Assessing the Usability of Six Data Entry Mobile Interfaces for Caregivers: A Randomized Trial

Frederic Ehrler¹, PhD; Guy Haller²,³, MD, PhD; Evelyne Sarrey⁵, RN; Magali Walesa⁶; Rolf Wipfli¹, PhD; Christian Lovis¹,⁷, MD, MPH

¹Division of Medical Information Sciences, Department of medical imaging and medical information sciences, University Hospitals of Geneva, Geneva, Switzerland
²Division of Anaesthesiology, University Hospitals of Geneva, Geneva, Switzerland
³Division of Clinical Epidemiology, University Hospitals of Geneva, Geneva, Switzerland
⁴Department of Epidemiology and Preventive Medicine, Monash University, Melbourne, Australia
⁵Direction of Nursing, University Hospitals of Geneva, Geneva, Switzerland
⁶School of Medicine, University of Geneva, Geneva, Switzerland
⁷Faculty of medicine, University of Geneva, Geneva, Switzerland

Corresponding Author:
Frederic Ehrler, PhD
Division of Medical Information Sciences
Department of medical imaging and medical information sciences
University Hospitals of Geneva
Rue Gabrielle-Perret-Gentil 4
Geneva, 1211
Switzerland
Phone: 41 223728697
Fax: 41 223728697
Email: frederic.ehrler@hcuge.ch

Abstract

Background: There is an increased demand in hospitals for tools, such as dedicated mobile device apps, that enable the recording of clinical information in an electronic format at the patient’s bedside. Although the human-machine interface design on mobile devices strongly influences the accuracy and effectiveness of data recording, there is still a lack of evidence as to which interface design offers the best guarantee for ease of use and quality of recording. Therefore, interfaces need to be assessed both for usability and reliability because recording errors can seriously impact the overall level of quality of the data and affect the care provided.

Objective: In this randomized crossover trial, we formally compared 6 handheld device interfaces for both speed of data entry and accuracy of recorded information. Three types of numerical data commonly recorded at the patient’s bedside were used to evaluate the interfaces.

Methods: In total, 150 health care professionals from the University Hospitals of Geneva volunteered to record a series of randomly generated data on each of the 6 interfaces provided on a smartphone. The interfaces were presented in a randomized order as part of fully automated data entry scenarios. During the data entry process, accuracy and effectiveness were automatically recorded by the software.

Results: Various types of errors occurred, which ranged from 0.7% for the most reliable design to 18.5% for the least reliable one. The length of time needed for data recording ranged from 2.81 sec to 14.68 sec, depending on the interface. The numeric keyboard interface delivered the best performance for pulse data entry with a mean time of 3.08 sec (SD 0.06) and an accuracy of 99.3%.

Conclusions: Our study highlights the critical impact the choice of an interface can have on the quality of recorded data. Selecting an interface should be driven less by the needs of specific end-user groups or the necessity to facilitate the developer’s task (eg, by opting for default solutions provided by commercial platforms) than by the level of speed and accuracy an interface can provide for recording information. An important effort must be made to properly validate mobile device interfaces intended for use in the clinical setting. In this regard, our study identified the numeric keyboard, among the proposed designs, as the most accurate.
interface for entering specific numerical values. This is an important step toward providing clearer guidelines on which interface to choose for the appropriate use of handheld device interfaces in the health care setting.

(JMIR Human Factors 2015;2(2):e15) doi: 10.2196/humanfactors.4093

KEYWORDS
data collection; mobile applications; computers, handheld; user-computer interface; vital signs; patient safety

Introduction

Electronic data collection and recording in the health care setting is performed increasingly at the patient’s bedside. Data (eg, medical prescriptions, medical summary reports, or daily recordings of body temperature, respiratory or cardiac frequency) can easily be collected on portable computers. Such devices have the advantage of being easy to store, manipulate, and use in emergency departments, outpatient clinics, or other crowded areas. Among the portable devices on offer, tablets and smartphones are becoming increasingly popular due to their handiness and resemblance to traditional paper-and-pencil data collection interfaces [1-3]. They also offer the advantage of providing apps designed especially for handheld devices, such as drug dosage calculators, electronic pharmacopeias, textbooks, or medical literature databases [1,4-7].

In the health care setting, many usability problems contribute to medical errors [8], of which those related to data entry are a major source. The quality of recorded information is of utmost importance because the life of a patient can easily be put at risk by the improper recording of a drug dosage or the incorrect labeling of a physiological or biological value [9-11]. Accuracy in the process of data recording can be significantly influenced by the design of an interface or by factors related to the type of data to be recorded, such as a number’s length, type, magnitude or frequency, and even font appearance [12]. This has already been demonstrated for specific entry devices, such as infusion pumps [12-14], and in the context of medical prescriptions [8]. The limited size and tactile interaction paradigm of handheld devices compared to traditional laptops has further emphasized the risk of increased errors in data entry [3].

The influence that specific characteristics of handheld devices can have on human-machine interactions has already been studied [15-27]. Earlier works investigated the impact of limited display size on users’ performances for tasks such as browsing, information retrieval, readability, or target selection [15-19]. These studies showed that the size of a mobile screen has no major impact on the user’s comprehension of the information displayed. However, a correlation was found between the ease of reading and the size of the screen. In this context, Duchnicky and Kolars [20] found that it takes users up to 25% longer to read a given text on a small-width display than on a regular desktop screen width. This was confirmed by another study by Resiel and Shneiderman [21] who reported that reading on a 22-line display compared with a 60-line display resulted in a 15% decrease in the speed of reading. The same is true for information retrieval; studies have indicated that information retrieval tasks are harder to perform on small screen devices [22]. Moreover, users are more likely to perform incorrect choices when selecting from possible links and waste more time carrying out additional scrolling type activities [23,24]. Screen size also influences the quality of information input. Most studies assessing users’ performance have confirmed that data input accuracy can be impacted by keyboard size or character setting, but also by other factors, such as a user’s finger size [25-27]. Generally, a familiar disposition of the display and large keyboards improve user performance.

All these tend to demonstrate that it is crucial to take into account screen size specificity in the design of mobile device interfaces. Therefore, we assessed 6 tactile interfaces representing many of the common interfaces used on tablet PC and smartphones as interfaces built based on user requirements.

Methods

In order to identify the most suitable interface for the effective and reliable recording of numerical data on handheld devices, we performed a crossover randomized trial assessing 6 handheld interfaces designed according to 6 paradigms ranging from commercially available solutions to experimental designs. Each participant had to record several vital numerical signs on each interface. The interfaces were provided automatically in a random order.

Participants

A sample most representative of health care professionals likely to collect, record, or use clinical data as part of their daily activity was recruited within the University Hospitals of Geneva, Switzerland. Recruitment was carried out following approval of the project by the “Commission cantonale d’éthique de la recherche,” the University hospitals’ human research and ethics committee. Previously published research on similar research questions were able to demonstrate significant differences between data entry interfaces with 30 participants involved [13]. However, because we were able to recruit among a large number of caregivers and had no prerequisites as to participants’ characteristics, we relied on a convenience sampling approach. In total, 150 hospital workers selected across 5 categories of professionals were recruited: nurses, assistant nurses, midwives, physicians, and administrative staff. Participation was on a voluntary basis and there was no exclusion criterion.

Instruments

To test the 6 interfaces, we used a common commercially available smartphone, the Samsung Galaxy Note, which has a 5.3-inch screen with a resolution of 1280x800 pixels. This type of smartphone is representative of contemporary smartphones. It is also characterized by a high level of flexibility, which reduces to the minimum the design constraints set by human-machine interfaces. The 6 interfaces that were tested were designed to represent the most complete range of options...
man-machine computer interfaces can offer (Figure 1). Four of the interfaces were chosen among existing interface design frameworks. The numeric keypad interface was chosen because it is recognized as being a very effective at reducing the number of entry errors [13]. The stepper and wheel interfaces were chosen because they are commonly available on 2 major commercial smartphone operating systems, Android and iOS. Finally, the character recognition interface was chosen because it was among the first interfaces proposed for touchscreen devices, such as the PalmPilot. The last two options were developed by two distinct expert committees based on end-user requirements for recording interfaces. The first committee included caregivers who designed a dynamic chart (named “column”) aimed at facilitating the identification of vital sign trends. The second committee included computer scientists at the hospital who developed a fast data entry system (the “circle”) built on an interaction principle similar to the SwiftKey keyboard. The interfaces were implemented by the research team and installed on all devices provided to participants. Experimental designs were previously extensively tested by volunteers to ensure usability of the recording interfaces.

Figure 1. Description of the six data entry interfaces.

Study Procedure and Data Collection

The trial was designed to minimize interaction biases. Based on a set of possible problems identified in a pilot phase, standardized procedures were defined so that any problem (e.g., a system crash) would by default always be handled in a similar way. Each session of the trial took place in a controlled environment where participants could perform the test without any interruption or other disturbance. Before beginning the trial, each participant was shortly briefed on the study purpose and overall procedure. Participants were also informed that there was no time constraint for entering the displayed numbers. Each participant was then provided with a smartphone. Because the study procedure and instructions were available on each device and the entire study process was regulated by a computer program, there was no need for further interaction with the research team. Although the experiment took place in a controlled environment, participants were asked to use only one hand to hold the device to simulate the real-life bedside recording procedure as closely as possible (see Multimedia Appendix 1 for the CONSORT checklist filled for this study).
Participants were first asked to read the instructions of the experiment that were displayed on the smartphone. A short questionnaire then popped up on the screen asking for details on participants’ demographics and computer literacy. Following these steps, the actual trial phase began. Each participant was requested to enter data on all 6 interfaces. The sequence in which interfaces came up was automatically computed and randomly defined for each participant. The procedure was the same for each interface: participants were first able to practice the procedure before entering the data to be recorded. For each interface design, participants were asked to enter 3 types of physiological measures (body temperature, respiratory rate, pulse rate) that were displayed at the top of the screen. We chose these 3 vital signs from among the most frequently collected signs during patient follow-up that could be described using a single numerical value. For this reason, measurements such as blood pressure (composed of 2 values) were discarded as were other types of data, such as dates or time, which we considered would require other types of dedicated entry interfaces. When entering a number, participants could correct it as many times as they wanted before validating it. The computer program randomly generated the type of physiological data to be recorded and a value to be entered. During the training period, participants had to record 2 physiological values with the interface displayed on the screen. Users could not skip this step and had to continue the selection process until they succeeded. This is the only time users were allowed to request external assistance if they did not understand how to use the interface. Once the 2 physiological values were successfully recorded, the testing process started and participants were asked to record 3 random values for each possible interface-sign combination. Data for pulse, respiratory rate, and body temperature were given a predefined range of recordable values (Table 1).

The task was repeated until 3 values were recorded for each of the 3 physiological signs in each of the 6 interfaces. Altogether, this totaled 54 data entries for each participant.

### Measured Variables

All variables needed for assessing the performance of each participant were recorded automatically on the smartphone. Among the variables that were measured were the number of actions performed by the participant, the number of corrections made, and the time until data entry validation. Data could be exported in a comma separated value file. Accuracy was measured by comparing the values participants were requested to enter as they appeared on the top of the smartphone’s screen with the ones actually entered by each participant. Success or failure outcomes were then summed and reported as a proportion of the maximum possible score for each category of data entry. Effectiveness was assessed by measuring the time used to record each of the values.
Analysis

Descriptive summary statistics of continuous variables included means (SD), or medians with ranges, depending on their distribution. Continuous variables were compared using the paired Student t test or the Wilcoxon signed rank test if not normally distributed. For categorical variables, we used frequencies and proportions. Participants’ satisfaction was assessed using the Friedman 2-way ANOVA test. After adjusting for age and familiarity of use of computers or smartphones, possible associations between the data entry time for each type of variable and the type of interface used or the sequence of data entry were examined using multilevel linear models. The same was done to evaluate the number of errors and corrections associated with each interface, this time using multilevel mixed logistic models. To obtain a normal distribution of the dependent variable, we used the log of speed of data entry. Interfaces were nested within the type of data entry, which were themselves nested within the coder.

A P value <.05 was considered statistically significant. We performed all analyses using the statistical package Stata version 12.

Results

Only 4 participants failed to complete the test due to technical problems or professional emergency causing them to leave the study before its completion. These incomplete records were removed from the study analysis. Demographics for the remaining 146 participants could be summarized as follows: mean age of participants was 43.6 years (SD 10.2) with two-thirds (64.4%, 94/146) younger than age 46 years. Most of participants enrolled in the study (92.5%, 135/146) were right-handed and nearly all (95.9%, 140/146) possessed a personal computer, although only 61.0% (89/146) already owned a smartphone. Most (75.3%, 110/146) were caregivers; the remaining 24.7% (36/146) were recruited among administrative personnel or computer scientists classified as “other participants.” Study participant characteristics are detailed in Table 2.

The time needed to enter values with the different interfaces differed depending on the type of vital sign recorded (Table 3). In most cases, it took less time to enter pulse than body temperature values, and even less time to enter respiratory rates (except for the wheel mode). The results provided in Table 3 show that the speed of data entry was also influenced by the type of interface used. Differences were statistically significant (P<.001). The fastest interface for data entry was the numeric keyboard, with a mean entry speed ranging from 2.81 (SD 0.06) to 4.34 (SD 0.08) seconds depending on the type of variable recorded. This was significantly quicker than the stepper (mean 4.31, SD 0.11 seconds to mean 7.23, SD 0.36 seconds), the wheeler (mean 5.13, SD 0.14 seconds to mean 8.35, SD 0.21 seconds), and the circle (mean 4.45, SD 0.01 seconds to mean 9.38, SD 0.28 seconds) models. Finally, the less efficient interfaces were the column (mean 5.25, SD 0.16 seconds to mean 10.49, SD 0.45 seconds) and character recognition (mean 6.32, SD 0.24 seconds to mean 14.68, SD 0.64 seconds) models.

Table 4 shows data entry accuracy for the different interfaces. The numeric keyboard was the most accurate interface (96.3%-99.3% accuracy). It was followed by the stepper (97.9%-98.4% accuracy) and the wheel (95.4%-98.6% accuracy). Other interfaces were associated with more data recording errors. The other interfaces yielded levels of accuracy ranging from 93.2% to 94.3% for the column model, 88.6% and 96.1% for the circle, and 81.5% to 86.8% for the character recognition interface.

Among the tested interfaces, the numeric keyboard achieved the highest level of accuracy for recording pulse data. When looking at the mean scores of all 3 measures, the overall performance of the numeric keyboard was comparable to that of the stepper and wheel models. With regard to speed of data entry, the numeric keyboard scored the highest, with results at least 1.5 times faster than all other interfaces regardless of the type of variable recorded. Interfaces built based on a participative design, such as the circle or column, were associated with additional errors. Character recognition was found to be the most inaccurate and slowest interface for clinical data recording (Table 5).

Discussion

These study results can be explained by several factors. First, the selection mechanism used to enter values into the system was the most influential factor for determining speed of data entry. With a deterministic selection mechanism, such as the keyboard, the actions performed by users are transformed unambiguously into the associated outcome. This differs from a nondeterministic selection mechanism, such as character recognition, where there is no guarantee that the action of the user is transformed into the desired outcome. The interfaces with such nondeterministic selection mechanisms require some level of expertise and training to perform accurate actions that can be transcribed by the system into the desired outcome. The number “1” could, for instance, be at misinterpreted as a “7” and therefore require additional time to be corrected. The second parameter influencing the speed of data entry is the number of actions required to record it. In this regard, the stepper is not as fast as the keyboard because it requires more actions to enter a given numerical value.

An analysis of the error rate enabled us to classify the interfaces into 2 main categories: those yielding a level of accuracy greater than 95% (numeric keyboard, stepper, and wheel) and those yielding levels less than 95% (character recognition, circle, and columns). Interfaces offering no immediate feedback and those where it was difficult to modify recordings (character recognition, wheel) were associated with a significantly increased risk of recording errors.
Table 2. Participant characteristics (N=146).

<table>
<thead>
<tr>
<th>Participant characteristics</th>
<th>n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age (years)</strong></td>
<td></td>
</tr>
<tr>
<td>22-34</td>
<td>42 (28.8)</td>
</tr>
<tr>
<td>35-45</td>
<td>52 (35.6)</td>
</tr>
<tr>
<td>≥46</td>
<td>52 (35.6)</td>
</tr>
<tr>
<td><strong>Gender</strong></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>96 (65.7)</td>
</tr>
<tr>
<td>Male</td>
<td>50 (34.3)</td>
</tr>
<tr>
<td><strong>Hand preference</strong></td>
<td></td>
</tr>
<tr>
<td>Left-handed</td>
<td>11 (7.5)</td>
</tr>
<tr>
<td>Right-handed</td>
<td>135 (92.5)</td>
</tr>
<tr>
<td><strong>Profession</strong></td>
<td></td>
</tr>
<tr>
<td>Assistant nurse</td>
<td>4 (2.7)</td>
</tr>
<tr>
<td>Nurse</td>
<td>76 (52.2)</td>
</tr>
<tr>
<td>Administrative</td>
<td>19 (13.0)</td>
</tr>
<tr>
<td>Midwife</td>
<td>11 (7.5)</td>
</tr>
<tr>
<td>Physician</td>
<td>19 (13.0)</td>
</tr>
<tr>
<td>Other</td>
<td>17 (11.6)</td>
</tr>
<tr>
<td><strong>Has a computer</strong></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>6 (4.1)</td>
</tr>
<tr>
<td>Yes</td>
<td>140 (95.9)</td>
</tr>
<tr>
<td><strong>Has a smartphone</strong></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>57 (39.0)</td>
</tr>
<tr>
<td>Yes</td>
<td>89 (61.0)</td>
</tr>
<tr>
<td><strong>Frequency of vital sign recording</strong></td>
<td></td>
</tr>
<tr>
<td>Never</td>
<td>45 (30.8)</td>
</tr>
<tr>
<td>Monthly</td>
<td>12 (8.2)</td>
</tr>
<tr>
<td>Every day</td>
<td>89 (61.0)</td>
</tr>
</tbody>
</table>

Table 3. Mean time to enter 3 variables for each interface-sign combination.

<table>
<thead>
<tr>
<th>Data type</th>
<th>Interface entry time (sec), mean (SD)</th>
<th></th>
<th></th>
<th></th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Character recognition</td>
<td>Circle</td>
<td>Columns</td>
<td>Numeric keyboard</td>
<td>Stepper</td>
</tr>
<tr>
<td>Pulse</td>
<td>10.50 (0.44)</td>
<td>6.27 (0.14)</td>
<td>10.06 (0.34)</td>
<td>3.08 (0.06)</td>
<td>6.53 (0.36)</td>
</tr>
<tr>
<td>Body temperature</td>
<td>14.68 (0.64)</td>
<td>9.38 (0.28)</td>
<td>10.49 (0.45)</td>
<td>4.34 (0.08)</td>
<td>7.23 (0.18)</td>
</tr>
<tr>
<td>Respiratory rate</td>
<td>6.32 (0.24)</td>
<td>4.45 (0.01)</td>
<td>5.25 (0.16)</td>
<td>2.81 (0.06)</td>
<td>4.31 (0.11)</td>
</tr>
</tbody>
</table>
Our results are in line with other research findings [13,28,29]. A study comparing 5 number entry interfaces for an infusion pump [13] showed lower error rates and speed of data entry when entering values on a number pad rather than a stepper. In a study evaluating data recording accuracy of a keyboard for electronic health recording systems compared with handwritten medical paper records, authors found that 25.6% of vital signs had one or more errors when documented on paper medical records compared to 14.9% errors when documented in an electronic format [28]. This confirms the improved accuracy of a keyboard compared to handwriting, as used in character recognition interfaces. Another study by Wager et al [29] compared error rates for 4 different types of entry devices (paper, computer on wheel, tablet, and direct feed from the monitoring system). Although no details were provided about the interfaces used, the error rate associated with a keyboard on a tablet was 5.6%, on average, similar to our study findings.

There are several limitations to our study. The first is the lack of exact knowledge about the prior level of familiarity users had for each of the 6 interfaces. However, one can reasonably assume that the novel interfaces implemented for the study were unknown to all users, whereas numeric keyboard and character recognition interfaces were probably familiar to computer and mobile phone users. Our study sample was mostly composed of caregivers, although we attempted to minimize this bias by asking users to carry the device in only one hand. The experimental setting cannot be considered as an ecologically valid one because the experiment took place in a controlled environment. Although we attempted to minimize this limitation by asking users to carry the device in only one hand, the fact that the numbers to enter were provided directly on the smartphone did not reflect real-life situations, where numbers are more likely to be read from another source. We chose this method of data delivery to exclude errors that might otherwise arise when transcribing data from one place to another [38].

A second limitation is the unknown influence of the original parameters set into the system on study outcomes [32,33]. Indeed, there are many factors that can influence the data entry process. For instance, the size of the keyboard buttons is known to influence a user’s performance during data entry [27,34]. On infusion pumps, such factors can generate up to 28 different implementations of the data entry interface [13]. In their study, Cauchi et al [35] confirmed that these parameters not only influence data entry performance, but also the severity/degree of recording errors. Likewise, the character recognition algorithm can facilitate or complicate user inputs [36]. Therefore, if our results show clear and statistically sound differences in interface usability and accuracy, other studies analyzing the same interfaces but using different settings may show different results.

Third, despite explicitly instructing participants to act as they would normally during the trial session, while at the same time forbidding observers to intervene after the testing phase, it cannot be excluded that external observers may have had an influence on participants’ behavior. For example, participants may have improved or modified an aspect of their behavior in response to the context in which they were acting rather than in response to the type of device used. This is known as the “Hawthorne effect” [37].

Fourth, the experimental setting cannot be considered as an ecologically valid one because the experiment took place in a controlled environment. Although we attempted to minimize this limitation by asking users to carry the device in only one hand, the fact that the numbers to enter were provided directly on the smartphone did not reflect real-life situations, where numbers are more likely to be read from another source. We chose this method of data delivery to exclude errors that might otherwise arise when transcribing data from one place to another [38].

Fifth, the purely quantitative approach adopted in our study cannot provide explanations as to why variations exist across different interfaces and what the implications are for the design and evaluation of mobile data entry tools. To clearly answer these questions, a more qualitative methodology needs to be adopted, such as the one used by Kushniruk et al [8] in their

### Table 4. Correct data entries for the 3 variables on each of the 6 interfaces.

<table>
<thead>
<tr>
<th>Data type</th>
<th>Correct data entries, n (%)</th>
<th>Mean time (sec)</th>
<th>Mean accuracy, %</th>
<th>Minimum number of actions measured</th>
<th>Selection</th>
<th>Design source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pulse</td>
<td>372 (84.9)</td>
<td>3.41</td>
<td>97.5</td>
<td>1 per digit</td>
<td>Deterministic</td>
<td>Literature</td>
</tr>
<tr>
<td>Body temperature</td>
<td>357 (81.5)</td>
<td>6.02</td>
<td>98.2</td>
<td>Sum of the digits</td>
<td>Deterministic</td>
<td>Android OS</td>
</tr>
<tr>
<td>Respiratory rate</td>
<td>380 (86.8)</td>
<td>8.6</td>
<td>93.9</td>
<td>1 per number</td>
<td>Nondeterministic</td>
<td>iOS</td>
</tr>
</tbody>
</table>

### Table 5. Data entry speed and accuracy for different interface characteristics.

<table>
<thead>
<tr>
<th>Interface characteristics</th>
<th>Mean time (sec)</th>
<th>Mean accuracy, %</th>
<th>Minimum number of actions measured</th>
<th>Design source</th>
<th>Selection</th>
</tr>
</thead>
<tbody>
<tr>
<td>Numeric keyboard</td>
<td>3.41</td>
<td>97.5</td>
<td>1 per digit</td>
<td>Literature</td>
<td>Deterministic</td>
</tr>
<tr>
<td>Stepper</td>
<td>6.02</td>
<td>98.2</td>
<td>Sum of the digits</td>
<td>Android OS</td>
<td>Deterministic</td>
</tr>
<tr>
<td>Circle</td>
<td>6.7</td>
<td>93</td>
<td>1 per digit</td>
<td>Computer scientists</td>
<td>Deterministic</td>
</tr>
<tr>
<td>Wheeler</td>
<td>6.9</td>
<td>97.2</td>
<td>1 per number</td>
<td>iOS</td>
<td>Nondeterministic</td>
</tr>
<tr>
<td>Columns</td>
<td>8.6</td>
<td>93.9</td>
<td>1 per number</td>
<td>Caregivers</td>
<td>Nondeterministic</td>
</tr>
<tr>
<td>Character recognition</td>
<td>10.5</td>
<td>84.4</td>
<td>1 per digit</td>
<td>Palm OS</td>
<td>Nondeterministic</td>
</tr>
</tbody>
</table>
analysis of the relationship between usability problems and prescription errors in handheld applications.

Finally, it is also worth pointing out that evaluating failure based on a binary measure does not make it possible to evaluate the type and potential severity of an error. Indeed, studies such as those carried out by Wiseman et al [39] or Oladimeji et al [13,40] reported up to 7 classes of errors. Due to this limitation, our results do not highlight the fact that interfaces that enable each digit to be chosen independently are much more likely to create errors of higher magnitude and usually critical in a clinical context [12,41,42].

Although there is a marked increase in use of handheld devices in the health care environment, not all interfaces are adapted to specific constraints, such as smaller, interactive tactile screens. The possibility of recording clinical information on mobile devices opens the door to many questions. Interfaces set on these devices are usually complex to use. In his study, Howard [43] showed that this complexity could be explained by 4 reasons: intricacy, equivalence, omniscience, and commitment.

To account for these different aspects, the design of a new interface should evaluate each parameter separately for effectiveness and accuracy before combining parameters into a more complex design. Moreover, the identification of the most appropriate handheld user interface to record numerical data in an electronic format is only one of the many aspects that need to be investigated. Further aspects, such as the manipulation of text data or the recording of graphical data, need to be analyzed also.

Acknowledgments

This work has been funded by the University Hospitals of Geneva.

Conflicts of Interest

The authors are owners or developers of the software used in this study.

Multimedia Appendix 1

CONSORT checklist.

[PDF File (Adobe PDF), 897KB-Multimedia Appendix 1]

References


©Frederic Ehrler, Guy Haller, Evelyne Sarrey, Magali Walesa, Rolf Wiipfli, Christian Lovis. Originally published in JMIR Human Factors (http://humanfactors.jmir.org), 15.12.2015. This is an open-access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on http://humanfactors.jmir.org, as well as this copyright and license information must be included.