Objective: Handheld computers (personal digital assistant, PDA) have the potential to reduce the logistic burden, cost, and error rate of paper-based health research data collection, but there is a lack of appropriate software. The present work describes the development and evaluation of PDACT, a Personal Data Collection Toolkit (www.healthware.org/pdact/index.htm) for the Palm™ Pilot handheld computer for interviewer-administered and respondent-administered data collection.

Methods: We developed Personal Data Collection Toolkit (PDACT) software to enable questionnaires developed in QDSTM Design Studio, a Windows™ application, to be compiled and completed on Palm™ Pilot devices and evaluated in several representative field survey settings.

Results: The software has been used in seven separate studies and in over 90,000 interviews. Five interviewer-administered studies were completed in rural settings with poor communications infrastructure, following one day of interviewer training. Two respondent-administered questionnaire studies were completed by learners, in urban secondary schools, after 15 min training.

Questionnaires were available on each handheld in up to 11 languages, ranged from 20 to 580 questions, and took between 15 and 90 min to complete. Up to 200 Palm™ Pilot devices were in use on a single day and, in about 50 device-years of use, very few technical problems were found. Compared with paper-based collection, data validation and cleaning times were reduced, and fewer errors were found.

PDA data collection is easy to use and preferred by interviewers and respondents (both respondent-administered and interviewer-administered) over paper. Data are compiled and available within hours of collection facilitating data quality assurance. Although hardware increases the setup cost of the first study, the cumulative cost falls thereafter,
1. **Introduction**

Field data collection and surveys are a fundamental requirement for public health, health services and intervention evaluation research. This is especially true in low and middle income countries where collection and analysis of administrative data are rare and postal surveys are limited by poor infrastructure and variable literacy levels. Most researchers use paper questionnaires followed by manual input into a computer database application for data collation and statistical analysis. This is time-consuming, error-prone and expensive, and acts as a barrier to increasing the volume and speed of needed research [1]. Direct digital data entry into a computer software application at the point of data collection could help solve these problems [2].

In principle, digital data collection has many advantages. Software applications could circumvent input errors through data type and range checking at the time of capture and by use of selection lists which restrict entry data to predefined options and thus prevent typing and data format errors. Software applications could also manage the user’s path through complex conditional logic and skip patterns, hiding irrelevant detail. Once a completed question, or an entire questionnaire has been committed to the computer, it need not be accessible without authorized access, improving privacy of sensitive data. Digital data capture can make data available rapidly enough to allow for error detection at the data collation point with immediate feedback to field-workers for correction. Encrypted digital data can be stored more easily and more securely than unencrypted paper forms.

Health data collection applications have been used in desktop and laptop computers, and in standalone, networked and web environments [3] but these technologies typically are expensive, fragile and not suitable for field use far from power supplies. Handheld computers (personal digital assistants, PDAs) are becoming attractive for field data collection in developing countries and resource-poor environments as they are pocket-sized and portable, incorporate all the advantages of electronic data collection and have reached the level of robustness, battery life, processing power and data storage capacity needed for large-scale survey research [4]. Although specially programmed PDAs have been used in developing countries for data collection ([5,6], www.satellife.org), there is still a lack of low-cost reusable, general purpose, user configurable software for this purpose.

1.1. **Aim**

We report on the development of a user configurable software survey tool for the Palm™ Pilot handheld computer and its field use for interviewer- and respondent-administered surveys in South Africa.

2. **Methods**

2.1. **Software design and development**

The Personal Data Collection Toolset (PDACT) was developed as an extension of the Questionnaire Development System (QDSTM; NOVA Research Company) and Computer Assisted Personal Interview (CAPI) application, operable on Palm™ Pilot PDAs. The PDACT system comprises three separate applications: (i) a Windows application that generates a questionnaire-specific PDACT control file from a compiled QDSTM questionnaire specification file; (ii) a runtime application that runs on the Palm™ operating system and inputs the PDACT control file to generate questionnaire-specific screens and data structures, and; (iii) a Palm™ conduit that enables data transfer between the Palm™ device and a Windows computer using the native Palm™ digital data and executable file transfer and synchronization (“hot-synching”) process.

Several utility programs to compress and extract data files were also programmed, including a compression algorithm that reduces individual binary data files to a few kilobytes before sending. PDACT questionnaires were implemented on the Palm™ Vx, 501, Tungsten 3/W/E as well as on the Handspring Visor and Palm™ Treo. Geo-positioning system information was recorded on a Magellan clip-on GPS device for the Handspring Visor. **Questionnaire Development and Data Collection**

Questionnaires are developed using the QDSTM Design Studio application and downloaded to the PDA. Individual questions are presented one at a time (Fig. 1), moving forward to the next screen on completion by entering data or tapping a ‘NEXT’ button which appears after a validated data type is entered.

For interviewer-administered operation field workers were given one day of training and ongoing support. For respondent-administered questionnaires in field settings, the interviewees were given a 15 min introduction to the Palm™ and PDACT before completing the questionnaires.

2.2. **Data uploading and collation**

Data was uploaded by direct serial or infrared connection to a standalone PC or server or, by using a clip-on modem, via standard telephone lines using toll-free numbers.

Uploaded data is decompressed, decrypted and then compiled using either the QDSTM Data Warehouse application or a desktop database or spreadsheet application (Microsoft Access or Excel) and exported to statistical applications (e.g., SAS and SPSS) for analysis.

2.3. **Evaluation of PDACT**

In Cape Town, during 2003 the SATZ pilot study was conducted among Grade 8 students in one public school to evaluate the
handheld application versus paper data collection in preparation for a cluster-randomized controlled trial of a sexual and reproductive health programme [8]. All students in the grade (N = approximately 200; average age, 14 years) were randomly allocated to complete either an electronic or a paper version of the identical sexual behaviour questionnaire.

Three weeks later, they completed questionnaire again, each student using the same method to which they had been randomly allocated. Questions about student opinions of the respective data collection methods and their preferences were added to the questionnaire.

Researchers trained the students in the group allocated to the electronic survey to use the handheld devices with a short demonstration questionnaire. Once students were comfortable using the device, they individually completed the questionnaire. Students were seated far enough apart to ensure that they could not view one other’s responses. The researchers were present throughout the survey to attend to questions.

Cost was estimated in one respondent-administered study [8] and excluded the cost of developing the PDACT software. Other software was priced at normal university purchase price, including academic discounts.

3. Results

3.1. Studies

To date, seven studies have been carried out using the PDACT system (Table 1).

3.2. Software

Development took approximately two programmer years. The application operates on versions of the Palm™ operating system up to and including version 3.5. The present version of the application is stable and bug-free after extensive field-testing.
### Table 1 – List of studies in which PDACT was used for data collection.

<table>
<thead>
<tr>
<th>Project</th>
<th>Description</th>
<th>Admin Type</th>
<th>Data Collection</th>
<th>Number of PDAs</th>
<th>Records completed</th>
</tr>
</thead>
<tbody>
<tr>
<td>PALSA (Practical Approach to Lung Health in South Africa)</td>
<td>Cluster randomized trial to evaluate the effectiveness and efficiency of a training intervention to promote quality respiratory care in the Free State province of South Africa</td>
<td>Interviewer</td>
<td>Remote upload to Cape Town</td>
<td>10</td>
<td>3499</td>
</tr>
<tr>
<td>Free State ARV Treatment Rollout and PALSA Plus (Practical Approach to Lung Health in South Africa Plus)</td>
<td>Operational clinical data collection attached to the rollout of antiretroviral therapy at specialist AIDS clinics in the Free State province of South Africa and survey of a training intervention for the rollout of antiretroviral therapy in the Free State province of South Africa</td>
<td>Interviewer</td>
<td>Remote upload to Cape Town</td>
<td>20</td>
<td>75000</td>
</tr>
<tr>
<td>SATZ Study (8)</td>
<td>Cluster-RCT to evaluate the effects of a teacher-driven HIV prevention programme on the sexual behaviour of grade eight students in Cape Town</td>
<td>Respondent</td>
<td>Local upload to PC</td>
<td>50</td>
<td>5646</td>
</tr>
<tr>
<td>HealthWise Study (9)</td>
<td>A cluster-RCT to evaluate the effects of the HealthWise high school curriculum on the HIV/AIDS risk behaviour, substance abuse and the use of leisure time</td>
<td>Respondent</td>
<td>Local upload to PC</td>
<td>100</td>
<td>4200</td>
</tr>
<tr>
<td>TB Defaulter study</td>
<td>Survey of reasons for treatment interruption among TB patients</td>
<td>Interviewer</td>
<td>Remote upload to Cape Town</td>
<td>10</td>
<td>664</td>
</tr>
<tr>
<td>Survey of Malaria Attitudes</td>
<td>Survey of attitudes to malaria prophylaxis among tourists</td>
<td>Interviewer</td>
<td>Remote upload to Cape Town with GPS coordinates</td>
<td>1</td>
<td>25</td>
</tr>
<tr>
<td>Integration Study (10)</td>
<td>Survey of substance use and sexual risk behaviour of patients presenting at primary health care services in Cape Town.</td>
<td>Interviewer</td>
<td>Local upload to PC</td>
<td>4</td>
<td>2608</td>
</tr>
</tbody>
</table>

### Table 2 – Student reports on the validity of their responses to the questions about sexual behaviour: comparison between electronic and paper questionnaires (all p values were <0.01).

<table>
<thead>
<tr>
<th>Question</th>
<th>Student's allocation</th>
<th>Student responses</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Electronic questionnaire</td>
<td>Paper questionnaire</td>
</tr>
<tr>
<td>Which method would you prefer for answering questions about sex?</td>
<td>Palm pilot group</td>
<td>62 (85%)</td>
</tr>
<tr>
<td></td>
<td>Paper group</td>
<td>39 (53%)</td>
</tr>
<tr>
<td>When answering questions about sex, which method do you think is more confidential?</td>
<td>Palm Pilot group</td>
<td>65 (89%)</td>
</tr>
<tr>
<td></td>
<td>Paper group</td>
<td>40 (55%)</td>
</tr>
<tr>
<td>Which method makes it easier to give truthful answers to questions about sex?</td>
<td>Palm Pilot group</td>
<td>66 (88%)</td>
</tr>
<tr>
<td></td>
<td>Paper group</td>
<td>24 (32%)</td>
</tr>
</tbody>
</table>
### Table 3 – Inter-item correlation and test-retest reliability of psychosocial variables-SATZ pilot study.

<table>
<thead>
<tr>
<th>Scale</th>
<th>No. of items</th>
<th>Sample item</th>
<th>Response format</th>
<th>Cronbach’s alpha</th>
<th>Spearman test–retest correlation coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Paper questionnaire</td>
<td>Palm pilot questionnaire</td>
</tr>
<tr>
<td>Knowledge of HIV/AIDS</td>
<td>13</td>
<td>Is there a cure for HIV/AIDS?</td>
<td>2 point scale: 1 = correct; 0 = incorrect</td>
<td>0.63</td>
<td>0.54</td>
</tr>
<tr>
<td>Attitudes towards sexual behaviour</td>
<td>6</td>
<td>If I had a boyfriend and he refused to have sex with me I will think that he does not love me</td>
<td>5-point scale: strongly agree-strongly disagree</td>
<td>0.59</td>
<td>0.66</td>
</tr>
<tr>
<td>Attitudes towards condom use</td>
<td>8</td>
<td>If my partner used a condom during sex, I would feel less pleasure</td>
<td>5-point scale: strongly agree-strongly disagree</td>
<td>0.6</td>
<td>0.68</td>
</tr>
<tr>
<td>Attitudes towards condom use</td>
<td>7</td>
<td>It is okay for girls my age to carry condoms if they plan to have sex</td>
<td>5-point scale: strongly agree-strongly disagree</td>
<td>0.77</td>
<td>0.7</td>
</tr>
<tr>
<td>Perceived risk of pregnancy, STI &amp; HIV infection</td>
<td>9</td>
<td>To what extent do you think HIV is a threat against your personal health?</td>
<td>5-point scale: strongly agree-strongly disagree</td>
<td>0.74</td>
<td>0.84</td>
</tr>
<tr>
<td>Behavioural norms regarding abstinence</td>
<td>7</td>
<td>Most of my friends think I should not have sexual intercourse</td>
<td>5-point scale: strongly agree-strongly disagree</td>
<td>0.65</td>
<td>0.73</td>
</tr>
<tr>
<td>Sexual behavioural norms</td>
<td>2</td>
<td>In your school, about how many learners do you think have had sex?</td>
<td>6 point scale: none–all of them</td>
<td>0.72</td>
<td>0.75</td>
</tr>
<tr>
<td>Social norms about condoms with reference to friends and parents</td>
<td>8</td>
<td>Most of my friends think that I should use a condom when having sex</td>
<td>5-point scale: strongly agree-strongly disagree</td>
<td>0.75</td>
<td>0.8</td>
</tr>
<tr>
<td>Self efficacy to abstain from sex</td>
<td>11</td>
<td>I would be able to refuse to have sex with my boyfriend if I did not feel like having sex</td>
<td>5-point scale: strongly agree-strongly disagree</td>
<td>0.8</td>
<td>0.87</td>
</tr>
<tr>
<td>Self efficacy regarding safe sex when under the influence of alcohol or marijuana</td>
<td>5</td>
<td>I would be able to refuse to have sexual intercourse even if I have drank alcohol</td>
<td>5-point scale: strongly agree-strongly disagree</td>
<td>0.8</td>
<td>0.81</td>
</tr>
<tr>
<td>Self efficacy to use a condom</td>
<td>10</td>
<td>I would be able to refuse to have sex if my partner did not want to use a condom</td>
<td>5 point</td>
<td>0.81</td>
<td>0.88</td>
</tr>
<tr>
<td>Self efficacy to use condoms when under the influence of alcohol or marijuana</td>
<td>3</td>
<td>I would be able to make sure that my partner uses a condom even if I have used dagga</td>
<td>5-point scale: strongly agree-strongly disagree</td>
<td>0.8</td>
<td>0.69</td>
</tr>
<tr>
<td>Self efficacy about obtaining condoms</td>
<td>3</td>
<td>I would be able to go to the clinic to fetch condoms</td>
<td>5-point scale: strongly agree-strongly disagree</td>
<td>0.81</td>
<td>0.82</td>
</tr>
</tbody>
</table>
and use on the Palm™ Tungsten W that supports a full working day of battery life between recharges. The application was programmed to use files generated only with an older version 2.1 of QDSTM. Compatibility with later versions would require reprogramming.

3.3. User opinions

Table 2 presents the opinions and preferences of the students who participated in the second survey of the SATZ pilot study and provided responses to the items in question. (There were students who were not at school on the day of the survey, or who did not present at the classroom in which the survey was administered and this resulted in attrition.) For each question, the responses of the students allocated to the electronic questionnaire were compared with the responses of the students allocated to the paper questionnaire, using a chi square statistic. Most students in both groups preferred to use electronic questionnaires when answering questions about sex, and thought that electronic questionnaires were more confidential and made it easier to provide truthful answers.

3.4. Data quality

3.4.1. Data uploading

We have uploaded more than 90,000 interview records using PDACT and have lost only two records in the process.

3.4.2. Data reliability

Table 3 compares the inter-item correlation and test–retest reliability of scaled psychosocial variables from the SATZ pilot study. The electronic questionnaire resulted in a higher Cronbach’s alpha statistic for 10 of the 13 scales, suggesting that the internal consistency was superior for this method. Electronic and paper questionnaires performed similarly in terms of test–retest reliability.

Table 4 compares the test–retest reliability of the self-reported adolescent risk behaviours from the SATZ pilot study. For both electronic and paper versions of the questionnaire, test–retest reliability on sexual risk behaviour questions was similar.

3.4.3. Cost comparison

A comparison of setup and running costs for the paper and handheld data collection systems for the SATZ study[8] is presented in Tables 5a and 5b. We analyzed the cost of the SATZ study involving 6000 participants completing a 20-page survey on paper questionnaires compared with a PDA-equivalent and then calculated the respective costs of completing identical follow-up surveys biannually for two years. Two licenses of QDS™ Design Studio and CAPI have been used by the two research groups since inception and, with few exceptions, we have used the same, relatively expensive, Palm™ Tungsten W devices on all the studies.

4. Discussion

Handheld data collection is suitable for a variety of research surveys in a developing country and preferable
Table 5a – Cost comparison of handheld and paper-based data collection—SATZ Study.

<table>
<thead>
<tr>
<th>Item</th>
<th>Study 1</th>
<th>Study 2</th>
<th>Study 3</th>
<th>Study 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Partial Running Costs (PDA)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Palm Tungsten W</td>
<td>4,000</td>
<td>615</td>
<td>100</td>
<td>15385</td>
</tr>
<tr>
<td>Licensing—QDSTM design studio</td>
<td>299</td>
<td>2</td>
<td>75</td>
<td>75</td>
</tr>
<tr>
<td>Licensing QDSTM CAPI</td>
<td>499</td>
<td>2</td>
<td>125</td>
<td>125</td>
</tr>
<tr>
<td>Maintenance (10% of capital)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3077</td>
<td>3077</td>
<td>3169</td>
<td>3169</td>
</tr>
<tr>
<td>Total</td>
<td>18,661</td>
<td>18,661</td>
<td>18,753</td>
<td>18,753</td>
</tr>
<tr>
<td>Partial Running Costs (Paper)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Printing</td>
<td>25,527</td>
<td>3927</td>
<td>1</td>
<td>3,927</td>
</tr>
<tr>
<td>Double data entry</td>
<td>96,000</td>
<td>14769</td>
<td>1</td>
<td>14,769</td>
</tr>
<tr>
<td>Total</td>
<td>18,696</td>
<td>18,661</td>
<td>19,257</td>
<td>19,835</td>
</tr>
</tbody>
</table>

Calculations: Total number of subjects per study is 6000. The cost of the PDA hardware and software was allocated to four studies over two years. The maintenance cost was calculated at 10% of capital cost per annum. CPIX = 3%. 1 US$ = 6.5 ZAR.

to paper data collection according to a number of criteria (Table 6).

We chose the Palm™ Pilot, running the Palm operating system (OS), over the Pocket PC (Compaq), running the Windows ME OS because of the superior battery life, power-efficient monochrome screens, and its smaller size. At the time we did this work, QDSTM software was not available for Windows ME devices.

We still experienced some difficulties with battery life, especially on the less expensive Palm™ Tungsten E device, and this issue should be taken into account in hardware choice.

Designing a questionnaire in QDSTM is intuitive, even for novice users implementing skips and conditional logic. The one question per screen design of QDSTM implemented in PDFACT frees the investigator from a dependency on programmers and screen design allowing low-cost, rapid development of questionnaires and was highly valued by investigators in our studies.

The PDFACT system has been a stable application in field setting. All of the expected advantages for electronic data collection have been manifested by our handheld data collection system. The finding that very few devices were damaged or stolen was an unexpected benefit as we expected that this would be a serious limitation of handheld data collection. In view of the low rate of hardware failures, the system is easily managed by having one or two devices on standby and using a courier service to replace field units at remote sites within a day or two.

From the outset, data uploading was seen to be a major potential point of failure in the system. Although some minor problems were experienced where telephone lines are particularly poor or switchboards make dialing out difficult, we found the uploading to be a very robust and reliable technology. We ascribe this to two factors, the inherent robustness of the Palm hot-synching technology and the fact that completed PDFACT questionnaires are uploaded one at a time as small compressed Palm™ database files of a few kilobytes each, increasing the chances of a file being sent, even if multiple dial-ups need to take place.

The two questionnaires lost during uploading were most likely due to unplanned termination of the call at a switchboard, or to electrical interference across long distance rural telephone lines. Also, the application does not inform the user when a data upload has been unsuccessful but does leave untransmitted or incompletely transmitted records on the PDA. For a period early in the use of the device, we mistakenly assumed that the file had been delivered, and then reprogrammed the device resulting in loss of hundreds of records still resident in the device memory. Paper forms were available at this stage, and so no data was actually lost, but in any case the problem was overcome by encouraging routine checking for retained files prior to reprogramming.

In all studies, both interviewer- and respondent-administered, field workers, interviewers and respondents alike preferred handheld data collection. The PDA was equally acceptable to computer literate and non-literate users, younger and older field workers. This is probably related to the ease of use in comparison with paper and substantial use of pick lists, which seems to mitigate the more cumbersome text entry using a stylus (Fig. 2).
This finding was confirmed in the PALSA study [7], where field workers reported that patients were more willing to disclose sensitive information regarding their health and socioeconomic status during PDA interviews than in paper-based interviews. Patients felt reassured that sensitive information, captured onto the PDA, could not be viewed by clinic staff at the health facility where they were receiving care, and where the interviews were held.

User satisfaction in the PALSA study was also attributed to the automation of skip instructions. The questionnaire for this study catered for collection of economic data, amounting to 60% of the questions, and half of the paper questionnaire’s 35 pages. This yields a bulky paper questionnaire, with a complex path which increased the risk of error. The PDA automated all skip instructions and only displayed relevant items to the interviewer, halving the interview time and reducing the actual (just 12% of all economic questions needed to be answered) and perceived burden of the interview.

Although the cost of the Palm solution is initially more expensive than the paper-based solution for the first study, our results show that the break-even point is manifest fairly early and, thereafter, for studies using the same devices, handheld data collection is less expensive than paper data collection. This will be particularly useful for institutions doing many such studies. In addition, the cost saving that could result from early error detection preventing a study being ruined could far outweigh the cost differential between PDA and paper.

PDA data collection has two other advantages, whose frequency may be low, but whose impact is so high that it may well tip the balance for some researchers. Due to frequent uploading, catastrophic data loss is unlikely, whereas paper data requires more arduous procedures like frequent physical collection to guarantee its safety. Frequent uploading (and early quality checking) allows detection of systemic data collection error, be it due to accident or design.

Upfront costs are obviously higher for the handheld data collection system, but running costs are lower. The cost of the PDA survey is slightly less than paper when the cost of the hardware is annualized over four studies and the programming cost excluded (Table 5a). When the original programming cost of PDACT itself is included, the upfront costs need to be discounted over eight studies to obtain a comparative cost with paper (Table 5b). The number of studies required for a return on investment is likely to be lower in other countries where the upfront PDA costs are significantly lower than in South Africa.

Many organizations are developing cell-phone solutions for field data collection in Africa because of the ubiquity, low cost and existing infrastructure associated with this technology and the perceived high cost of PDAs. Our PDACT application design with one question per screen lends itself to implementation on cell phones and we are already using PDACT on the Palm Treo smartphone. The combination of PDA power and cell phone convenience will likely prove a powerful combination for future applications, particularly in developing countries where cellular infrastructure is well-developed and paper-based alternatives are difficult to administer.

From our field experiences [7,8] and that of other colleagues using the tool [9,10], we conclude that this novel and
<table>
<thead>
<tr>
<th>Criterion</th>
<th>Handheld device</th>
<th>Paper-based</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Questionnaire development and formatting</strong></td>
<td>Benefits 1. Intuitive and simple to design and reuse questionnaire specifications with QDS 2. Questionnaire layout and logic is automated and simple to develop with the one-screen-at-a-time design Limitations 1. Cannot include graphical elements (but can in the Pocket PC version sold by NOVA)</td>
<td>1. Familiar and widely-used (intuitive) with significant infrastructure for supporting development of paper-questionnaires 2. Simple to include graphics 1. Questionnaire development is familiar but laborious and time-consuming 2. Questionnaire development needs to be done manually with limited support from word-processing and desktop publishing software</td>
</tr>
<tr>
<td><strong>Data collection</strong></td>
<td>Benefits 1. Form completion is simple and leaves a clear audit trail 2. Multiple questionnaires and languages are readily included on one device 3. May be more usable in inclement weather (using a suitable covering) Limitations 1. Limitations in terms of power requirements</td>
<td>1. Familiar to most interviewees and interviewers 2. No additional requirements to use (e.g. power, upload server) 1. Questionnaire may be bulky and difficult to navigate 2. Easily lost</td>
</tr>
<tr>
<td><strong>Data quality</strong></td>
<td>Benefits 1. Select lists, compulsory fields and range-checking are readily-included and contribute significantly to improved accuracy 2. Skip logic automates navigation to relevant questions 3. Audit trails can be included to track changes 4. Effective error correction at time and source of data capture Limitations 1. Lack of familiarity can result in input errors 2. Ease of input may facilitate garbage input</td>
<td>1. Familiar to most interviewees and interviewers 2. Can be easier to include ad hoc notes 1. Limited ways to reduce input errors 2. No automated ways of detecting and preventing data input errors</td>
</tr>
<tr>
<td><strong>User preferences</strong></td>
<td>Benefits 1. Preferred by most interviewers and interviewees, alike Limitations 1. Can be more difficult to use for less computer-literate users</td>
<td>1. Familiar to most interviewees and interviewers 1. Less confidential</td>
</tr>
<tr>
<td><strong>Data collation</strong></td>
<td>Benefits 1. Data collation is reliable and can be rapid enough to allow early correction of errors and omissions in the data Limitations 1. Need to check for residual data left on devices and, periodically, for transmission failures 2. Increased IT infrastructure, expertise and support is required</td>
<td>1. Familiar and easy to set up 1. Data is not collated in real time and turnaround time is high 2. The logistic, transport and data collation burden is high 3. It is difficult to maintain control of the questionnaires 4. Significant burden of double data capture</td>
</tr>
</tbody>
</table>
Table 6 (Continued)

<table>
<thead>
<tr>
<th>Criterion</th>
<th>Handheld device</th>
<th>Paper-based</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Data reliability</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Benefits</td>
<td>1. For the studies cited here PDA and paper-data collection were equally reliable</td>
<td>1. Paper is inherently more reliable</td>
</tr>
<tr>
<td>Limitations</td>
<td>1. Electronic devices are inherently less reliable</td>
<td>1. Anecdotal accounts of significant loss of clinical and survey data</td>
</tr>
<tr>
<td></td>
<td>2. In some cases, there may be a theft or loss issue</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3. Need to ensure adequate backup system to guard against data loss</td>
<td></td>
</tr>
<tr>
<td><strong>Cost</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Benefits</td>
<td>1. Costs can be recovered within a fairly small number of studies</td>
<td>1. Startup costs are low</td>
</tr>
<tr>
<td>Limitations</td>
<td>1. There is a significant startup cost, particularly for studies requiring a large number of field workers</td>
<td>2. Little or no barriers to entry</td>
</tr>
<tr>
<td></td>
<td>2. IT infrastructure costs are higher but may be lessened by the presence of existing infrastructure and expertise</td>
<td>1. Costs will accumulate over time with little hope of economies of scale</td>
</tr>
</tbody>
</table>

Summary points

“What was already known before our study”

- Handheld computers (PDAs) had been shown to be useful for field data collection but the programming of software for PDA data collection required special programming skills and was not generally available to investigators without access to programming resources.
- It was thought that PDA technology was not sufficiently robust and reliable for large-scale field data collection, particularly in rural settings.
- It was thought that PDA surveys may require specialized background or training for general use which would be difficult to find in resource-constrained settings.
- The cost of using PDAs for surveys was either not known or thought to be prohibitively expensive.

“What did our study add to our body of knowledge”

- Software can be developed that permits end-users with no particular computer programming skill or knowledge to develop PDA-based survey applications and structured medical record form applications.
- PDAs and remote telephone-based data collection can be deployed reliably in the field for extended periods.
- PDA surveys are readily and reliably used by novice users with no special background and little formal training and increase the convenience and reliability of data collected.
- The capital cost of PDA technology can be recovered after a relatively low number of studies using the same technology.

apparently sophisticated technology designed for time-pressed senior executives in developed countries, is an affordable, robust, and extremely useful method for data collection in lower and middle income countries.

5. Conclusion

From our field experiences, we conclude that this novel and apparently sophisticated technology designed for time-pressed senior executives in developed countries, is an affordable, robust, and extremely useful method for data collection in lower and middle income countries.

Acknowledgements

The development of PDACT was funded by an Innovation Fund grant from the South African Department of Science and Technology and the South African Medical Research Council. The PALSA and PALSA Plus studies are funded by the Governance Equity and Health (GEH) programme of the International Development Research Centre (IDRC), Canada. The Free State Antiretroviral Treatment data collection is funded by the ACACIA and Connectivity Africa programs of the IDRC. NOVA Research Company provided free QDS™ software to the Palm™ PDACT development team and heavily discounted software to South African researchers using the application. Some of the data reported in this paper are derived from a larger project titled: “Promoting sexual and reproductive health. School-based HIV/AIDS prevention in Sub-Saharan Africa”. The acronym for this larger project is “SATZ”, which stands for South Africa Tanzania. The SATZ study is funded by the EC - INCO research programme (Fifth Framework Programme—Contract number ICA4-CT-2002-10038). The partners and principal investigators include: University of Cape Town (Alan J. Flisher), Muhimbili University College of Health Sciences, (Sylvia Kaaya), University of
the North (Hans Onya), Karolinska Institute (Minou Fugle-
sang), University of Maastricht (Herman Schaalma), University
of Oslo (Knut-Inge Klepp), World Population Foundation and
Youth Incentives (Jo Reinders), University of Bergen (Leif
Edvard Aarø – coordinator). None of the funders had any
involvement in the work presented here or in the decision to
publish.

Contributions.

Authors: Merrick Zwarenstein originated the idea of PDACT,
contributed to the design, development and implementation
of the application, to the design and implementation of the
cited PALSAR, PALSAR Plus and Free State Antiretroviral Treatment
rollout studies and to the writing of this manuscript.
Chris Seebregts designed the software, managed its develop-
ment and implementation in most of the studies and
contributed to the writing of the manuscript. Lara Fairall
contributed to the design and implementation of the applica-
tion, conducted the cited PALSAR, PALSAR Plus and Free State
Antiretroviral Treatment rollout studies and contributed to the
writing of the manuscript. Catherine Mathews conducted the
 cited SATZ studies and contributed results and analyses and
to the writing of the manuscript. Alan Flisher supervised the
SATZ and Integration studies and contributed to the writing of
the manuscript. Clive Seebregts developed all the PDACT and
associated software and contributed to the implementation
of most of the studies cited in the manuscript. Wanjiru
Mukoma contributed results and analyses from the SATZ study.
Knut-Inge Klepp contributed results and analyses from the SATZ
study.

PDACT user group: Venessa Timmerman, Michael Ndloko-
vane and Mpumelelo Msimanga contributed technically to
the implementation of many of the studies cited in the
manuscript. Catherine Ward contributed experiences using
PDACT on several studies. Lisa Wegener contributed experi-
ences using PDACT on the cited HealthWise study. Heather
Jaspan and Linda-Gail Bekker contributed experiences using
PDACT on the studies in Masiphumelela. Karin Weyer con-
tributed experiences using PDACT on the TB defaulter studies.
Burnie Naun, Ischen Seocharan, Natasha Morris, Francois
Maartens and Brian Sharp contributed to the design and
field testing of the GIS module. Paul Young and Susan Sheck
assisted with understanding of QDS™ and the software
design.

REFERENCES

[1] N. Siegfried, M. Clarke, J. Volmink, Randomised controlled
trials in Africa of HIV and AIDS: descriptive study and
[2] S.M. Powsner, J.C. Wyatt, P. Wright, Opportunities for and
challenges of computerization, Lancet 352 (November)
computerized field data collection system for health
collection in developing countries, Health Pol. Plann. 7 (1992)
67–71.
download/pdapaper1.pdf.
use of handheld computers for healthcare in Ghana,
Uganda, and Kenya, 28 February 2003,
http://www.bridges.org/satellife/.
Lombard, B. Majara, G. Joubert, R.G. English, A. Bheekie,
H.C.J. van Rensburg, P. Mayers, A.C. Peters, R.D. Chapman,
Educational outreach to nurses improves tuberculosis case
detection and primary care of respiratory illness: a
pragmatic cluster randomized controlled trial, BMJ 331
Flisher, C. Mathews, HealthWise South Africa: development
of a life skills curriculum for young adults, World Leisure J.
G.B. Distiller, C.M. Weisner, Substance abuse and HIV risk
behaviours in Cape Town primary health care service users,