-inferiority of a new product compared with an existing approved competitor has become an acceptable practice".

Non-inferiority testing requires the up-front identification of an allowable tolerance between the two methods known as the non-inferiority margin. The non-inferiority margin was calculated as 20% (Allen & Seaman 2007) of the expected error rate of 5%, giving a non-inferiority margin of 1%.

For an existing product, X, and a new product, Y, the hypothesis can be stated as follows:

$$H_o: X - Y > \Delta, \textit{ in which } \Delta \textit{ strictly} > 0$$

$$H_a: X - Y \leq \Delta$$

where Δ represents the non-inferiority margin.

The above null hypothesis states that the new product, Y, is inferior to the current product, X, while the alternate hypothesis states that Y is equal to or superior to X.

For this study, the null hypothesis was that the proportion correct for the mobile phone would be less than that for the paper CRF method. The alternate hypothesis was that the proportion correct for the mobile phone would be greater than or equal to that for the paper CRF method.

To test for non-inferiority, the confidence intervals for the difference between two products are constructed. The null hypothesis that product Y is inferior cannot be rejected if the lower limit of the confidence interval is less than -Δ. On the other hand, non-inferiority (and superiority) is shown when the lower limit of the confidence interval of the mean difference is greater than -Δ bounds (Figure 8).

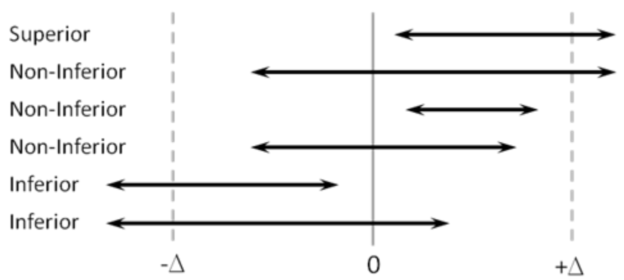

Figure 8 Non-inferiority limits

As non-inferiority is a one-sided test, the equivalence limit was calculated as two times the 95% confidence interval, i.e. equal to the lower limit for a two-sided 90% confidence interval.

The secondary analysis would test various subsets of the data for independence. As it was possible that in some instances one or more samples in the subset could have relatively few observations, the Chi-square Test of Independence (χ^2) was used to evaluate these subsets. If, however, any of the samples had fewer than five observations, the Fisher's Exact Test of Independence was used to determine independence.

3.2 Study Participants

The study participants were all volunteer CRWs based at the Worcester site of the SATVI,[2] the largest dedicated TB vaccine research group on the African continent. The CRWs are responsible for the recruitment of, follow-up on and retention of clinical trial participants and go out into the field each day to conduct interviews, recording the answers on paper forms. They do not capture the data on mobile devices or on a computer back at the office, although some may use email for work purposes.

A preliminary workshop was held with interested participants where the aims of the study and the proposed methods were described and questions were answered. In order to avoid possible bias caused by lack of familiarity with the phone handset, only those volunteers who actually owned a generic feature phone or had recently 'upgraded' to a Blackberry phone were considered for the study. After this workshop, volunteers put their names forward to the operational manager of the site. Fortuitously, only 10 staff members volunteered and all were selected for the study, thus avoiding possible selection bias towards more or less computer literate users.

[2] SATVI's website can be visited at http://www.satvi.uct.ac.za/.

3.2.1 Training of Participants

Training material was prepared covering:

- The CRF (paper form) and the expected questions;
- Downloading of the applications to the mobile phone;
- Using the custom application; and
- Using openXdata.

Training was done in a workshop environment: first as a pilot with two participants, followed by two workshops with four participants each. Time was allocated as follows:

- 30 minutes for explaining the CRF and allowing each participant to conduct a practice interview;
- 15 minutes for downloading both applications (MyApp and openXdata) to the users' phones;
- One hour for teaching how to use the custom application, including a practice interview by each participant; and
- Two and a half hours for teaching how to use openXdata, including a practice interview by each participant[3].

3.3 Randomisation

3.3.1 Participants (interview pairs)

The participants were allocated identification numbers by the SATVI manager and these were captured in a Microsoft Access database. A Graeco-Latin square was created for each method of capture (3x3) and then replicated nine times to create a matrix of nine interviews for each of the 10 participants. This construction resulted in each interviewer conducting three interviews with each method and interviewing each interviewee only once, thus avoiding bias in the allocation of interview methods to the interview pairs.

[3] The recommended training period is one and a half to two hours for 10 people (Workman [n.d.]).

3.3.2 Interview language

The allocation of interview language selection was not in itself randomised, as strict randomisation could have led to too few 'home language' interviews. Instead, the home language of the two participants in each interview pair was compared; if the same, the interview language was set to that common language, until the necessary number of home language interviews had been reached. All other interviews were set to the English language, being the common language of all participants. While a number of Xhosa speaking staff said they understood all three languages, Afrikaans speaking staff did not feel they were sufficiently fluent in Xhosa.

3.3.3 Generated answers

Interview answer sheets were created with randomly generated answers for all the interview questions using the Microsoft Access random number generator. No check was done on preceding answers, e.g. an answer was generated for the question "Specify the adverse event", even if the answer to the preceding question was that there were no adverse events. This was done to avoid that the interviewer would receive a hint that the previous answer had been incorrectly captured.

Having defined the requirements for this study and the methods to be used, the next chapters will cover the requirements for the applications that were used in the study, namely the custom application, MyApp, and the open source application, openXdata.

4. REQUIREMENTS FOR THE CUSTOM APPLICATION

This section covers the requirements for the custom application developed in this study.

- Section 4.1 defines the business transaction(s) required to cover the scope of this research, i.e. capturing data for clinical research.
- Section 4.2 then analyses and models the process(es), tasks and actions required for capturing data, taking into account the usability guidelines and rules identified in Section 2.4.
- The user interface is designed by creating a paper prototype which is covered in Section 4.3. As there are numerous formats to express dates, some of which can confuse the less skilled user, Section 4.3.4 identifies the best format to capture dates when using the custom application. The use of predictive text in the context of data capture is defined in Section4.3.5.

The definition of the requirements followed a standard waterfall system development lifecycle with iterations where required. For example, if the user interface design was not optimal, the process definition could be restarted. It was felt that using an agile lifecycle with its inherent incremental change approach would not cater sufficiently for the required analysis and reflection.

As it was important to create an application which was simple to use and enabled the user to complete the desired task as efficiently as possible, a Task Analysis approach (Diaper & Stanton 2004) was used so as to optimise the tasks required, rather than an approach that explored all possible functionalities.

The steps in this approach are:

1. 'Business Transaction Discovery', which identifies the end-to-end business transaction(s) required for the defined scope of work.
2. 'Business Process Definition', which defines each transaction in detail identifying the individual tasks or logical units of work, the flow between the tasks and any exception conditions. Included in this definition are the actions required to be executed during the performance of each task.

3. 'User Interface Design', which uses an interactive paper prototyping method to lay out the user interface and to confirm that it is practical.
4. 'Application Design', where the technical architecture, data types and structures and flow of the application are defined, taking into account the preceding definitions and designs. ('Application Design' is covered in Chapter 5.)

As the aim of this study was to test the ability of a cellphone to capture data, various tasks that would be required for a commercial application were identified, but then omitted from the detailed definition.

4.1 Business Transaction Discovery

This activity identified the only business transaction of interest as 'Capture Subject Data' where the data for one specific subject would be recorded and saved on a central server for later analysis.

4.2 Business Process Definition

A typical business transaction is completed by following a defined process which is made up of one or more tasks. A task is a single unit of work made up of one or more steps. A task can be interrupted, provided the intermediate state of the task is saved. Actions are performed by the user to complete a step.

The Business Process for 'Capture Subject Data' is a simple process consisting of three tasks once the user has logged in to the application as shown in Figure 9.

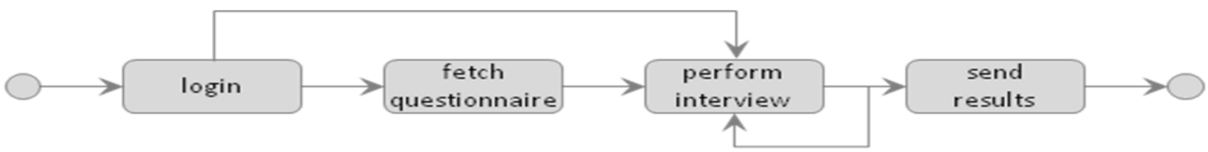

Figure 9 Process diagram

4.2.1.1 Login

Log in locally on the handset. Connection to the network is not required, unless send or receive tasks are performed.

4.2.1.2 Fetch Questionnaire

Fetch a new questionnaire or new version of a questionnaire and save it on the handset. Steps are:

1. Display questionnaires currently saved on the handset.
2. Prompt to confirm that questionnaires are to be fetched from the server.
3. Connect to server and download questionnaires, saving each to the handset storage.
4. Confirm when operation is complete.

The task can be interrupted and abandoned before the questionnaires are saved.

4.2.1.3 Perform Interview

Select a questionnaire that has been saved previously on the handset and perform the interview by asking the questions in the questionnaire and recording the answers. The interview answers are automatically saved as a single set of results on the handset upon completion of the interview. The interview can be paused, the answer to the previous question erased and the previous question displayed if the interviewer realises the preceding answer was incorrect.

4.2.1.4 Send Results

Send the saved results to the central server. Results need not be sent upon completion of the questionnaire, as it may be inconvenient or the network connection may be too poor or non-existent. Steps are:

1. Display list of saved results.
2. Prompt to select the result to send.
3. Connect to server and upload result.
4. Confirm when operation is complete.

The task can be interrupted and abandoned before the result is sent.

Additional tasks that could be included in a commercial application include:

- 'Review Answers'; and
- 'Delete Saved Result'.

The design of the application must allow for these tasks to be included without compromising the user experience, otherwise this research would be of no value.

4.2.2 Design Guidelines

As Nielsen notes, "Feature phones [are] primitive handsets with tiny screens and very limited keypads that are suited mainly for dialling phone numbers." (Nielsen 2011a). Two key aspects of usability that are limited by these 'primitive' handsets are multitasking/task switching and action selection (telling the device what to do). Guidelines for implementing these on the feature phone are provided below.

4.2.2.1 Multitasking/task switching

The phone display does not provide space for multiple windows, nor for a task bar to display the various running process/applications. So, although it is technically possible to multitask the business process, the user will have to remember each active task and its state. This additional memory load makes the user's job far more complex and so, in this instance, the user will be limited to performing a single task at a time, i.e. the design will be procedural, rather than event driven. This limitation does conflict with Shneiderman's rule relating to the locus of control, but as noted above (Section 2.4.2) it does accord with Ballard's guideline of efficiency as it enables the user to focus on a single task at a time.

4.2.2.2 Action selection

In order for users to interact with the feature phone, they need to be able to enter commands to the device to perform an action.

A feature phone has the following restrictions:

- The display is too small to allow space for a menu bar, which is extensively used in WIMP (windows, icons, menus, pointer) interfaces. Thus, drop-down (or pull-up) menus cannot be provided for action selection as is commonly done on WIMP interfaces;
- Chorded key press actions (pressing multiple keys simultaneously) are difficult on the cramped keypad, and are therefore not considered as a viable option; and
- The keypad has two or usually three soft keys at the top of the keypad together with four directional keys. As the directional keys are used for moving the cursor, the number of actions that can be selected by pressing a soft key is limited.

With the limited keypad, the only realistic command mechanisms that can be used are the:

- Function key: Press a soft key to perform the identified action – this action can occur at any time;
- Menu form: Select an option by using the directional keys to navigate a list and then press a soft key to select the desired option. Sometimes the numeric keys can be used to select an option number; and
- Context menu: Same as above. The context menu is displayed when a soft key is pressed. This mechanism is not always available.

In order to implement a usable application while addressing the issues around action selection, the available actions were classified and prioritised and the implementation of each class of action was defined to be dissimilar.

4.2.2.2.1 Classification of actions

The actions were classified as follows:

- Task initiation actions, which initiate a business task, e.g. start task 'Perform Interview';
- Process actions, which affect the process state of a task; e.g. 'Select' picks the highlighted answer, saves it and displays the next question. Process actions can be sub-classed into 'forward', 'backward' or 'holding', based on whether they change the process state toward completion, away from completion, or render no change.
- Display actions, which aid the user in completing the answer, e.g. the left directional key deletes the preceding character in a text box.

This classification is consistent with Shneiderman's rules, as the separation supports predictability and reduction of short-term memory load.

This differentiation is commonly implemented by initiating tasks from a standard menu structure, while triggering desired actions through command buttons or function keys.

When the menu structure defines the complete business process through suitable ordering and nesting, the user obtains an understanding of what the tasks are that (s)he has to perform. Menu items to which the user has access are distinguished in one way or another, thus providing even more information. This display of the business process supports goal-seeking behaviour by providing a (partial) model of the system (Dix *et al.* 2004).

As the display actions require navigation or cursor movement, they are triggered by the central directional keys.

Process actions are triggered by left and right soft keys and as there are only two soft keys, the process actions are limited to two per form. The forward process action should be implemented on the right soft key and the backward process action on the left soft key. This is the natural order for users who use left-to-right text formalisms to represent language.

4.2.2.2.2 Prioritisation of actions

As the aim of the design should be to complete a business task efficiently (Section 2.4.3), 'forward' actions are desired and should be prioritised over and above less desired 'backward' or 'holding' actions. The exception to this prioritisation is an action to undo the previous action, which should be as easy to trigger as possible. Examples of this prioritisation are:

- 'Move to Next Question' would normally be the desired action. This action should be easy to trigger;
- 'Back to Previous Question' would undo the previously saved answer. This action too should be easy to trigger.
- 'Pause Interview' is a holding action, which is not usually desired. If necessary, this action can be defined in a manner that is less easy to trigger, or it may be omitted if there is another alternative. It was omitted from this application, as it would be unlikely that a health related interview would be interrupted except in an emergency.
- 'Cancel Interview' is not a desired action. It too may be designed to be less easy to trigger. It was catered for in this application by allowing the user to repeatedly undo previous answers using the 'Back to Previous Question'

action until the first question was reached, and then undo the first answer to cancel the interview.

4.3 User Interface Design

The primary aim in designing the user interface was ensuring 'consistency with simplicity' and was done in order to enhance learnability as defined by Dix (Dix *et al.* 2004), and thereby minimise errors.

4.3.1 Initial design goals

The initial design goals were the following:

- **Provide a single consistent home form**. The home form should present the menu for the overall process, thus providing the user with a recognisable 'starting point' for all tasks to comply with Ballard's predictability guideline.
- **Minimise user decisions during any task by eliminating less desired options**. This curtailment implies a reduction in flexibility, but is thought to be beneficial as it allows the user to focus on the task at hand. This goal is an application of Shneiderman's rule to 'reduce short-term memory load'.
- **Classify and prioritise the actions** (see Sections 4.2.2.2.1 and 4.2.2.2.2). This goal complies with Shneiderman's rule for consistency by grouping similar items in a class. It also complies with Ballard's efficiency guideline by allowing easier triggering of more frequently used actions.
- **Use consistent display objects** for
    - Questions;
    - Answers (where entered by the user);
    - Option lists; and
    - Progress indicators for long-running tasks, e.g. fetching the questionnaires.

  This goal complies with the Shneiderman's rule of consistency.
- **Lay out the various forms in a consistent manner**.
    - Place questions at the top of the various forms used for the interview;
    - Place the answer space below the question; and
    - Place lists of options below the answer space if present, otherwise below the question.

This goal complies with the Shneiderman's rule of consistency.

- **Show progress indicators with text indicating process states.** For example, connecting > registering > downloading > complete as a sequence of states for fetching questionnaires. This goal supports Shneiderman's rule of providing informative feedback.
- **Avoid multi-tap keying**. Where necessary, use predictive text for textual answers or make use of the * (asterisk) and # (hash) keys by re-mapping them to symbols as required. This goal complies with Ballard's efficiency guideline.

4.3.2 Action definitions

In accordance with the prioritisation of actions above, soft key actions will be limited to:

- Back action ('Back', 'Cancel') – left-hand soft key;
- Forward action ('Select', 'OK') – right-hand soft key;
- Cursor up/down – up and down directional keys;
- Mark (multi-select;) – centre soft key; and
- Backspace (text correction) – left directional key.

Ideally there should be three 'back'-type actions:

- 'Back to Previous Form' – the 'Back' action;
- 'Backspace Deleting the Character before the Cursor' – 'Backspace' action; and
- 'Cursor Left' – traversing the text without deletion.

Given the limited number of function keys, 'Cursor Left' was omitted in the application developed in this study. The workaround for users would be to backspace and re-type if they needed to correct an intermediate character in the character string. As this action is a non-desired action, this restriction would probably be acceptable to the users.

4.3.3 Initial designs

4.3.3.1 Home form and 'Main Menu'

The application should have a single consistent home form, to which the user is returned after completion of each task. The ideal form would be a 'Main Menu' from which the user can select the task (s)he wishes to initiate. The initial design for the menu is an unnumbered list of tasks plus 'Exit' as shown in Figure 10.

An optional display item could be the number of saved results, which could serve as a warning to the user. This item was not included in the initial design due to a concern that it would then take up some of the scant space from the 'Main Menu' itself.

Figure 10 Main Menu

Figure 11 List of Questionnaires

4.3.3.2 'List of Questionnaires' or 'List of Results'

The 'List of Questionnaires' or 'List of Results' would be similar to the menu, apart from allowing multiple lines for each option as shown in Figure 11. The multiple lines are required to unambiguously identify the required questionnaire/result by including the language and version of the questionnaire or the interview number and date of the result.

4.3.3.3 Questions and answers

The initial design for a numeric answer, a date or a similar answer that is keyed in by the user, is straightforward. The design for both single-select and multi-select questions are also straightforward. The designs are shown below in Figure 12,

Figure 13, and Figure 14. For a single-select question with predictive text, the design can be visualised as a combination of a keyed-in answer and a dynamic list of options, as shown in Figure 15.

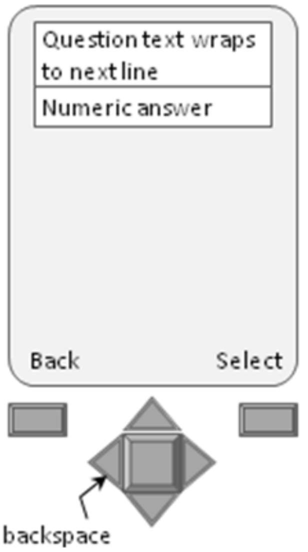

Figure 12 Numeric answer

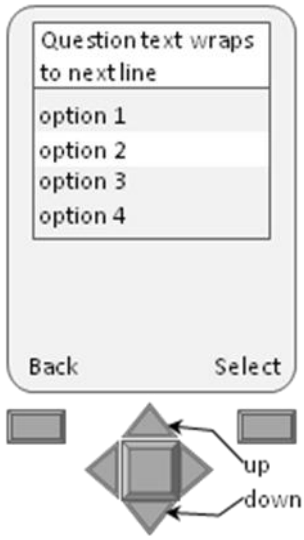

Figure 13 Single-select

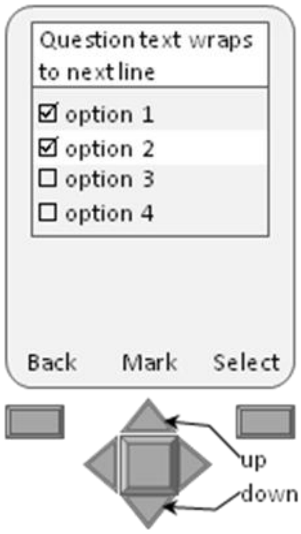

Figure 14 Multi-select;

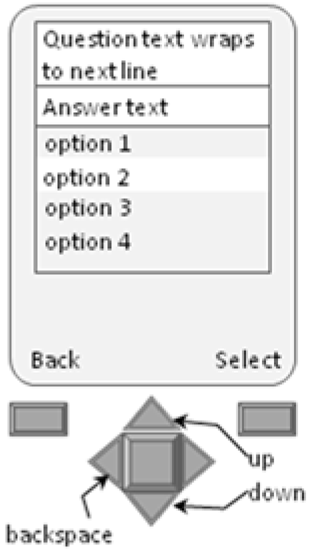

Figure 15 Predictive text

For numeric and similar answers, the * and # keys on the keypad are re-assigned to . (decimal point) and / (slash). This re-assignment allows users to capture real numbers (e.g. 37.7), dates (e.g. 2011/05/09) or values such as blood pressure (e.g. 70/120). Times will be captured as a four digit number, such as 1200 to represent midday.

4.3.3.4 'Progress Form'

The long-running task form displays the task name, a progress bar and the current task state as shown in Figure 16. The 'OK' soft key action returns the user to the previous form. The 'Cancel' soft key action pops a confirmatory dialog, and if the OK option is selected, the long-running task is stopped and the application returns to the preceding form.

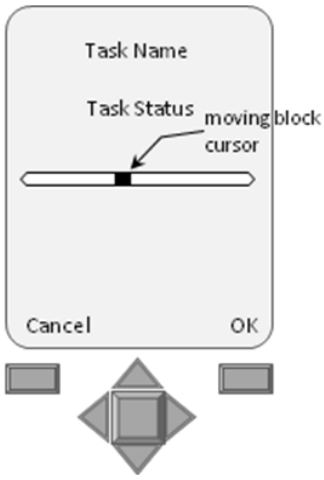

Figure 16 Progress form

The user interface as defined above was implemented in the custom application.

4.3.4 Date Formats

To simplify the capture of dates using the 12-key keypad, one option is to use a fully numeric date format as it is difficult to capture the month as a word. Alternatively, predictive text could be considered to capture the month as a word.

The familiar formats used in South Africa are:

1. yyyy/mm/dd – four-digit year followed by two-digit month and two-digit day, e.g. 2013/07/11. This is the International Organisation for Standardisation (ISO) standard for writing dates (ISO 2004), to which South Africa is a signatory;
2. dd/mm/yy – two-digit day followed by two-digit month and two-digit year, e.g. 11/07/13;
3. yy/mm/dd – two-digit year followed by two digit month and two-digit day, e.g. 13/07/11;
4. dd/mm/yyyy – two-digit day followed by two digit month and four-digit year, e.g. 11/07/2013;
5. dd mmm yyyy – two-digit day followed by three-character abbreviated name of the month and four-digit year, e.g. 11 Aug 2013; and
6. dd mmmm yyyy – two-digit day followed by month's name in full and four-digit year, e.g. 11 August 2013.

Options 2 and 3 were discarded as they can be confused with each other. As shown in the examples, it is not clear whether the day is the 13th or the 11th and whether the year is 2011 or 2013.

Option 5, using predictive text, was discarded as 'Feb' and 'Dec' both correspond to the key sequence 3, 3, 2

Option 6 was discarded as being cumbersome, as even using predictive text would still require four key presses to disambiguate February and December.

Finally, Option 1 was chosen as being the national and international standard date format. (In hindsight, this decision was not the best choice as the CRWs have been extensively trained to record dates in the format of Option 5 (dd mmm yyyy) and reversing the order of year and day confused them – see Section 7.2.3)

4.3.5 Predictive text specification

4.3.5.1 Form

The predictive text form is built with a text area for the question at the top of the form, the answer text box just below the question and the list of options below the answer, as shown in Figure 17

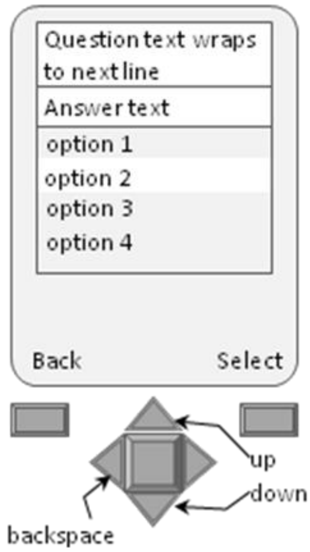

Figure 17 Predictive text

4.3.5.2 Predictive data list

For each question identified as validated text, the phone memory contains lists (vectors) for the following allowable options:

- The text to be displayed: e.g. 'Option 1';
- The numeric key mapping for this text: e.g. '67846601' ('Option 1' where 0 maps to space);
- Optional shortcut mapping: e.g. 6#1 where 6# could be equated to 'Option 1'; and
- The return code for the option: e.g. 'Ans1'.

The return code for each option is consistent across the various languages the questionnaires are presented in, even if the order of options changes. For example, the English questionnaire may contain the options 'arm', 'chest', 'leg' … (in alphabetical order), while the Afrikaans equivalent contains 'arm', 'been', 'bors' … (also in alphabetical order). The answer code mapping for 'chest' and 'bors' is 'Ans2' and the mapping for 'leg' and 'been' is 'Ans3' as they represent the same body part in the different languages.

4.3.5.3 Variables

The variables used to explain the functionality of the interface are as follows:

- inputString – a string containing the sequence of keys pressed;
- numChar – the current number of characters to match and display; and
- selectedOption – the currently selected option, initialised to the optionText of the first position in the vector.

- optionVector – a vector of the list of options.
- numericMapping – a vector of the mapping between the numeric code and the vector of options.
- shortcutMapping - a vector of the mapping between the numeric code including shoprtcuts and the vector of options.

  Example of the mapping:

  optionVector item: pneumonia

  numericMapping: 763866642

  shortcutMapping: 7#642

  where the shortcut '7#' or 'P#' represents the string 'pneumo'

4.3.5.4 Functionality

The following functionality will be used to implement predictive text:

1. Display the question in the protected text area and the list of options for the question in the list. Position the cursor in the answer text box. Highlight the first item in the option list. Initialize numChar = 0.
2. Trap all key presses; i.e. do not display the pressed key in the answer box.
3. When a non-directional key is pressed, perform the following actions:
   a. Add the key value to inputString and increment numChar.
   b. Select all options from the optionVector where either the inputString maps to the first numChar characters of the numericMapping or the inputString maps to the first numChar characters of the shortcutMapping.
   c. Build a new list from the selected options and replace currently displayed list.
   d. If the current selectedOption is no longer in the new list, reset selectedOption to the first value in the list.
   e. Display substring of first numChars of selectedOption in answer text box.

   Example: Assume the options consist of 'arm', 'chest', 'leg'… and the '2' key (abc) is pressed. The numChar will be set to 1, the list will be reduced to 'arm', 'chest' and the answer will display 'a'. If the '4' key (ghi) is then pressed, numChar will be incremented to 2, the list will be truncated to 'chest' and the answer will display 'ch'.
4. If an up or down directional key is pressed, perform the following actions:

a. Identify the previous (up) or next (down) option in the currently displayed list based on the current value of selectedOption and set selectedOption to this value.
b. Remove the previous highlight and highlight this option in the list displayed.
c. Display substring of first numChars of selectedOption in answer text box.
d. Ignore the key press if there is no previous or next option (top or bottom of list) as it will probably cause confusion if handled as a circular list

Example: Assume the options consist of 'arm', 'chest', 'leg' … with the 'arm' option highlighted. If the '2' key (abc) is pressed, numChar will be (re)set to 1, the list will be reduced to 'arm', 'chest', the 'arm' option will remain highlighted and the answer will display 'a'. If the down key is then pressed, the option 'chest' will be highlighted and 'c' will be displayed in the answer text box.

5. If the left direction key is pressed (this action is treated as a backspace), perform the following actions:
    a. If numChar $\neq$ 0, decrement numChar and remove the last character from inputString.
    b. Perform steps b to e of 3 above.
6. If the 'Select' key is pressed, perform the following actions:
    a. If the inputString does not match the currentOption (i.e. only part of the key has been entered), display the currentOption in full in the answer text box, set numChars to the length of the currentOption and set inputString = currentOption. This set of actions allows the user to verify if they have identified the correct item.
    b. If the inputString matches the currentOption, accept the option and return the returnCode for the option to the calling application.

5. MYAPP CUSTOM APPLICATION

The custom application, MyApp, was created specifically for this study using CodenameOne. It is an Internet-based application and saves the captured answers to a MySQL database on the website, as shown in Figure 18.

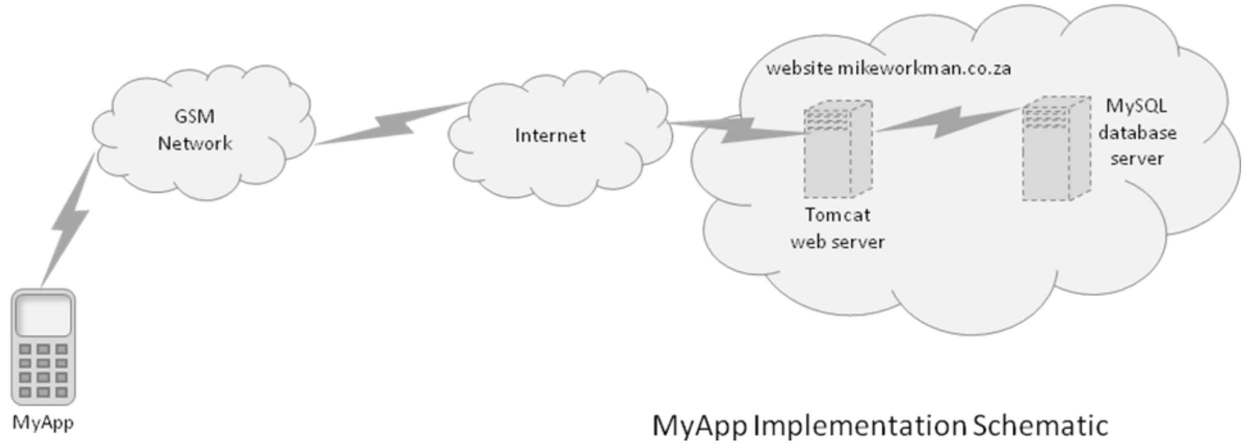

Figure 18 MyApp implementation schematic

MyApp provides the simple functionality to download a questionnaire, perform an interview and to send each result as a delimited string to the website. The usual server-side services required for a field survey, such as user management and access control, loading data into a study database (CRDB), statistical validation of the data, analysis tools, etc. are not available.

As the target platform for this study is the feature phone, J2ME was the obvious choice to use for implementing the custom application. However, it was found during the development of the custom application that the J2ME implementation inhibited the creation of the predictive text feature (see Section 4.3.4). Further, it was found that the different manufacturers implement the soft key functionality in different ways, which prevented the proposed minimalistic design (see Section 2.3).

It was therefore not possible to achieve the desired user interface with J2ME, or to create it in a consistent manner across manufacturers. Hence, alternatives were sought and CodenameOne was selected as it would allow the desired user interface to be created (Section 2.2.1).

5.1 Internal structures

The questionnaires and results are saved in the handset's internal store as delimited strings with multi-level delimiters. This structure was chosen as it is easy to parse, and it was felt to be a compact structure that could be more consistently transmitted over low-speed networks (Section 1.5.1) than a more verbose structure such as XML. This compactness was verified with the size of the saved results for the study CRF (English version) being less than 400 bytes.

Each questionnaire and each result (all answers for a single interview) is saved as a separate record in the store. For performance reasons, a questionnaire index and a result index are saved as vectors, each in its own record in the store.

5.2 Questionnaire definition

The entire questionnaire is assembled as a single, multi-level delimited string. The constituent parts of the string are:

- Question Identifiers:
    - ID;
    - Description;
    - Version; and
    - Language.
- Question Number: A three-character question number. The question number is not displayed and is not used for ordering the questions during the interview, apart from the interview always starting at question 001. Order is only controlled by the 'Next Question' parameter (see below).
- Question Text: The text as it is to be displayed on the handset display.
- Question Type: The type of question:
    - s-s: single select;
    - m-s: multi-select;;
    - vtx: free-form text with validation; and
    - ptx: predictive text.
- Parameters:
    - Question types s-s, m-s and ptx: the options to be displayed; and

- Question type vtx: parameter 1 is the regular expression used to evaluate the answer, parameter 2 is the error message to be displayed if regex validation fails.

- 'Next Question': The 'Question Number' of the next question to be displayed if this option is selected. The specification of this value controls the question flow through the questionnaire. Repeating groups are handled in a loop with the next question number of the last question of the group directed at the question that is at the start of the repeating group.
- The delimiter characters selected are §, ¥ and ¢ and are used to separate questions, parameters and options within parameters. These characters were selected as being unlikely to be used in questions as well as not being reserved within Java or regular expressions.

A partial sample of the questionnaire structure is shown in Table 1.

Table 1 Questionnaire sample

Question number	Question text	Question type	Question parameters	Next question	Delimited string
003	Participant's date of birth	vtx	^(19|20)\d\d/(0[1-9]|1[012])/(0[1-9]|[12][0-9]|3[01])$	004	§003¥Participant Date of Birth¥vtx¥^(19|20)\d\d/(0[1-9]|1[012])/(0[1-9]|[12][0-9]|3[01])$¢004
			Date format is yyyy/mm/dd, e.g. 2012/05/21	004	¢Date format is yyyy/mm/dd e.g. 2012/05/21¢004
004	Consent signed?	s-s	Yes	005	§004¥Consent signed?¥s-s¥Yes¢005
			No	007	¢No¢007
005	Version of ICF signed	vtx	^[1-9]\.\d$	006	§005¥Version of ICF signed¥vtx¥^[1-9]\.\d$¢006
			Version number is e.g. 1.4	006	¢Version number is e.g. 1.4¢006

5.3 Display objects

The display objects (widgets) for a generic mobile handset should be as simple as possible and should be activated by, or receive input from, the keypad with ease. Typical PC display objects which require a pointing device (e.g. screen button) or scroll wheel (e.g. spinner) should be avoided. To this end, the display objects that can be readily used with a 12-key keypad are:

- 'Text Field' – For user input from the keypad displaying up to 255 characters. Input can be limited to certain character combinations (e.g. numeric only). Scrolling behaviour of the text is possible where the text is longer than the display width of the text field, or text can be wrapped to display on multiple lines.
- 'Text Area' – For longer user input from the keypad, typically up to 2 048 displaying multiple lines of text. The text wraps at word boundaries, where possible.
- 'List' – A set of predefined text values (usually) arranged vertically, which the user navigates using the up/down navigation keys on the keypad. Lists can be:
  - Exclusive – Only one item can be selected, the selected item being indicated by a radio button;
  - Implicit – Only one item can be selected, the selected item being highlighted (no radio button is displayed); and
  - Multiple – More than one item can be selected, the selected item(s) being indicated by a checkbox.
- 'Progress Indicator' – A widget which shows that background activity is occurring by a sliding or rotating icon. For this implementation a horizontally sliding device was selected, as it takes up less space on the display and can be implemented on even the simplest graphical display.

5.4 Form Design

5.4.1 Home Form and 'Main Menu'

A home form is important as it provides a consistent user experience. For this application, the home form was the 'Main Menu' as shown in Figure 19, with the various tasks in the order in which they would be performed, viz. 'Fetch Questionnaire', 'Perform Interview', 'Send Results' and 'Exit'. As 'Exit' is a task that may have to do some clean up, it is set as a menu

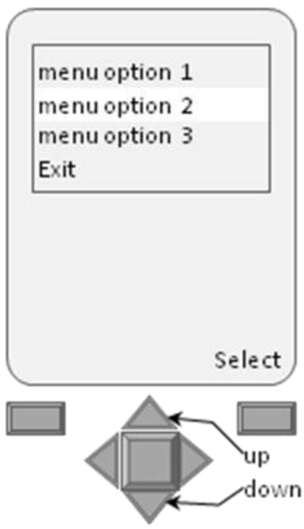

Figure 19 Main Menu

item and not as a task action soft key. Therefore, there is no back soft key.

5.4.2 ‘Fetch Questionnaire’ task

To avoid any possible confusion with out-of-date questionnaires, all current questionnaires are fetched from the server whenever the action 'Fetch Questionnaire' is performed. The questionnaires previously saved on the handset are discarded. The form displays a continuously moving slider to indicate the busy state, and displays the status : connecting > verifying > fetching > done. When complete, a dialog "All questionnaires saved" is displayed. Soft keys are 'Cancel' (left-hand soft key), which aborts the task, and 'OK' (right-hand soft key), which closes the dialog and form and returns to the 'Main Menu' once the task is complete.

5.4.3 'Perform Interview' task

The following forms are provided for the capture of answers to each question in the questionnaire, depending on the type of answer expected. All forms display the question in full in a multi-line text area at the top of the display.

The soft keys defined for all questions are:

- 'Back' – Discard last saved answer and display the last previously shown question (left-hand soft key); and
- 'Select' – Save answer, look up the next question depending on the value of answer just saved and display the question (right-hand soft key).

5.4.3.1 Continuous values

Continuous values (integer or real numbers), such as weight or temperature, were implemented as numeric text fields associated with a regular expression that would restrict the input to certain patterns. The action for the * key on the keypad was over-ridden to create a decimal point for real number values. The text field is displayed immediately under the question as shown in Figure 20.

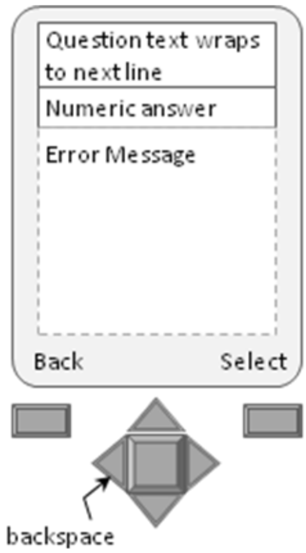

Figure 20
Continuous values

Below the 'Answer' text field, a text area is displayed with an error message if the selected answer were to fail validation according to the regular expression.

The value is accepted when the user presses the right-hand soft key, 'Select'.

The left directional key is defined as 'Backspace', meaning 'move left, deleting the preceding character'. The other directional keys are not used.

5.4.3.2 Continuous values with special formatting

Continuous values with special formatting cover data items such as dates or blood pressure. These types were implemented as continuous values as described above with a regular expression that would restrict the input to a date pattern (yyyy/mm/dd) or blood pressure pattern (nnn/nnn). The action of the # key on the handset keypad was overridden to provide the / separator, which is used for dates and blood pressure. The text field is displayed immediately under the question.

Below the 'Answer' text field, a text area is displayed with an error message if the selected answer fails validation according to the regular expression.

The captured value was accepted when the user presses the right-hand soft key 'Select'.

The left directional key was defined as 'Backspace', meaning 'move left, deleting the preceding character'. The other directional keys are not used.

5.4.3.3 Single selection

An implicit (unordered) list was implemented for the selection of single items from the list. The list was displayed immediately under the question as shown in Figure 21. The list could be navigated using the up and down directional keys.

The highlighted item could be selected either by the centre soft key (unlabelled), or the right-hand soft key, 'Select'. The other directional keys are not used.

5.4.3.4 Multiple selection

A list displaying checkboxes was implemented for the selection of multiple items from the list as shown in Figure 22. The centre soft key could be used to mark the items, which is the usual behaviour for the mobile applications that were reviewed, as well as other instances on the mobile phone. The list could be navigated using the up and down directional keys. The selected items are accepted when the user presses the right-hand soft key, 'Select'.

5.4.3.5 Predictive (coded) selection

The predictive text form (see Figure 23) is laid out as follows:

- The question text area at the top of the display;
- The answer text box below the question; and
- The list of options below the answer.

(See Section 4.3.4 for the detailed specification.)

The list can be navigated using the up and down directional keys, while the left directional key implements the 'Backspace' function.

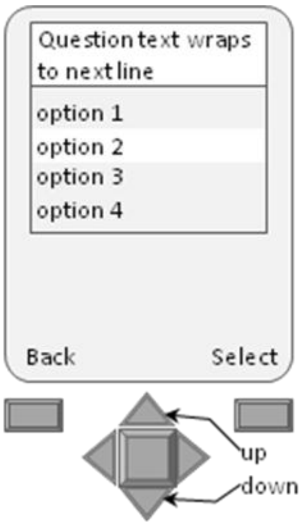

Figure 21 Single-select

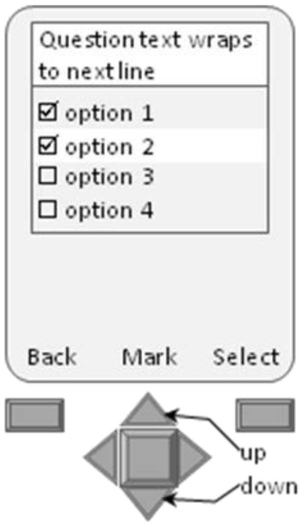

Figure 22 Multi-select

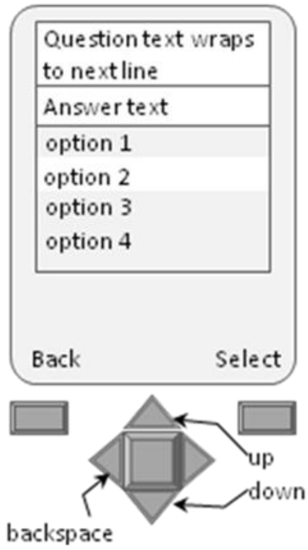

Figure 23 Predictive select

5.4.3.6 Repeating Groups

A repeating group is a group of questions where all questions in the group may be repeated more than once.   For example, e.g. the category 'Medical condition' consists of zero or more repetitions of data relating to each symptom. The four questions here are as follows:

Sign or symptom	Ongoing?	Date of diagnosis	Date of resolution

With a personal or tablet computer, it is usual to present the entire group in tabular format with one row for each repetition and a column per question. However, with a small form factor, such as that of a feature phone, one cannot use the tabular layout which requiress each question to be shown in isolation. The repetition is managed as a looping sequence of questions where the first question (in this case "Sign or symptom?") has an option "none/none further", which causes the question flow to break out of the loop when selected.

The consequence of this approach is that the application will store multiple question/answer pairs for these repeating groups and a server-side application will have to unpack them after the results have been sent.

5.4.4 'Send Results' task

There are two forms for the 'Send Results' task:

- 'Select Result': Results are selected individually. The form displays a list of saved results showing the interview number and date of interview for each result. The user selects the result to send and presses the 'Select' soft key (right-hand). The 'Back' (left-hand) soft key returns to the 'Main Menu'.
- 'Progress Form': This form is similar to the 'Fetch Questionnaire' progress form and is opened upon the selection of the result to send. When the send is complete, the sent result is deleted and the user is returned to the 'Select Result' form. If there is a failure in the sending process (e.g. connection time-out), an error message is displayed on the form. The 'Cancel' (left hand) soft key cancels the 'Send' action and returns to the 'Select Result' form. There is no right-hand soft key function for this form.

5.4.5 'Exit Form' task

The 'Exit Form' task displays a dialog, "Are you sure you want to exit?". The 'OK' (right-hand) soft key closes the application and the 'Cancel' (left-hand) soft key returns one to the 'Main Menu'.

6. OPEN SOURCE APPLICATION

openXdata is an Internet-based application for recording data for various types of survey and research studies. mforms is a Java-based application for feature phones that interfaces with openXdata. Together, they have been used to capture clinical research data in a number of studies.

6.1 Implementation of openXdata and mforms

For this study, openXdata was implemented as follows and as shown in Figure 24.

1. An Amazon Web Services (AWS) Elastic Compute Cloud (EC2) instance was created on the AWS in their eu-west-1c zone hosted in Ireland using the AWS New Instance wizard. Due to the author's familiarity with Windows, a Windows Server 2008 instance was configured rather than the more common Linux instances.
2. The necessary security keys were created and installed using the EC2 Management Console.
3. Although the instance includes an Apache HTTP web-server, openXdata requires the Apache Tomcat web-server. Tomcat was downloaded, installed and started using a remote desktop connection to the EC2. As Apache HTTP was already listening on Port 80, it was configured to forward all calls to Tomcat on port 8080.

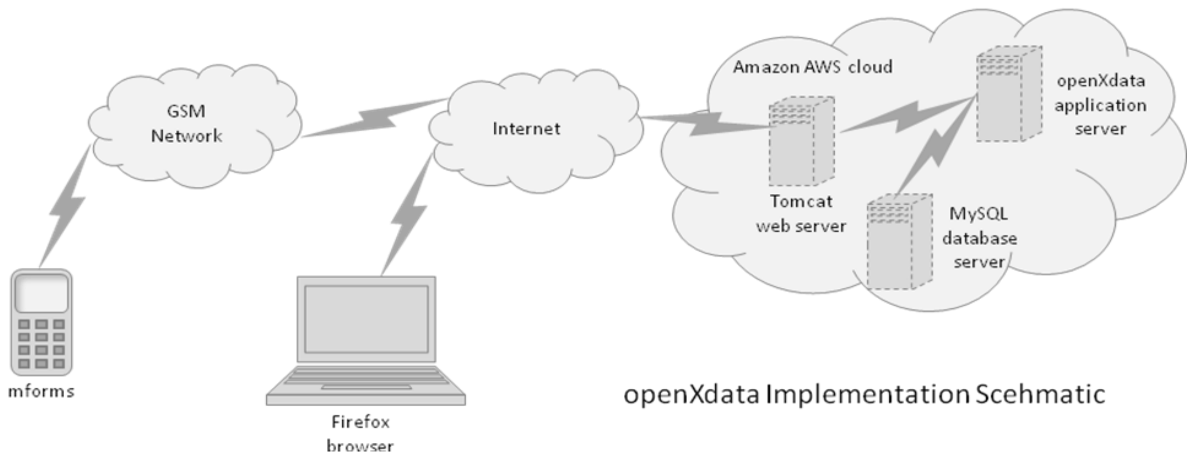

Figure 24 openXdata Implementation Schematic

4. Although the instance includes a Microsoft SQL Server database, openXdata requires a MySQL database. SQL Server was stopped and MySQL was downloaded, installed and started using a remote desktop connection to the EC2.

5. An elastic IP address, or EIP, was registered in order to access the EC2 instance over the Internet via the openXdata forms, or via the mforms mobile application.
6. openXdata was downloaded and installed using a remote desktop connection to the EC2.
7. Tomcat was stopped and restarted and openXdata was deployed from the Tomcat Management Console.

6.2 openXdata form management

openXdata provides a management form from which one can manage study users, create, edit and manage questionnaire forms and download saved questionnaire results in comma-separated variable (csv) format.

Questions are defined with the following elements:

- Question Text.
- Question Type.
    - Number (Integer or Real);
    - Date, Time or DateTime;
    - Boolean;
    - Single Select;
    - Dynamic Single Select (list is dependent on preceding answer in hierarchy of Single Select);
    - Multi-select;; and
    - Repeating Group.
- List options for single or multi-select lists.
- Range checking for numeric values.
- Mandatory/Optional.

By default questions are defined as 'enabled'. In order to control the question flow and skip some questions, a question can be disabled according to the answer captured for one or more preceding questions.

Validation rules can be defined for each variable and the rules can reference the value of a preceding answer.

openXdata also includes a graphical web form designer which allows the form to be laid out according to the requirements and preferences of the user.

6.3 Using openXdata and mforms

The use of openXdata and mforms for capturing data is described below. The application is flexible in that a user can elect to skip questions and still save the results, even though the skipped questions may be mandatory. The same mechanism (soft keys and context menus) are used for initiating tasks and for performing actions during the tasks.

6.3.1 Capturing data with openXdata

When a user logs on to openXdata, (s)he is presented with a list of forms to which (s)he has been assigned and then (s)he can capture the data for the selected questionnaire. On completion of the questionnaire, the data are saved to the database.

6.3.2 Viewing and downloading data

A user with the role of manager may view all captured data and may download the data as a .csv file for subsequent analysis.

6.3.3 mforms

mforms is a Java-based (J2ME) application for feature phones. It has the following functions:

- Logging on locally to the openXdata mforms application;
- Downloading openXdata questionnaires, including dynamic list, validation and skip rules, but ignoring the openXdata form layout. A background logon to the openXdata server is done using the credentials with which the user logged on to mforms;
- Capturing data for the selected questionnaire. This task is performed offline (on the handset), but requires local logon; and
- Uploading saved questionnaires to openXdata. The answers are saved as if they had been captured using the openXdata web forms.

6.4 mforms users, studies and forms

When using openXdata and mforms, a user must always log on. Then mforms validates the user locally against a saved list of users and credentials and only when an action is requested that requires connection to openXdata is the user revalidated against the master user list.

Studies are defined by an administrator and users are linked to studies. Forms for data entry are created for a particular study, although they can be subsequently shared between studies. When the user logs on, (s)he is presented with a list of forms to which (s)he has access. When logging on to mforms, the user may select to fetch new or revised forms from the server, or may continue using the forms currently saved on the handset.

7. RESEARCH RESULTS

7.1 Non-adherence to the Study Design

During the course of the study, it was found that one of the aspects of the original study design had to change. It was impossible to cater for the users'/participants' home language in every instance.

The initial design was to evaluate the impact of home language interviews by allocating Xhosa questionnaires to interviews where both participants were Xhosa speaking; Afrikaans questionnaires to Afrikaans-speaking pairs and English questionnaires to interviews where the pair did not have a common home language. Adjustments would then have been made to conduct approximately half the interviews in a home language and the other half in English.

An Afrikaans CRF was prepared by a SATVI senior research officer (Dr Hennie Geldenhuys) who is Afrikaans speaking, while the Xhosa CRF was prepared by a senior member of the SATVI staff (Ms Tembakazi Daki) who is Xhosa speaking.

However, when the Xhosa CRF was distributed at training, the questions puzzled the Xhosa speaking participants and they did not always understand what was being asked. It transpired that many medical terms do not have equivalent terms in the Xhosa language. The translator, in conjunction with the senior research officer, had created words and phrases which they felt best represented the questions, but these words and phrases were not recognised by the participants, or they were confused by them.

Accordingly, the Xhosa CRFs were dispensed with and the home language was considered to be Afrikaans where both participants in an interview were Afrikaans speaking and all other interviews were conducted in English. This assignment gave an approximate 40:60 split between Afrikaans home language and English. Note that no interviews in English would be conducted between the two participants who spoke Afrikaans and English at home. They were classified as Afrikaans speakers, and so would conduct interviews in Afrikaans when interviewing each other.

7.2 Problem areas

Problems were encountered while performing the study, but these problems only came to light after the data capture exercise had been completed.

7.2.1 Low number of openXdata results

The previously identified issue with openXdata not having a standard start screen had a serious impact on the study. When performing the interviews on 16 March, the network connection was not reliable and the participants were advised to perform all interviews without uploading the results. These results would have been uploaded at a later stage.

When there is a saved result on openXdata, the home screen for openXdata shows the list of saved results and the visible commands on the soft keys are 'Options', 'Select' and 'Back'. When the 'Select' command is pressed, the highlighted saved interview is opened for editing. Some participants thought they were entering a new interview and did not realise that by pressing the 'Select' key they would overwrite the previously saved values. This error resulted in a number of openXdata interviews being overwritten and lost.

In addition, two participants found openXdata more difficult to use than MyApp and performed six interviews using MyApp instead of openXdata before this behaviour was discovered and stopped.

7.2.2 Multi-select questions

During training, a bug in MyApp was found with the handling of multi-select questions. This bug was fixed in time for the study, but unfortunately not all participants downloaded the revised application, which led to invalid answers being received for a number of the questionnaires.

As there was only one multi-select question in the questionnaire, it was decided omit multi-select questions from the analysis of all methods of capture.

7.2.3 Difficulty with dates when using MyApp

It was noticed that when participants were using MyApp, they seemed to experience difficulty with capturing dates. MyApp was configured to capture dates as yyyy/mm/dd, e.g. 2013/01/01 (see Section 4.3.4). However, subsequent to the completion of the exercise, participants discussed issues they encountered and the unfamiliar date format was a common issue. It turns out that the participants have been extensively trained to use the dd mmm yyyy format (e.g. 01 Jan 2013) and reversing the order of the date components confused them. The results were analysed to see if date fields had an impact on the outcome.

7.2.4 Repeating groups when using openXdata

The repeating group of data items relating to signs and symptoms troubled the participants when using openXdata. The behaviour of openXdata (as noted in Section 2.3.2) is to open a sub-questionnaire for each group of data items, and it appears unrelated to the main questionnaire. The results were analysed to see if these repeating data items had an impact on the outcome.

7.3 Participants

Ten participants took part in the study and were classified according to the following criteria:

- Home language – Afrikaans or Xhosa. Two participants indicated that they used both English and Afrikaans at home. For the purposes of this study, they were classified as Afrikaans speaking to integrate them into the group who could perform home-language interviews with all other Afrikaans speakers;
- Education – Grade 12/Matric or below Grade 12. None of the participants had tertiary education. All CRWs were required, as part of their job, to have undergone training in Good Clinical Practice and similar work-related courses. Therefore, there was no differentiation between them based on post-secondary training; and

- Computer literacy – The participants were either classified as 'computer literate' or 'non-computer literate'. The result was based on a three-part question:
  - Do you surf the net at home?
  - Do you use a computer at work?
  - Do you use email?

  If any one of these questions was answered positively, the user was classified as computer literate.

The profile of the group of participants is shown in Table 2 below:

Table 2 Participant profile

Computer literate	Level of education	Afrikaans home language	Xhosa home language	All languages
No	Grade 10/Junior Certificate	1	1	2
	Grade 12/Matric	1	1	2
Yes	Grade 12/Matric	4	2	6

7.4 Interviews

Despite there being 90 interviews defined (30 for each method of capture), only 64 results (sets of answers) were received electronically or on paper. In addition, a few interviews were performed using the incorrect capture method, as noted above. The profile of the interviews is shown below:

Table 3 Profile of interviews

	Specified method			
Actual method	MyApp	openXdata	Paper	Total
MyApp	22	6	1	29
openXdata		11		11
Paper	1	1	22	24
Not received	7	12	7	26
Total	30	30	30	90

The reduction in the number of interviews by each method may reduce the ability to detect significance in some of the secondary results, but it is unlikely to influence the primary result of the study, as the number of interviews exceeded the number suggest by the sample size calculation.

It is possible that the results captured by the incorrect method may lead to bias as the allocation is no longer entirely random. The detail of the incorrectly captured interviews is listed in Table 4 below.

Table 4 Incorrect interview methods

Specified method	Actual method	Interview language	Computer literacy	Education	Home language
Paper	MyApp	English	Yes	Grade 12/Matric	Afrikaans
openXdata	MyApp	English	No	Grade 12/Matric	Xhosa
openXdata	MyApp	English	No	Grade 12/Matric	Xhosa
openXdata	MyApp	Afrikaans	Yes	Grade 12/Matric	Afrikaans
openXdata	Paper	Afrikaans	Yes	Grade 12/Matric	Afrikaans
openXdata	MyApp	Afrikaans	Yes	Grade 12/Matric	Afrikaans
openXdata	MyApp	Afrikaans	No	Grade 10/Junior Cert	Afrikaans
openXdata	MyApp	English	No	Grade 10/Junior Cert	Afrikaans
MyApp	Paper	Afrikaans	Yes	Grade 12/Matric	Afrikaans

As can be seen, the characteristics of the interview show no clear pattern, and it is felt that these incorrectly captured interviews will not bias the results.

7.5 Quantitative Results

The measurements of accuracy are shown in Appendix C. They are stratified by various factors, as indicated.

7.6 Qualitative Results

The results of the post-study qualitative analysis are depicted in Figure 25 below.

Overall, ease of use was rated highly for both applications. Where the questions differentiate between MyApp and openXdata, openXdata fares less well, although still at an acceptable level. Points to note are:

- The MyApp residential question which made use of predictive text had an average of 4.9 out of 5, an indication that this feature was well received; and

- openXdata could not capture blood pressure in the special format, e.g. 120/70. Instead, the entries were defined as two questions for the systolic pressure and diastolic pressure. This abnormal way of recording the blood pressure was somewhat disliked, which was illustrated by the relatively high error rate for this question.

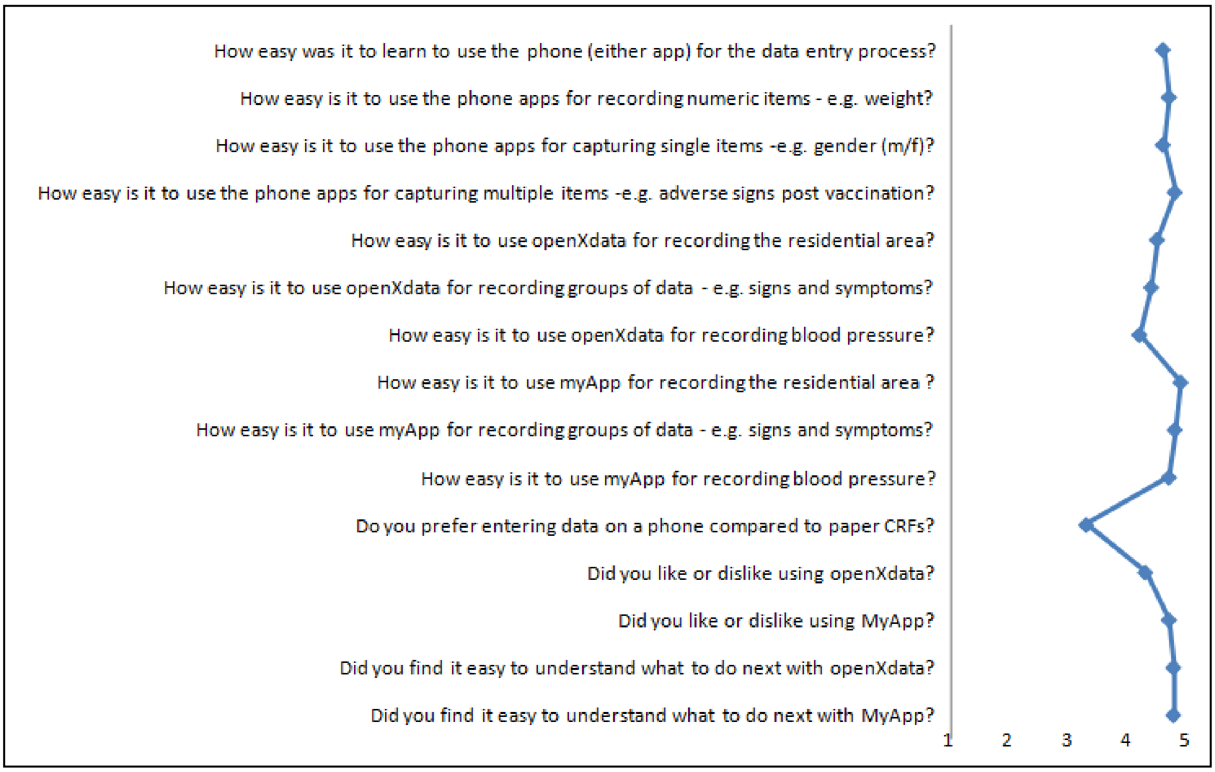

Figure 25 Qualitative Results

The original aim of the qualitative analysis was to determine if there were areas of the MyApp user interface that required reworking. Given the relatively high ratings, it seems that MyApp is perceived as being easy to use with no usability problems.

8. FINDINGS AND ANALYSIS

8.1 Exploratory analysis

An exploratory analysis was performed on proportion correct by level of education, computer literacy, home language, method of capture and answer data type. Further exploration was done for a combination of method of capture vs. level of education, computer literacy and home language. The results that were obtained, are profiled in Appendix C.

8.2 Quantitative results of interest

The results of interest obtained from the exploratory analysis (see Appendix C) are shown below.

The error rate across the study is unacceptably high as shown in Table 5, thus appearing to support previous research.

Table 5 All answers

Total answers	Correct	Correct %	Error	Error %
2080	1980	95.2%	100	4.8%

However, inspection shows that the proportion correct for the answers captured using openXdata are much lower than the remainder of the results. More detailed profiling of the answers by method of capture shown in Table 6 confirms that openXdata proportion correct is much lower than for paper or MyApp.

Table 6 Answers by method of capture

	Total	Correct	Correct %	Error	Error %
MyApp	958	931	97.2%	27	2.8%
Paper	759	739	97.4%	20	2.6%
openXdata	363	310	85.4%	53	14.6%
All Methods	2080	1980	95.2%	100	4.8%

Tables 7, 8 and 9 show the results profiled by level of education, computer literacy and interview language. A result of note is shown in Table 8, which profiles computer literacy vs. method of capture and shows a lower proportion correct for non-computer literate participants using openXdata than for those who are computer literate.

Table 7 Answers by level of education and data capture method

Education	Method	Total	Correct	Correct %	Error	Error %
Grade12	MyApp	891	864	97.0%	27	3.0%
	Paper	594	578	97.3%	16	2.7%
	openXdata	330	282	85.5%	48	14.5%
Grade10 or 11	MyApp	67	67	100.0%	0	0.0%
	Paper	165	161	97.6%	4	2.4%
	openXdata	33	28	84.8%	5	15.2%
All grades	MyApp	958	931	97.2%	27	2.8%
	Paper	759	739	97.4%	20	2.6%
	openXdata	363	310	85.4%	53	14.6%
Total		2080	1980	95.2%	100	4.8%

Table 8 Answers by computer literacy and data capture method

Computer literate	Method	Total	Correct	Correct %	Error	Error %
Yes	MyApp	627	602	96.0%	25	4.0%
	Paper	462	452	97.8%	10	2.2%
	openXdata	264	231	87.5%	33	12.5%
No	MyApp	331	329	99.4%	2	0.6%
	Paper	297	287	96.6%	10	3.4%
	openXdata	99	79	79.8%	20	20.2%
Both	MyApp	958	931	97.2%	27	2.8%
	Paper	759	739	97.4%	20	2.6%
	openXdata	363	310	85.4%	53	14.6%
Total		2080	1980	95.2%	100	4.8%

Table 9 Answers by interview language and data capture method

Education	Method	Total	Correct	Correct %	Error	Error %
Home language	MyApp	330	328	99.4%	2	0.6%
	Paper	330	323	97.9%	7	2.1%
	openXdata	165	143	86.7%	22	13.3%
English	MyApp	628	603	96.0%	25	4.0%
	Paper	429	416	97.0%	13	3.0%
	openXdata	198	167	84.3%	31	15.7%
Both	MyApp	958	931	97.2%	27	2.8%
	Paper	759	739	97.4%	20	2.6%
	openXdata	363	310	85.4%	53	14.6%
Total		2080	1980	95.2%	100	4.8%

From the above, it is clear that the answers captured using openXdata are problematic. It was hypothesised that answers may have been captured from the wrong interview sheet (e.g. expected answer = 07 September 1996, while the captured answer = 04 April 1996), but inspection revealed this hypothesis does not hold.

8.2.1 Results comparing Paper to MyApp

Profiling the MyApp answers vs. the answers recorded on paper (the control) gives the results shown in Appendix C.2, from which pertinent results are shown below.

Table 10 All answers

Total Answers	Correct	Correct %	Error	Error %
1717	1670	97.3%	47	2.7%

Table 11 Answers by method of capture

	Total	Correct	Correct %	Error	Error %
MyApp	958	931	97.2%	27	2.8%
Paper	759	739	97.4%	20	2.6%
All methods	1717	1670	97.3%	47	2.7%

Table 12 Answers by interview language and data capture method

Interview language	Method	Total	Correct	Correct %	Error	Error %
Home language	MyApp	330	323	97.9%	7	2.1%
	Paper	330	328	99.4%	2	0.6%
English	628	429	416	97.0%	13	3.0%
	Paper	628	603	96.0%	25	4.0%
Both languages	MyApp	759	739	97.4%	20	2.6%
	Paper	958	931	97.2%	27	2.8%
Total	Both	1717	1670	97.3%	47	2.7%

Table 13 Answers by computer literacy and data capture method

Computer literate	Method	Total	Correct	Correct %	Error	Error %
Yes	Paper	462	452	97.8%	10	2.2%
	MyApp	627	602	96.0%	25	4.0%
No	Paper	297	287	96.6%	10	3.4%
	MyApp	331	329	99.4%	2	0.6%
Both	Paper	759	739	97.4%	20	2.6%

	MyApp	958	931	97.2%	27	2.8%
Total	Both	1717	1670	97.3%	47	2.7%

Table 14 Answers by answer type and data capture method

Answer type	Method	Total	Correct	Correct%	Error	Error %
Date	MyApp	145	136	93.8%	9	6.2%
	Paper	115	111	96.5%	4	3.5%
	Both	260	247	95.0%	13	5.0%
Integer	MyApp	58	57	98.3%	1	1.7%
	Paper	46	44	95.7%	2	4.3%
	Both	104	101	97.1%	3	2.9%
Text	MyApp	117	114	97.4%	3	2.6%
	Paper	92	90	97.8%	2	2.2%
	Both	209	204	97.6%	5	2.4%
Real	MyApp	87	84	96.6%	3	3.4%
	Paper	69	67	97.1%	2	2.9%
	Both	156	151	96.8%	5	3.2%
Single select	MyApp	464	457	98.3%	7	1.5%
	Paper	368	358	97.3%	10	2.7%
	Both	832	815	97.8%	17	2.0%
Time	MyApp	87	83	95.4%	4	4.6%
	Paper	69	69	100.0%	0	0.0%
	Both	156	152	97.4%	4	2.6%
All types	Both	1717	1670	97.2%	47	2.7%

The above profiling now shows that there seems to be little difference between error rates when data are recorded on paper or captured using MyApp, generally in the 3% to 4% range. The one exception is capturing dates using MyApp, where the error rate is in excess of 6%. This result will be explored further in the analysis.

8.2.2 Results comparing Paper to openXdata

Profiling the openXdata answers vs. answers recorded on paper (the control) gives the results shown in Appendix C.3, from which pertinent results are shown below.

Table 15 All answers

Total answers	Correct	Correct %	Error	Error %
1122	1049	93.5%	73	6.5%

Table 16 Answers by method of capture

	Total	Correct	Correct %	Error	Error %
openXdata	363	310	85.4%	53	14.6%
Paper	759	739	97.4%	20	2.6%
All methods	1122	1049	93.5%	73	6.5%

Table 17 Answers by interview language and data capture method

Home language	Method	Total	Correct	Correct %	Error	Error %
Home language	Paper	330	323	97.9%	7	2.1%
	openXdata	165	143	86.7%	22	13.3%
English	Paper	429	416	97.0%	13	3.0%
	openXdata	198	167	84.3%	31	15.7%
Both languages	Paper	759	739	97.4%	20	2.6%
	openXdata	363	310	85.4%	53	14.6%
Total		1122	1049	93.5%	73	6.5%

Table 18 Answers by computer literacy and data capture method

Computer literate	Method	Total	Correct	Correct %	Error	Error %
Yes	Paper	462	452	97.8%	10	2.2%
	openXdata	264	231	87.5%	33	12.5%
No	Paper	297	285	96.0%	12	4.0%
	openXdata	99	79	79.8%	20	20.2%
Both	Paper	759	737	97.1%	22	2.9%
	openXdata	363	310	85.4%	53	14.6%
Total		1122	1047	93.3%	75	6.7%

The above results point to openXdata being possibly inferior to paper and will be verified by the analysis presented in the following section.

8.3 Analysis of results

The results were analysed using the STATA statistical package as defined in the study design (Section 3.1.12).

8.3.1 Analysis of MyApp vs. Paper

The null and alternate hypotheses for non-inferiority are stated as:

$$H_0: X - Y < -\Delta$$

$$H_A: X - Y \geq \Delta$$

where X represents recording on paper and Y represents data capture using MyApp. Analysing the results using a two sample test of proportion provides the following data:

Table 19 Analysis by method of capture

	Total	Correct	Proportion correct	One Tail 95% confidence interval
MyApp	958	931	0.97182	0.96302
Paper	759	739	0.97365	0.96409
Difference			-0.00183	-0.01483

As can be seen, the defined equivalence range (-0.010) is greater than the lower bound of the 95% confidence interval (-0.01601). Therefore, the null hypothesis cannot be rejected, and so it cannot be said that MyApp is not inferior to the paper CRF at this level.

However, as was noted above, the capture of dates using MyApp was problematic and the error rate for these answers was higher than all other answer types. When the date answer type for MyApp is excluded, the test for non-inferiority produced the following results:

Table 20 Analysis by method of capture excluding dates

	Total	Correct	Proportion correct	95% confidence interval
MyApp	813	795	0.97786	0.96791
Paper	759	739	0.97365	0.96409
Difference			0.00421	-0.00858

The defined equivalence range (-0.010) is now less than the lower bound of the 95% confidence interval (-0.00858). Therefore, the null hypothesis can be rejected, and it can be said that MyApp is not inferior to the paper CRF, provided that the problematic, high-error rate of the date-type answers can be reduced successfully.

8.3.2 Analysis of openXdata vs. Paper

Because the results for openXdata appeared to be much worse than those for paper, the non-inferiority test was not used and the two methods were compared using the Chi-square test of independence. The results of this test are shown in Table 21

Table 21 Analysis by method of capture

	Total	Correct	Proportion correct	95% confidence interval	
openXdata	363	310	0.85399	0.81767	0.89032
Paper	759	739	0.97365	0.96225	0.98504
Difference	-396	-427	-0.11966	-0.15773	-0.08158

P = 0.0000

As can be seen in Table 21, the 95% confidence intervals do not overlap, the 95% confidence intervals of the Difference do not cross the null, and $p < 0.05$. The null hypothesis is hence rejected and it can be accepted that openXdata is inferior to using paper to record data.

8.4 Secondary analysis

When exploring the results, it was observed that there may be results of interest relative to the computer literacy of the participants and the language the interview was conducted in (Afrikaans home language or English). These aspects were consequently investigated as a secondary part of the analysis.

8.4.1 Computer literacy

It was observed that there seemed to be a greater proportion correct when non-computer literate participants used MyApp to capture the answers relative to capturing them on paper. As the sample size was below five, the Fisher's Exact Test was used to verify independence.

Table 22 Analysis by computer literacy

Computer literate	Total	Correct	Proportion correct	Incorrect	Proportion incorrect
No	331	329	0.99396	2	0.00604
Yes	297	287	0.96633	10	0.03367

2-tail *p*-value = 0.0461

As the *p*-value < 0.05, the two samples are independent and it can be stated that the non-computer literate participants were more likely to capture data correctly using MyApp than recording them on paper.

The inverse effect was observed when non-computer literate participants who used openXdata to capture the answers were compared to those who were computer literate.

Table 23 Proportion correct for computer literacy vs. data capture method

	Non-computer literate	Computer literate
MyApp	328/331 = 99.1%	602/627 = 96.0%
openXdata	79/99 = 79.8%	231/264 = 87.5%

The proportion correct when using openXdata was lower for computer illiterate participants (79.8%) than for computer literate participants (87.5%).
This result was explored further using the Chi-square test to verify independence which produced the following outcome

Table 24 Analysis by computer literacy

Computer literate	Total	Correct	Proportion correct	Incorrect	Proportion incorrect
No	99	79	0.79798	20	0.20202
Yes	264	231	0.87500	33	0.12500

p-value = 0.0642

As the *p*-value > 0.05, the two samples are not independent and it cannot be stated that the non-computer literate participants were more likely to capture data incorrectly using openXdata than the computer literate participants.

8.4.2 Interview Language

It was hypothesised that interviews in the home language of both participants may produce more accurate answers as there would be no need for 'on-the-fly' translation by the participants. When inspecting the data, it appeared that this might be true to a very slight extent. A Chi-square test revealed that the difference was not significant at the $p < 0.05$ level as shown in Table 25.

Table 25 All answers by interview language

	Total	Correct	Correct %	Error	Error %
Home language	825	794	96.2%	31	3.8%
English	1255	1185	94.4%	70	5.6%
Both	2080	1979	95.1%	101	4.9%

p-value = 0.0589

9. CONCLUSION

This research aimed to show whether CRWs could capture data with sufficient accuracy using a feature phone, for this method to be by the clinical research profession.

Despite prior research in this area finding that error rates were excessive, it was hypothesised that, provided the CRWs used devices with which they are familiar, and provided the data capture user interface was simple enough and followed a paradigm with which the CRWs were familiar, then the CRWs should be able to capture data with sufficient accuracy to meet the requirements for clinical research. The grounds for this hypothesis are that the feature phone is in everyday use in the communities from which the CRWs are drawn and the use of the feature phone has grown significantly beyond simply making phone calls, to the point where complex transactions such a mobile banking, which requires absolute accuracy, are frequently performed.

To provide for comparison, we evaluated an existing open source application used for clinical research and other surveys globally including the Western Cape, as well as a custom application designed and built with a simplified user interface. Both of these applications were compared to recording data on paper CRFs.

The outcomes when using the two applications were very different, and the factors influencing the outcomes are discussed below.

9.1 The custom application

The outcome when capturing data using the custom application has proven to be satisfactory, but this outcome was not found unequivocally. The date data type showed higher error rates than all other data types. If the date data type is excluded from the results for the custom application, the outcome at the 95% confidence level is that, providing the problem with dates is overcome, it is possible for typical CRWs to capture data using a feature phone as accurately as when recording the data on paper forms.

An error in judgment when designing the CRF for the custom phone application resulted in dates being captured in a format unfamiliar to the participants and it is believed that this unfamiliarity lead to the high error rates for this type of data. It would have been trivial to change to a more familiar format, if uncovered earlier, but a drawback of having a single session for performing the mock interviews was that interim analyses of the data was not possible. A further study should be performed to verify that use of a date format familiar to the participants would lead to a better outcome. This emphasised how important it is for a system designer to follow the principle of 'knowing your user'.

The rewarding aspect of this outcome is that non-computer literate users were found to be able to capture data with at least the same accuracy as computer literate users.

The participants performed the bulk of the mock interviews some months after a very brief, one-hour training session, which leads to the conclusion that the user interface guideline of learnability was comprehensively met as the participants were still able to use the application after some time has passed. Reasons for the fact that this guideline could be met, include:

- The user interface design limited the number of actions that could be performed at any one time so that the use of sub-menus or context-sensitive menus was avoided. This limitation of actions followed from an analysis of the data capture process to ensure that this limitation was not undue and that real-world requirements could be achieved;
- In some instances, less desired features were not provided where the inclusion of these features would negatively impact the more desired features. For example, a 'Cursor Left without Delete' function was not provided as provision of both 'Cursor Left without Delete' and 'Backspace with Delete' would have required a more complex implementation for both functions;

- A consistent home screen was provided where the complete process was arranged in a menu structure. This consistency also enhanced the user predictability characteristic;
- All forms adhered to the same conceptual layout; and
- Elimination of free text capture and its replacement with predictive text was a definite success, not only limiting errors, but also mitigating the requirement for subsequent manual coding of the free text for analysis.

Determining the principles and factors to take into account for the user interface design was not a straightforward task. There are a number of books and documents available on user interface design, but most deal with PC and smartphone user interfaces. There are very few user interface design guides available for a feature phone, and of these guides, most seem to relate to manufacturer-specific instructions to achieve the manufacturer's desired look and feel, such as Nokia's Symbian guidelines. In the end, the set of rules defined by Shneiderman for PC user interfaces was used as a basis for the user interface design but tempered by the complications related to the small display and the cramped 12-key keypad.

When designing the application, the capabilities of the feature phone relative to the requirements were not known, i.e. how much memory and processing power would be required and whether the performance when transmitting and receiving data would be adequate. The decision to use a multi-level delimited text string both for storage and transmission of questionnaires and results may have contributed to the adequate performance of the custom application.

These points identified above, indicate that the overall design of the custom application was instrumental in the ability of the users to capture data with sufficient accuracy.

9.2 openXdata mforms

The outcome when capturing data using openXdata was not satisfactory and error rates were significantly greater than for both the paper forms and the custom application. During the mock interviews, it was not possible to send all results to the server, and consequently these results remained on the phone handset. This behaviour confused the participants and some of the results were 'lost' being overwritten by those of a later interview. Despite the loss of results, the primary outcome was still statistically significant.

Participants were trained for two and a half hours in workshops of groups of four participants, which is even more training time than is recommended. It is, however, possible that the informal style of training was not optimal for teaching openXdata and a more formal lecture style would have been better.

A number of issues were identified during the evaluation of open source applications, and these issues seem to have contributed to the unsatisfactory outcome. The lower-than-desired number of results has probably led to the secondary analysis not producing statistically significant outcomes. Therefore, it can only be said that the following are indications of possible causes of unacceptable error rates.

- Data in repeating groups were skipped or were incorrectly captured more often than for the equivalent data in the rest of the questionnaire. This erroneous result was possibly caused by the repeating group appearing as a sub-questionnaire without the context of the main questionnaire;
- There was no consistent home form – After completing a questionnaire, the application either started a new interview, or displayed the list of saved results with a different set of optional actions. This inconsistency led to user confusion and was the probable cause of interviews being overwritten; and
- Action options were inconsistent. On some forms there were two positive actions ('OK' and 'Select'), three negative actions ('Delete', 'Back', 'Exit') and other actions which led to a relatively large sub-menu requiring two clicks to complete an action. This inconsistency possibly led to a relatively large number of skipped values or incorrect values chosen from a list.

The error rates for non-computer literate participants (20%) were higher than for computer literate participants (12,5%). This outcome was not statistically significant (*p* value= 0.0642 compared to the desired *p*-value of 0.05), which was possibly the result of a too small a sample, following the loss of results, as noted before.

The loss of data that occurred would be unacceptable in an actual clinical research study. There are various factors to that led to this loss:

- The root cause of this loss was the inability to send results, even though the area was well served by the mobile network operators and the custom application was able to send (albeit somewhat slowly);
- The users had been specifically trained to send results back after each interview, and were not trained how to handle results that were retained on the handset;
- The lack of consistency in the user interface depending on whether there were unsent results or not, led to confusion;
- There was no warning when the user was about to edit previously saved data; and
- There is no audit trail of the changes that would have allowed the original data to be reconstructed.

9.3 Summary

This study has shown that the typical CRW, when using an application similar to the custom application (MyApp), can capture data with an accuracy equal to that of recording data on paper forms. However, it is felt that, given the large number of sites using openXdata, it would be unwise to state that openXdata is necessarily inadequate to capture data accurately. It is possible that further studies involving more skilled openXdata users and trainers would have a different outcome. However, what is clear from the results is that the simpler user interface of the custom application is eminently more learnable and hence conducive to lower training requirements and lower error rates.

A point of interest is the ability of the custom application to send results back to the server when openXdata could not. It is possible that the chosen minimalistic data structure was instrumental in achieving this outcome.

9.4 Usability guidelines for feature phone applications

As there seems to be a lack of usability guidelines for feature phones, and given the success of this research, the usability guidelines established in Section 4.2.2 could be considered for use in other applications. They are the following:

- Provide a single home form, which should display the business process in the form of a menu.
- Avoid multitasking by users. Multiple threads are acceptable for background processes but not for users.
- Classify the user actions into:
  - Task initiation actions which initiate a business task and which are triggered from a menu on the home form;
  - Process actions which affect the process state of the task – these actions are triggered by the left, right and possibly centre soft keys, e.g. 'Select' picks the highlighted answer, saves the answer and moves to the next form. Process actions can be sub-classed into 'Forward', 'Backward' or 'Holding', based on whether they change the process state toward completion, away from completion or if there is no change; and
  - Display actions which aid the user in completing the answer, e.g. the left directional key deletes the preceding character in a text box.
- Prioritise user actions. The priority sequence should be:
  1. 'Undo Previous Action';
  2. 'Most Desirable Forward Action; and
  3. 'Next most Desirable Action(s) – This could be 'Forward', 'Back' or 'Holding'.
- Avoid context menus (option menus). This guideline avoids multi-press actions but does require the elimination of the lower-priority user actions to a point where there are only two (left and right soft keys) and a possible third (centre soft key). Workarounds for the eliminated actions should be devised where required. For example, stepping back through the previously completed forms to the beginning could generate a cancel action.

- Implement the 'Back' ('Undo') action on the left soft key and 'Forward' action on the right soft key. This order is the natural order for left-to-right text formalisms but is contrary to the Nokia guidelines. For right-to-left languages, such as Arabic, this order should probably be reversed.

Application of these guidelines during user interface design could improve the usability of, what Jakob Nielsen calls, the 'primitive handset'.

9.5 Future directions

The custom user interface design can be accepted and taken forward. Additional work required is the following:

1. In order to meet the legal and regulatory requirements,
    - A strong user authentication and authorisation function is required for the application, with the ability to remotely deactivate a user's rights. This functionality will require careful consideration as the ability to use the application while out of contact with the main server must not be inhibited;
    - The data, or at least the identifiers in the data, must be encrypted to industry standards while saved on the phone handset and while being transmitted; and
    - An audit trail with an irrefutable electronic signature is required as part of the result payload sent to the server. This audit trail need not be a log of all keystrokes but should include all data entered or selected by the user. The implication of this requirement is that stepping back through the questionnaire will require the retention of the 'old' data with an indicator that the data has been 'deleted'.
2. To cater for the ability to operate in areas with poor and intermittent connections:
    - The current design should be examined and improved, where necessary, so that the design is optimal in terms of data transmission;

    - Good monitoring and user feedback is required during the transmission of data, so the user may make a decision to abort the transmission if desired; and
    - Handshaking or a form of '2-phase commit' is required to ensure data is not lost or duplicated.
3. Integration with an industry-standard CRDB, such as OpenClinica, is required for the handset application to be of use in the more rigorous clinical research trials. This would be done in concert with the implementation of the regulatory requirements to provide continuity across both platforms.

Before such an application can be used for clinical research, a business unit must take ownership of this application. This business unit will need to provide ongoing support and training as well as develop the necessary quality assurance and standard procedures in terms of FDA Section 21 Part 11 (FDA 2013a).

10. LIST OF REFERENCES

Academy of Science of South Africa 2009, *Revitalising Clinical Research in South Africa*, Academy of Science of South Africa.

Akaza Research 2012, "OpenClinica Product Features," viewed 10 June, 2013, <https://www.openclinica.com/product-features>.

Allen, IE & Seaman, CA 2007, "Statistics Roundtable: Superiority, Equivalence and Non-inferiority," *Quality Progress - American Society for Quality*, viewed 28 May, 2013, <http://asq.org/quality-progress/2007/02/statistics-roundtable/superiority-equivalence-and-non-inferiority.html>.

Ballard, B 2005, *User Interface Design Guidelines for J2ME MIDP 2.0*, Lulu.com.

Delmas, N (ITU) 2013, "Key 2006-2013 ICT data," viewed 15 May, 2013, <http://www.itu.int/ITU-D/ict/statistics/at_glance/KeyTelecom.html>.

Diaper, D & Stanton, NA 2004, *The Handbook of Task Analysis for Human-Computer Interaction*, Lawrence Ehrlbaum Associates, Publishers.

Dix, A, Finlay, J, Abowd, GD & Beale, R 2004, *Human-Computer Interaction*, Pearson Education Limited, pp. 260–283.

Dunlop, MD & Crossan, A 2000, "Predictive Text Entry Methods for Mobile Phones," *Personal Technologies*, vol. 4, no. 2-3, pp. 134–143.

EDM Forum 2004, *An Approach to Benchmarking EDC Performance*.

FDA 2013a, *CFR - Code of Federal Regulations Title 21 PART 11 Electronic Records; Electronic Signatures*, United States of America, viewed 16 June, 2013, <http://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfcfr/cfrsearch.cfm?cfrpart=11>.

FDA 2013b, *Guidance for Industry Electronic Source Data in Clinical Investigations Guidance for Industry Electronic Source Data in Clinical Investigations*, United States of America.

Fox, N, Hunn, A & Mathers, N 2007, *Sampling and Sample Size Calculation*.

Free, C, Phillips, G, Watson, L, Galli, L, Felix, L, Edwards, P, Patel, V & Haines, A 2013, "The effectiveness of mobile-health technologies to improve health care service delivery processes: a systematic review and meta-analysis.," T Cornford (ed.), *PLoS medicine*, vol. 10, no. 1, Public Library of Science, p. e1001363, viewed 28 May, 2013, <http://dx.plos.org/10.1371/journal.pmed.1001363>.

Godlee, F, Smith, J & Marcovitch, H 2011, "Wakefield's article linking MMR vaccine and autism was fraudulent.," *BMJ (Clinical research ed.)*, vol. 342, no. jan05_1, p. c7452, viewed 16 June, 2013, <http://www.bmj.com/content/342/bmj.c7452>.

Grimes, DA & Schulz, KF 2002, "An overview of clinical research: the lay of the land.," *Lancet. 1/5/2002*, vol. 359, no. 9300, p. 57.

Haller, G, Haller, DM, Courvoisier, DS & Lovis, C 2009, "Handheld vs. laptop computers for electronic data collection in clinical research: a crossover randomized trial.," *Journal of the American Medical Informatics Association : JAMIA*, vol. 16, no. 5, pp. 651–9.

Hawkridge, A, Hatherill, M, Little, F, Goetz, MA, Barker, L, Mahomed, H, Sadoff, J, Hanekom, W, Geiter, L & Hussey, G 2008, "Efficacy of percutaneous versus intradermal BCG in the prevention of tuberculosis in South African infants: randomised trial.," *BMJ (Clinical research ed.)*, vol. 337, p. a2052, viewed 26 May, 2013, <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2583390&tool=pmcentrez&rendertype=abstract>.

Hoeller, J & McHenry, A 2011, "Insights into Mobile Telecoms in Africa."

Hogan, WR & Wagner, MM 1996, "Free-text fields change the meaning of coded data.," *Proceedings : a conference of the American Medical Informatics Association / ... AMIA Annual Fall Symposium. AMIA Fall Symposium*, pp. 517–21, viewed 26 October, 2013, <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2233180&tool=pmcentrez&rendertype=abstract>.

Iarossi, G 2013, *The Africa Competitiveness Report 2013 - Benchmarking Africa's Costs and Competitiveness*, p. 91, viewed 14 May, 2013, <http://siteresources.worldbank.org/EXTAFRSUMAFTPS/Resources/chapter4.pdf>.

IDC 2013, "Europe, Middle East and Africa Quarterly Mobile Phone Tracker," viewed 30 June, 2013, <http://www.idc.com/tracker/showproductinfo.jsp?prod_id=181>.

International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use 1996, *Guideline for Good Clinical Practice E6(R1)*.

ISO 2004, "iso8601," viewed 9 August, 2013, <http://www.iso.org/iso/home/standards/iso8601.htm>.

ITU-E.161 2001, "E.161 : Arrangement of digits, letters and symbols on telephones and other devices that can be used for gaining access to a telephone network," viewed 25 April, 2013, <http://www.itu.int/rec/T-REC-E.161/en>.

Kleppinger, CF & Ball, LK 2010, "Building quality in clinical trials with use of a quality systems approach.," *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America*, vol. 51 Suppl 1, no. Supplement 1, Oxford University Press, pp. S111–6, viewed 24 May, 2013, <http://cid.oxfordjournals.org/content/51/Supplement_1/S111.full>.

Lu, Z & Su, J 2010, "Clinical data management : Current status , challenges , and future directions from industry perspectives," *Open Access Journal of Clinical Trials*, pp. 93–105.

Nahm, ML, Pieper, CF & Cunningham, MM 2008, "Quantifying data quality for clinical trials using electronic data capture.," RW Scherer (ed.), *PloS one*, vol. 3, no. 8, Public Library of Science, p. e3049, viewed 13 June, 2013, <http://dx.plos.org/10.1371/journal.pone.0003049>.

Nielsen, J 2011a, "Mobile Usability Update (Jakob Nielsen's Alertbox)," viewed 10 October, 2012, <http://www.useit.com/alertbox/mobile-usability.html>.

Nielsen, J 2011b, "Mobile UX Sharpens Usability Guidelines," viewed 23 April, 2013, <http://www.nngroup.com/articles/mobile-sharpens-usability-guidelines/>.

Nokia 2012, "Feature phone," *Nokia Developer Wiki*, viewed 1 July, 2013, <http://www.developer.nokia.com/Community/Wiki/Feature_phone>.

Nokia 2013, "How to use Commands in Java ME," *Nokia Developer Wiki*, viewed 19 July, 2013, <http://developer.nokia.com/Community/Wiki/How_to_use_Commands_in_Java_ME>.

Ogunmefun, C, Mothibe, N & Friedman, I 2010, *Electronic Community Care Giver (CCG) Study*.

Parliamentary Communication Services 2013, "Protection of Personal Information Bill Approved," viewed 7 October, 2013, <http://www.parliament.gov.za/live/content.php?Item_ID=3919>.

PC Magazine Encyclopaedia n.d., "feature phone Definition from PC Magazine Encyclopedia," *PC Magazine*, viewed 1 July, 2013, <http://www.pcmag.com/encyclopedia/term/62894/feature-phone>.

Portio Research 2013, "Mobile Factbook 2013," viewed 15 May, 2013, <http://www.portioresearch.com/media/3986/Portio Research Mobile Factbook 2013.pdf>.

Republic of South Africa 2012, *Electronic Communications and Transactions Act*, Republic of South Africa.

Republic of South Africa 2009, *Protection of Personal Information Bill*, Republic of South Africa.

Richardson, D 2006, *The Internet and rural and agricultural development: an integrated approach*, viewed 14 May, 2013, <http://www.fao.org/docrep/W6840E/w6840e06.htm>.

Santos, AF & Soares, MA 2010, "HIV Genetic Diversity and Drug Resistance.," *Viruses*, vol. 2, no. 2, pp. 503–31, viewed 4 October, 2012, <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3185604&tool=pmcentrez&rendertype=abstract>.

SAQA 2012, "National Certificate: Community Health Work," viewed 7 October, 2013, <http://allqs.saqa.org.za/showQualification.php?id=64749>.

Scurtu, M 2013, "Moving beyond just EDC," viewed 2 October, 2013, <http://www.linkedin.com/groupItem?view=&gid=1066557&type=member&item=255468301&qid=1fe27728-91a0-4bf0-a63e-fb332e8b45e3&trk=groups_most_popular-0-b-ttl&goback=%2Egmp_1066557>.

Shelby-James, TM, Abernethy, AP, McAlindon, A & Currow, DC 2007, "Handheld computers for data entry: high tech has its problems too," *Biomed Central*, viewed 23 October, 2012, <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1804282/>.

Shneiderman, B & Plaisant, C 2005, *Designing the User Interface*, Addison-Wesley.

Skuse, A & Cousins, T 2007, "Managing Distance: Rural Poverty and the Promise of Communication in Post-Apartheid South Africa," *Journal of Asian and African Studies*, vol. 42, no. 2, pp. 185–207, viewed 16 May, 2013, <http://jas.sagepub.com/cgi/doi/10.1177/0021909607074867>.

Tatalović, M 2013, "Evidence lacking on mHealth effectiveness in poor countries," *SciDev.Net*, viewed 26 May, 2013, <http://www.scidev.net/en/health/news/evidence-lacking-on-mhealth-effectiveness-in-poor-countries.html>.

The European Parliament & The Council of the European Union 2001, *Directive 2001/20/EC of the European Parliament and of the Council of the European Union*, European Union, pp. 34–44.

Tomlinson, M, Solomon, W, Singh, Y, Doherty, T, Chopra, M, Ijumba, P, Tsai, AC & Jackson, D 2009, "The use of mobile phones as a data collection tool: a report from a household survey in South Africa.," *BMC medical informatics and decision making*, vol. 9, no. 1, p. 51, viewed 10 October, 2012, <http://www.biomedcentral.com/1472-6947/9/51>.

United States of America House of Representatives & Senate, US of A 1996, *HEALTH INSURANCE PORTABILITY AND ACCOUNTABILITY ACT OF 1996*, United States of America.

Walther, B, Hossin, S, Townend, J, Abernethy, N, Parker, D & Jeffries, D 2011, "Comparison of electronic data capture (EDC) with the standard data capture method for clinical trial data.," G Biondi-Zoccai (ed.), *PloS one*, vol. 6, no. 9, Public Library of Science, p. e25348, viewed 6 October, 2012, <http://dx.plos.org/10.1371/journal.pone.0025348>.

WHO 2008, *Ethopia's Human Resources for Health Programme*, viewed 15 August, 2013, <http://www.who.int/workforcealliance/knowledge/case_studies/Ethiopia.pdf>.

WHO 2012a, "Update on artemisinin resistance - April 2012."

WHO 2012b, "WHO | HIV/AIDS Fact Sheet," World Health Organization, viewed 24 October, 2012, <http://www.who.int/mediacentre/factsheets/fs360/en/index.html>.

WHO 2013a, "WHO | Malaria Fact Sheet," World Health Organization, viewed 24 October, 2012, <http://www.who.int/mediacentre/factsheets/fs104/en/>.

WHO 2013b, "WHO | Tuberculosis Fact Sheet," World Health Organization, viewed 24 October, 2012, <http://www.who.int/mediacentre/factsheets/fs104/en/>.

Workman, M 2012, "How much training? - Google Groups," *Google Group Forum*, viewed 1 November, 2012, <https://groups.google.com/forum/?fromgroups=#!topic/openxdata-users/2Xr2U_GktII>.

Ziphorah, RM 2011, "The Provision of Basic Levels of Computer Literacy and Education to the Neediest Community in South Africa," *International Journal for e-Learning Security*, vol. 1, no. December, pp. 96–103, viewed <http://infonomics-society.org/IJeLS/The Provision of Basic Levels of Computer Literacy and Education to the Neediest Community in South Africa.pdf>.

APPENDIX A. CASE REPORT FORMS

A.1 English CRF

PARTICIPANT ENROLMENT NUMBER: _ _ _ _
PARTICIPANT RESIDENTIAL LOCATION: _ _ _ _ _ _ _ _ _ _ _
PARTICIPANT DATE OF BIRTH: _ _ /_ _ _ / _ _ _ _

INFORMED CONSENT	
CONSENT signed?	Yes No
Version ICF signed (version number e.g. 1.0,2.1, or 3.2 etc)	_ . _
Date ICF signed	_ _ /_ _ _ / _ _ _ _

MEDICAL HISTORY AND PHYSICAL EXAMINATION	
Examination performed?	Yes No
Were there abnormalities?	Yes No

	SCREENING RESULTS (circle applicable)		
CHEST X-RAY		NORMAL	ABNORMAL
PREGNANCY TEST	N/A	NEGATIVE	POSITIVE
QUANTIFERON	INDETERMINATE	NEGATIVE	POSITIVE
		QFT VALUE (numeric value):	_ _ . _

BASELINE MEDICAL CONDITIONS				
Signs or symptoms		Ongoing? (circle)	Date of diagnosis	Date of resolution (if ongoing, leave blank)
1.		Yes / No	_ _/ _ _ _/_ _	_ _/ _ _ _ /_ _
2.		Yes / No	_ _/ _ _ _/_ _	_ _/ _ _ _ /_ _
3.		Yes / No	_ _/ _ _ _/_ _	_ _/ _ _ _ /_ _
4.		Yes / No	_ _/ _ _ _/_ _	_ _/ _ _ _ /_ _
5.		Yes / No	_ _/ _ _ _/_ _	_ _/ _ _ _ /_ _

INCLUSION AND EXCLUSION	
ALL INCLUSION CRITERIA MET?	Yes No
ANY EXCLUSION CRITERIA PRESENT?	Yes No
IS PARTICIPANT ELIGIBLE?	Yes No

SAMPLE COLLECTION	
BLOOD SAMPLES TAKEN FOR (circle 1):	Safety Bloods and Immunology Safety Bloods only Immunology only No Samples
IF NO SAMPLES, CHOOSE REASON (circle 1):	Misbleed Participant refused Not required Lab closed
TIME OF BLOOD TAKEN	_ _ h _ _
INITIAL OF TEAM MEMBER TAKING BLOOD:	_ _

VACCINATION	
PRESCRIPTION SIGNED? (circle)	Yes No
TIME OF VACCINE PREPERATION	_ _ h _ _
DATE OF VACCINATION	_ _ / _ _ _ / _ _
VACCINE ADMINISTERED? (circle)	Yes No
TIME OF VACCINATION	_ _ h _ _
SITE OF VACCINATION (circle)	LEFT ARM / RIGHT ARM
INITIAL OF VACCINATOR	_ _
ADVERSE EVENTS WITHIN 60 MINUTES OF VACCINATION?	Yes No
IF AE OCCURRED, WHAT IS NATURE OF AE? (circle)	Redness Pain Rash Swelling Other

VITAL SIGNS AT 60 MINUTES POST-VACCINATION	
BLOOD PRESSURE	_ _ _ / _ _ _ mm Hg
HEART RATE	_ _ _ / per minute
TEMPERATURE	_ _ . _ °C
RESPIRATORY RATE	_ _ per minute

	INITIAL	DATE
COMPLETED BY		

A.2 Afrikaans CRF

DEELNEMER INSKRYWINGSNOMMER: _ _ _ _
DEELNEMER SE BLYPLEK: _ _ _ _ _ _ _ _ _ _ _
DEELNEMER GEBOORTEDATUM: _ _ /_ _ _ / _ _ _ _

INGELIGTE TOESTEMMING	
TOESTEMMING geteken?	Ja Nee
Weergawe ICF geteken (weergawe nommer bv. 1.0,2.1, or 3.2 ens.)	_ . _
Datum ICF geteken	_ _ /_ _ _ / _ _ _ _

MEDIESE GESKIENDENIS EN FISIESE ONDERSOEK	
Ondersoek goedoen?	Ja Nee
Was daar abnormaliteite?	Ja Nee

	SIFTINGS UITSLAE (omkring toepaslike)		
BORS X-STRAAL		NORMAAL	ABNORMAAL
SWANGERSKAP TOETS	NVT	NEGATIEF	POSITIEF
QUANTIFERON	INDERTIMINATE	NEGATIEF	POSITIEF
		QFT WAARDE (numeriese waarde):	_ _ . _

BASISLYN MEDIESE TOESTANDE				
Simptome of tekens		Deurlopend? (omkring)	Datum van diagnose	Datum van Opklaring (indien deurplopend, los oop)
1.		Ja / Nee	_ _/ _ _ _/_ _	_ _/ _ _ _ /_ _
2.		Ja / Nee	_ _/ _ _ _/_ _	_ _/ _ _ _ /_ _
3.		Ja / Nee	_ _/ _ _ _/_ _	_ _/ _ _ _ /_ _
4.		Ja / Nee	_ _/ _ _ _/_ _	_ _/ _ _ _ /_ _
5.		Ja / Nee	_ _/ _ _ _/_ _	_ _/ _ _ _ /_ _

INSLUTING EN UITSLUITING	
ALLE INSLUITINGSKRITERIA NAGEKOM?	Ja Nee
ENIGE UITSLUITINGSKRITERIA TEENWOORDIG?	Ja Nee
IS DEELNEMER GESKIK?	Ja Nee

MONSTER VERKRYGING	
BLOED MONSTERS VERKRY VIR (omkring 1):	Veiligsheidbloede and Immunologie Net Veiligshedibloede Net Immunologie Geen Monsters
INDIEN GEEN MONSTERS, KIES REDE (omkring 1):	Bloed onverkygbaar (Misbleed) Deelnemer weier Nie nodig nie Lab toe
TYD BLOED GENEEM	_ _ h _ _
VOORLETTER VAN SPAN LID WAT BLOED NEEM:	_ _

INENTING	
VOORSKRIF GETEKEN? (omkring)	Ja Nee
TYD VAN VOORBERIEDING VAN VAKSIENE	_ _ h _ _
DATUM VAN INENTING	_ _ / _ _ _ / _ _
INENTING TOEGEDIEN? (omkring)	Ja Nee
TYD VAN INENTING	_ _ h _ _
PLEK VAN INENTING (omkring)	LINKER ARM / REGTER ARM
VOORLETTER VAN PERSOON WAT INENT	_ _
NADELIGE EFFEKTE BINNE 60 MINUTE NA INENTING?	Ja Nee
INDIEN NADELIGE EFFEK TEEN WOORDIG IS, WAT IS AARD VAN DIE EFFEK? (omkring)	Rooiheid Pyn Veluitslag Swelsel Ander

VITALE TEKENS BY 60 MINUTE NA INENTING	
BLOEDDRUK	_ _ _ / _ _ _ mm Hg
HARTSPOED	_ _ _ / per minuut
TEMPERATUUR	_ _ . _ °C
AEMHALINGSSPED	_ _ per minuut

	VOORLETTER	DATUM
VOLTOOI DEUR		

APPENDIX B. END OF PROJECT SURVEY FORM

CELL PHONE PROJECT - END OF STUDY SURVEY

How easy was it to learn to use the phone (either app) for the data entry process?	① ② ③ ④ ⑤ difficult easy
How easy is it to use the phone applications for capturing single items -e.g. gender (m/f)?	① ② ③ ④ ⑤ difficult easy
How easy is it to use the phone applications for capturing multiple items -e.g. adverse signs post vaccination?	① ② ③ ④ ⑤ difficult easy
How easy is it to use the phone applications for recording numeric items - e.g. weight?	① ② ③ ④ ⑤ difficult easy
How easy is it to use MyApp for recording the residential area ?	① ② ③ ④ ⑤ difficult easy
How easy is it to use openXdata for recording the residential area?	① ② ③ ④ ⑤ difficult easy
How easy is it to use MyApp for recording blood pressure?	① ② ③ ④ ⑤ difficult easy
How easy is it to use openXdata for recording blood pressure?	① ② ③ ④ ⑤ difficult easy
How easy is it to use MyApp for recording groups of data - e.g. signs and symptoms?	① ② ③ ④ ⑤ difficult easy
How easy is it to use openXdata for recording groups of data - e.g. signs and symptoms?	① ② ③ ④ ⑤ difficult easy
Do you prefer entering data on a phone compared to paper CRFs?	① ② ③ ④ ⑤ phone same paper
Did you find it easy to understand what to do next with MyApp?	① ② ③ ④ ⑤ difficult easy
Did you find it easy to understand what to do next with openXdata?	① ② ③ ④ ⑤ difficult easy
Did you like or dislike using MyApp?	① ② ③ ④ ⑤ disliked even liked
Did you like or dislike using openXdata?	① ② ③ ④ ⑤ disliked even liked

APPENDIX C. TABLES OF RESULTS

C.1 All results

Table 26 All Answers

Total Answers	Correct	Correct %	Error	Error %
2080	1980	95.2%	100	4.8%

Table 27 Answers by Level of Education

	Total	Correct	Correct %	Error	Error %
Grade12	1815	1724	95.0%	91	5.0%
Grade10 or 11	265	256	96.6%	9	3.4%
All Grades	2080	1980	95.2%	100	4.8%

Table 28 Answers by Interview Language

	Total	Correct	Correct %	Error	Error %
Home Language	825	794	96.2%	31	3.8%
English	1255	1186	94.5%	69	5.5%
Both	2080	1980	95.2%	100	4.8%

Table 29 Answers by Home Language

	Total	Correct	Correct %	Error	Error %
Afrikaans	1222	1174	96.1%	48	3.9%
Xhosa	858	806	93.9%	52	6.1%
Both	2080	1980	95.2%	100	4.8%

Table 30 Answers by Computer Literacy

	Total	Correct	Correct %	Error	Error %
Literate	1353	1285	95.0%	68	5.0%
Not Literate	727	695	95.6%	32	4.4%
Both	2080	1980	95.2%	100	4.8%

Table 31 Answers by Answer Type

	Total	Correct	Correct %	Error	Error %
date	315	288	91.4%	27	8.6%
integer	126	120	95.2%	6	4.8%
real	189	183	96.8%	6	3.2%
time	189	181	95.8%	8	4.2%
text	253	238	94.1%	15	5.9%

s-s	1008	970	96.2%	38	3.8%
All Types	2080	1980	95.2%	100	4.8%

Table 32 Answers by method of Capture

	Total	Correct	Correct %	Error	Error %
MyApp	958	931	97.2%	27	2.8%
Paper	759	739	97.4%	20	2.6%
openXdata	363	310	85.4%	53	14.6%
All Methods	2080	1980	95.2%	100	4.8%

Table 33 Answers by Level of Education and Data Capture Method

Education	Method	Total	Correct	Correct %	Error	Error %
Grade12	MyApp	891	864	97.0%	27	3.0%
	paper	594	578	97.3%	16	2.7%
	openXdata	330	282	85.5%	48	14.5%
Grade10 or 11	MyApp	67	67	100.0%	0	0.0%
	Paper	165	161	97.6%	4	2.4%
	openXdata	33	28	84.8%	5	15.2%
All Grades	MyApp	958	931	97.2%	27	2.8%
	Paper	759	739	97.4%	20	2.6%
	openXdata	363	310	85.4%	53	14.6%
Total		2080	1980	95.2%	100	4.8%

Table 34 Answers by Computer Literacy and Data Capture Method

Computer Literate	Method	Total	Correct	Correct %	Error	Error %
Literate	MyApp	627	602	96.0%	25	4.0%
	paper	462	452	97.8%	10	2.2%
	openXdata	264	231	87.5%	33	12.5%
Not Literate	MyApp	331	329	99.4%	2	0.6%
	Paper	297	287	96.6%	10	3.4%
	openXdata	99	79	79.8%	20	20.2%
Both	MyApp	958	931	97.2%	27	2.8%
	Paper	759	739	97.4%	20	2.6%
	openXdata	363	310	85.4%	53	14.6%
Total		2080	1980	95.2%	100	4.8%

Table 35 Answers by Interview Language and Data Capture Method

Education	Method	Total	Correct	Correct %	Error	Error %
Home	MyApp	330	328	99.4%	2	0.6%

Language	paper	330	323	97.9%	7	2.1%
	openXdata	165	143	86.7%	22	13.3%
English	MyApp	628	603	96.0%	25	4.0%
	Paper	429	416	97.0%	13	3.0%
	openXdata	198	167	84.3%	31	15.7%
Both	MyApp	958	931	97.2%	27	2.8%
	Paper	759	739	97.4%	20	2.6%
	openXdata	363	310	85.4%	53	14.6%
Total		2080	1980	95.2%	100	4.8%

C.2 Results for Paper vs. MyApp

Table 36 All Answers

Total Answers	Correct	Correct %	Error	Error %
1717	1670	97.3%	47	2.7%

Table 37 Answers by Level of Education

	Total	Correct	Correct %	Error	Error %
Grade12	1485	1442	97.1%	43	2.9%
Grade10 or 11	232	228	98.3%	4	1.7%
All Grades	1717	1670	97.3%	47	2.7%

Table 38 Answers by Interview Language

	Total	Correct	Correct %	Error	Error %
Home Language	660	651	98.6%	9	1.4%
English	1057	1019	96.4%	38	3.6%
Both	1717	1670	97.3%	47	2.7%

Table 39 Answers by Computer Literacy

	Total	Correct	Correct %	Error	Error %
Literate	1089	1054	96.8%	35	3.2%
Not Literate	628	616	98.1%	12	1.9%
Both	1717	1670	97.3%	47	2.7%

Table 40 Answers by Answer Type

	Total	Correct	Correct %	Error	Error %
date	260	247	95.0%	13	5.0%
integer	104	101	97.1%	3	2.9%
real	156	151	96.8%	5	3.2%
time	156	152	97.4%	4	2.6%

text	209	204	97.6%	5	2.4%
s-s	832	815	98.0%	17	2.0%
All Types	1717	1670	97.3%	47	2.7%

Table 41 Answers by Method of Capture

	Total	Correct	Correct %	Error	Error %
MyApp	958	931	97.2%	27	2.8%
Paper	759	739	97.4%	20	2.6%
All Methods	1717	1670	97.3%	47	2.7%

Table 42 Answers by Level of Education and Data Capture Method

Education	Method	Total	Correct	Correct %	Error	Error %
Grade12	Paper	594	578	97.3%	16	2.7%
	MyApp	891	864	97.0%	27	3.0%
Grade10 or 11	Paper	165	161	97.6%	4	2.4%
	MyApp	67	67	100.0%	0	0.0%
All Grades	Paper	759	739	97.4%	20	2.6%
	MyApp	958	931	97.2%	27	2.8%
Total		1717	1670	97.3%	47	2.7%

Table 43 Answers by Computer Literacy and Data Capture Method

Computer Literate	Method	Total	Correct	Correct %	Error	Error %
Yes	Paper	462	452	97.8%	10	2.2%
	MyApp	627	602	96.0%	25	4.0%
No	Paper	297	287	96.6%	10	3.4%
	MyApp	331	329	99.4%	2	0.6%
Both	Paper	759	739	97.4%	20	2.6%
	MyApp	958	931	97.2%	27	2.8%
Total		1717	1670	97.3%	47	2.7%

Table 44 Answers by Interview Language and Data Capture Method

Home Language	Method	Total	Correct	Correct %	Error	Error %
Home Language	Paper	330	323	97.9%	7	2.1%
	MyApp	330	328	99.4%	2	0.6%
English	Paper	429	416	97.0%	13	3.0%
	MyApp	628	603	96.0%	25	4.0%
Both Languages	Paper	759	739	97.4%	20	2.6%
	MyApp	958	931	97.2%	27	2.8%

Total		1717	1670	97.3%	47	2.7%

Table 45 Answers by Answer Type and Data Capture Method

Answer Type	Method	Total	Correct	Correct%	Error	Error %
Date	MyApp	145	136	93.8%	9	6.2%
	Paper	115	111	96.5%	4	3.5%
	Both	260	247	95.0%	13	5.0%
Integer	MyApp	58	57	98.3%	1	1.7%
	Paper	46	44	95.7%	2	4.3%
	Both	104	101	97.1%	3	2.9%
Text	MyApp	117	114	97.4%	3	2.6%
	Paper	92	90	97.8%	2	2.2%
	Both	209	204	97.6%	5	2.4%
Real	MyApp	87	84	96.6%	3	3.4%
	Paper	69	67	97.1%	2	2.9%
	Both	156	151	96.8%	5	3.2%
Single Select	MyApp	464	457	98.3%	7	1.5%
	Paper	368	358	97.3%	10	2.7%
	Both	832	815	97.8%	17	2.0%
Time	MyApp	87	83	95.4%	4	4.6%
	Paper	69	69	100.0%	0	0.0%
	Both	156	152	97.4%	4	2.6%
All Types	Both	1717	1670	97.2%	47	2.7%

C.3 Results for Paper vs. MyApp after removing date question type

Removal of the date question type gave the following profile of the answers:

Table 46 All Answers

Total Answers	Correct	Correct %	Error	Error %
1572	1534	97.6%	38	2.4%

Table 47 Answers by Level of Education

	Total	Correct	Correct %	Error	Error %
Grade12	1350	1316	97.5%	34	2.5%
Grade10 or 11	222	218	98.2%	4	1.8%
All Grades	1572	1534	97.6%	38	2.4%

Table 48 Answers by Interview Language

	Total	Correct	Correct %	Error	Error %
Home Language	610	601	98.5%	9	1.5%
English	962	933	97.0%	29	3.0%
Both	1572	1534	97.6%	38	2.4%

Table 49 Answers by Computer Literacy

	Total	Correct	Correct %	Error	Error %
Literate	994	968	97.4%	26	2.6%
Not Literate	578	566	97.9%	12	2.1%
Both	1572	1534	97.6%	38	2.4%

Table 50 Answers by Question Type

	Total	Correct	Correct %	Error	Error %
Date	115	111	96.5%	4	3.5%
Integer	104	101	97.1%	3	2.9%
Real	156	151	96.8%	5	3.2%
Time	156	152	97.4%	4	2.6%
Text	209	204	97.6%	5	2.4%
Single Select	832	815	98.0%	17	2.0%
All Types	1572	1534	98.0%	38	2.0%

Table 51 Answers by Method of Capture

	Total	Correct	Correct %	Error	Error %
MyApp	813	795	97.8%	18	2.2%
Paper	759	739	97.4%	20	2.6%
All Methods	1572	1534	97.6%	38	2.4%

Table 52 Answers by Level of Education and Data Capture Method

Education	Method	Total	Correct	Correct %	Error	Error %
Grade12	Paper	594	578	97.3%	16	2.7%
	MyApp	756	738	97.6%	18	2.4%
Grade10 or 11	Paper	165	161	97.6%	4	2.4%
	MyApp	57	57	100.0%	0	0.0%
All Grades	Paper	759	739	97.4%	20	2.6%
	MyApp	813	795	97.8%	18	2.2%
Total		1572	1534	97.6%	38	2.4%

Table 53 Answers by Computer Literacy and Data Capture Method

Computer Literate	Method	Total	Correct	Correct %	Error	Error %
Yes	Paper	462	452	97.8%	10	2.2%
	MyApp	532	516	97.0%	16	3.0%
No	Paper	297	287	96.6%	10	3.4%
	MyApp	281	279	99.3%	2	0.7%
Both	Paper	759	739	97.4%	20	2.6%
	MyApp	813	795	97.8%	18	2.2%
Total		1572	1534	97.6%	38	2.4%

Table 54 Answers by Interview Language and Data Capture Method

Home Language	Method	Total	Correct	Correct %	Error	Error %
Home Language	MyApp	280	278	99.3%	2	0.7%
	Paper	330	323	97.9%	7	2.1%
English	MyApp	533	517	97.0%	16	3.0%
	Paper	429	416	97.0%	13	3.0%
Both Languages	Paper	813	795	97.8%	18	2.2%
	MyApp	759	739	97.4%	20	2.6%
Total		1572	1534	97.6%	38	2.4%

Table 55 Answers by Question Type and Data Capture Method

Question Type	Method	Total	Correct	Correct %	Error	Error %
Date	MyApp	0	0		0	
	Paper	115	111	96.5%	4	3.5%
Integer	MyApp	58	57	98.3%	1	3.4%
	Paper	46	44	95.7%	2	4.3%
Real	MyApp	87	84	96.6%	3	3.4%
	Paper	69	67	97.1%	2	2.9%
Time	MyApp	87	83	95.4%	4	4.6%
	Paper	69	69	96.6%	0	0.0%
Text	MyApp	117	114	97.4%	3	2.6%
	Paper	92	90	97.8%	2	2.2%
Single Select	MyApp	464	457	98.5%	7	1.5%
	Paper	368	358	97.3%	10	2.7%
All Types	MyApp	813	795	97.8%	18	2.2%
	Paper	759	739	97.4%	20	2.6%
Total		1572	1534	97.6%	38	2.4%

C.4 Results for Paper vs. openXdata

Profiling openXdata results vs. the paper control give the following

Table 56 All Answers

Total Answers	Correct	Correct %	Error	Error %
1122	1049	93.5%	73	6.5%

Table 57 Answers by Level of Education

	Total	Correct	Correct %	Error	Error %
Grade12	924	860	93.1%	64	6.9%
Grade10 or 11	198	189	95.5%	9	4.5%
All Grades	1122	1049	93.5%	73	6.5%

Table 58 Answers by Interview Language

	Total	Correct	Correct %	Error	Error %
Home Language	495	466	94.1%	29	5.9%
English	627	583	93.0%	44	7.0%
Both	1122	1049	93.5%	73	6.5%

Table 59 Answers by Computer Literacy

	Total	Correct	Correct %	Error	Error %
Literate	726	683	94.1%	43	5.9%
Not Literate	396	366	92.4%	30	7.6%
Both	1122	1049	93.5%	73	6.5%

Table 60 Answers by Question Type

	Total	Correct	Correct %	Error	Error %
date	170	152	89.4%	18	10.6%
integer	68	63	92.6%	5	7.4%
real	102	99	97.1%	3	2.9%
time	102	98	96.1%	4	3.9%
text	136	124	91.2%	12	8.8%
s-s	544	513	94.3%	31	5.7%
All Types	1122	1049	93.5%	73	6.5%

Table 61 Answers by Method of Capture

	Total	Correct	Correct %	Error	Error %
openXdata	363	310	85.4%	53	14.6%
Paper	759	739	97.4%	20	2.6%
All Methods	1122	1049	93.5%	73	6.5%

Table 62 Answers by Level of Education and Data Capture Method

Education	Method	Total	Correct	Correct %	Error	Error %
Grade12	Paper	594	578	97.3%	16	2.7%
	openXdata	330	282	85.5%	48	14.5%
Grade10 or 11	Paper	165	161	97.6%	4	2.4%
	openXdata	33	28	84.8%	5	15.2%
All Grades	Paper	759	739	97.4%	20	2.6%
	openXdata	363	310	85.4%	53	14.6%
Total		1122	1049	93.5%	73	6.5%

Table 63 Answers by Computer Literacy and Data Capture Method

Computer Literate	Method	Total	Correct	Correct %	Error	Error %
Yes	Paper	462	452	97.8%	10	2.2%
	openXdata	264	231	87.5%	33	12.5%
No	Paper	297	287	96.6%	10	3.4%
	openXdata	99	79	79.8%	20	20.2%
Both	Paper	759	739	97.4%	20	2.6%
	openXdata	363	310	85.4%	53	14.6%
Total		1122	1049	93.5%	73	6.5%

Table 64 Answers by Interview Language and Data Capture Method

Home Language	Method	Total	Correct	Correct %	Error	Error %
Home Language	Paper	330	323	97.9%	7	2.1%
	openXdata	165	143	86.7%	22	13.3%
English	Paper	429	416	97.0%	13	3.0%
	openXdata	198	167	84.3%	31	15.7%
Both Languages	Paper	759	739	97.4%	20	2.6%
	openXdata	363	310	85.4%	53	14.6%
Total		1122	1049	93.5%	73	6.5%

Table 65 Answers by Question Type and Data Capture Method

Question Type	Method	Total	Correct	Correct %	Error	Error %
Date	openXdata	55	41	74.5%	14	25.5%
	Paper	115	111	96.5%	4	3.5%
Integer	openXdata	22	19	86.4%	3	3.0%
	Paper	46	44	95.7%	2	4.3%
Real	openXdata	33	32	97.0%	1	3.0%
	Paper	69	67	97.1%	2	2.9%

Time	openXdata	33	29	87.9%	4	12.1%
	Paper	69	69	97.0%	0	0.0%
Text	openXdata	44	34	77.3%	10	22.7%
	Paper	92	90	97.8%	2	2.2%
Single Select	openXdata	176	155	88.1%	21	11.9%
	Paper	368	358	97.3%	10	2.7%
All Types	openXdata	363	310	85.4%	53	14.6%
	Paper	759	739	97.4%	20	2.6%
Total		1122	1049	93.5%	73	6.5%

C.5 Results for Paper vs. openXdata excluding Repeating Groups

Given the problems with the repeating groups, these questions were removed to give the following results.

Table 66 All Answers

Total Answers	Correct	Correct %	Error	Error %
986	930	94.3%	56	5.7%

Table 67 Answers by Method of Capture

	Total	Correct	Correct %	Error	Error %
openXdata	319	281	88.1%	38	11.9%
Paper	667	649	97.3%	18	2.7%
All Methods	986	930	94.3%	56	5.7%

Table 68 Answers by Level of Education and Data Capture Method

Education	Method	Total	Correct	Correct %	Error	Error %
Grade12	Paper	522	508	97.3%	14	2.7%
	openXdata	290	254	87.6%	36	12.4%
Grade10 or 11	Paper	145	141	97.2%	4	2.8%
	openXdata	29	27	93.1%	2	6.9%
All Grades	Paper	667	649	97.3%	18	2.7%
	openXdata	319	281	88.1%	38	11.9%
Total		986	930	94.3%	56	5.7%

Table 69 Answers by Computer Literacy and Data Capture Method

Computer Literate	Method	Total	Correct	Correct %	Error	Error %
Yes	Paper	406	397	97.8%	9	2.2%
	openXdata	232	208	89.7%	24	10.3%

No	Paper	261	252	96.6%	9	3.4%
	openXdata	87	73	83.9%	14	16.1%
Both	Paper	667	649	97.3%	18	2.7%
	openXdata	319	281	88.1%	38	11.9%
Total		986	930	94.3%	56	5.7%

Table 70 Answers by Interview Language and Data Capture Method

Home Language	Method	Total	Correct	Correct %	Error	Error %
Home Language	Paper	290	284	97.9%	6	2.1%
	openXdata	145	129	89.0%	16	11.0%
English	Paper	377	365	96.8%	12	3.2%
	openXdata	174	152	87.4%	22	12.6%
Both Languages	Paper	667	649	97.3%	18	2.7%
	openXdata	319	281	88.1%	38	11.9%
Total		986	930	94.3%	56	5.7%

C.6 Results for openXdata including and excluding Repeating Groups

Comparing results including Repeating Groups and excluding Repeating Groups gives the following results.

Table 71 Answers including and excluding repeating groups

	Total	Correct	Correct %	Error	Error %
excl rpt grp	319	281	88.1%	38	11.9%
incl rpt grp	363	310	85.4%	53	14.6%

Table 72 Answers including and excluding repeating groups by Computer Literacy

Computer Literate	Repeating Grp	Total	Correct	Correct %	Error	Error %
yes	excl rpt grp	232	208	89.7%	24	10.3%
	incl rpt grp	264	231	87.5%	33	12.5%
no	excl rpt grp	87	73	83.9%	14	16.1%
	incl rpt grp	99	79	79.8%	20	20.2%
both	excl rpt grp	319	281	88.1%	38	11.9%
	incl rpt grp	363	310	85.4%	53	14.6%