Usability of a Web-based Software Tool for History Taking in the Emergency Department

JF Renggli, C Eken, V Siegrist, RN Ortega, C Nickel, C Rosin, R Hertwig & R Bingisser

Abstract
Medical history taking is an important step within the diagnostic process. This study aims to assess the quality and usability (effectiveness, satisfaction, efficiency) of a web-based medical history taking app in the emergency department. During three weeks, patients and junior physicians filled out study questionnaires about the app. Senior physicians rated the quality of medical histories taken by junior physicians and app. In 241 patients, the studied app showed excellent usability with patients not in need of immediate medical attention. Senior physicians rated medical histories as more complete when app was used by patients in comparison to conventional history taking alone (p<0.01). Current app could not substitute medical history taking by physicians, but could definitely rather be used to gather ancillary information.

Keywords
Emergency care, History taking, Quality improvement, Information technology, Patient safety

Key learning points
• A medical history taking app could currently not substitute a physician.
• A medical history taking app should be used to gather ancillary information.
• The use of medical history taking app doesn’t affect patient and physician satisfaction.

Introduction
Improvements in healthcare have mostly relied on technological advances. So far, progress was made mainly in the fields of clinical chemistry (e.g. early detection of myocardial infarction), imaging (e.g. early detection of tumours) and interventions (e.g. molecular targeting in cancer treatment). However, certain areas in clinical medicine are still in need of research and improvement, such as the very beginning of the diagnostic process, namely medical history taking.

Medical history taking is a pivotal step in the diagnostic process, but it has not undergone much change since Hippocrates, except for the tools used for recording. Electronic documentation of medical history is nowadays common practice. However, there is insufficient evidence for the benefits of structuring and/or coding patient histories using electronic applications. Several factors, such as the resources available, the differences between information given and recorded and the delays in recording may contribute to this lack of evidence. First, the time available for direct patient contact has substantially decreased. Second, physicians tend to screen information according to their experience, emphasising information useful for immediate decision-making, focusing on common specific symptoms, such as chest pain, and potentially underreporting nonspecific symptoms, such as weakness. Third, documentation is often not synchronized with history taking. Such delayed documentation is hampered by working memory capacity. As soon as new information becomes apparent, such as results of imaging or blood work, physicians are prone to deviate from their unbiased histories. Furthermore, repetitive questioning of patients, common in teaching hospitals, may alter the memories of patients. These problems may even exacerbate under circumstances of high stress to patients and caregivers, such as in the emergency setting. Thus, a standardised medical history taken by the patient himself with the support of a web-based software tool might help overcome some of the stated problems and improve the quality of care.

Quality can be defined as the degree to which a set of inherent characteristics fulfils requirements. In medicine, requirements are usually defined by experts outlining indicators that can be measured and compared to a specific target. Quality can be measured using different dimensions, such as patient satisfaction, effectiveness (i.e. achieving an expected result), or completeness of measures taken or information recorded. Obviously, there is a complex interaction between patient satisfaction,
requirements defined by experts, and efficiency (i.e. achieving a result with minimum expenditure). Usability is a concept taking into account the interaction between satisfaction, effectiveness and efficiency, and optimal usability is achieved if all three components are perfectly aligned. The goal of the current study was to investigate the quality and the usability of a web-based software tool (i.e. app) for medical history taking, adding measures of efficiency, in a cohort of patients with lower acuity presenting to an urban emergency department (ED). Additionally, we compared patients’ and physicians’ satisfaction regarding care in both junior physician- (i.e. conventional) and patient-generated-medical history taking. It has been shown that patient recorded histories are highly accurate, and that patients are more likely to write down intimate details rather than share them orally with their physicians. Thus, we hypothesised that quality defined by the primary endpoint of effectiveness of information gathering would increase when using an app for medical history taking. Since people tend to have a preference for the current state of affairs (i.e. status quo bias), satisfaction might decrease in patients and physicians. To the best of our knowledge, only the feasibility of such app has been reported so far, but quality and usability were never assessed in an emergency setting.

Materials and Methods

Study population
Patients presenting to the ED were included in the study during a period of 3 weeks if they spoke English or German and gave informed consent. We excluded patients in need of immediate medical attention (Emergency Severity Index, ESI 1-2) or with minor complaints who did not require any resources (ESI 5). Patients with an ESI score of 3 or 4 were seated in the waiting room of the ED, where screening for the study took place 24/7. The study was approved by the regional ethics committee (EKNZ 2016-02091).

Patient-generated medical history taking application
We used a commercially available web-based software tool (app) named Sublimd provided by a medtech company (sublimd.com). This app is available in English and German, and proposes specific questions to patients in order to generate their medical histories. The content of the app, 150 presenting symptoms with over 3000 corresponding questions, was aligned with the protocols provided by medStandards (www.medstandards.com), the decision support tool owned by the University Hospital Basel, Switzerland, providing over 1300 diagnostic and therapeutic protocols and algorithms based on current medical evidence and clinical guidelines.

As patients answered the questions presented by the app, the software dynamically adapted the selection of further questions. On this basis, detailed patient histories consisting of the presenting symptoms, medications, allergies, personal and family histories, as well as systematic history of symptoms were obtained. When all necessary information was gathered, the app automatically generated a written report of patients’ medical histories.

Study procedure
The study was designed as an intervention comparing a baseline (week 1: junior physicians performed history-taking and patients did not record their medical histories electronically) and an intervention (week 3: all included patients recorded their medical histories electronically, and junior and senior physicians had access to this information). Week 2 was a run-in period, in which patients recorded their medical histories on tablets, but only senior physicians had access to this information.

It is common practice in our ED that history taking is performed by junior physicians after patients are assigned a treatment bay: bedside questioning is followed by recording patients’ medical histories in the electronic health record (EHR) system on nearby computers. All patient histories are presented to senior physicians by the junior physicians.

The study process was adapted each week as follows (Figure 1):

Baseline (week 1): Medical history taking was only carried out conventionally.
Run-in (week 2): Patients received tablets to record their medical history. They were briefly
instructed on how to use the app by the study team. Junior physicians did not have access to the patient-generated histories, whereas senior physicians had access to those once junior physicians presented their patient histories. Thus, senior physicians rated the quality of patient- vs junior physician-generated medical histories on a visual analogue scale (VAS) from 0 to 10. Additionally, they selected history categories with missing information and rated how strongly the lack of this information would affect patients’ successful treatment.

**Intervention** (week 3): As in the run-in week, patients received tablets to record their medical histories and were briefly instructed how to use the app. In contrast, junior physicians now had access to the patient-generated histories before approaching patients. Senior physicians had access to both medical histories.

**Study questionnaires**

After ED work-up, all included patients and their physicians were asked to fill out different questionnaires on tablets, related to the three components of usability: effectiveness, satisfaction and efficiency\(^{15}\) (Table 1). Answers were given either on VAS, typed-in numbers or as yes/no options. Questionnaires were translated into English using forward translation and reconciliation followed by back translation.\(^{23}\)

The primary endpoint defined as the key question testing effectiveness was “Have you obtained helpful information from another source than your own medical history?” of the junior physicians’ questionnaire (Question 1, Table 1). Secondary outcomes were all remaining questions in the patients’, junior and senior physicians’ questionnaires.

The questions 2-3 and 9-11 were asked to evaluate satisfaction. The effectiveness questions were evaluated with questions 4-6 and 12-15. As a measure of efficiency, questions 7-8 and 16-17 were included in the questionnaires. Finally, questions 18 to 20 were added to measure and control for patients’ digital experience.

Additionally, senior physicians rated the completeness of the medical histories gathered by junior physicians, and the importance of missing information (questions 21-23).

During the run-in and intervention weeks, each patient completed the System Usability Scale (SUS) questionnaire.\(^{24}\) The SUS is a ten-item questionnaire with five options (from “strongly agree” to “strongly disagree”) measuring the overall usability of a system. Moreover, every junior physician completed the Technology Acceptance Model (TAM) questionnaire\(^{25}\) at the end of the study period, which assesses how users accept and adopt a new technology. For our purpose, we only used the “perceived usefulness” part of the TAM (6 questions on a 7-Point Likert-scale).

**Statistical analyses**

Descriptive data are presented as mean values with standard deviation (SD) for normally distributed variables, median (+/- interquartile range; IQR) for non-normally distributed variables, or as absolute numbers and percentages of the study population for categorical variables. Chi-squared Test and Fisher’s Exact Test were used for hypothesis testing in count data.

Generalised linear regression analysis with negative binomial distribution was used to model the association between continuous dependent variables (questionnaire answers) and the study groups. Logistic regression analysis was performed for binary variables. A multivariate adjusted analysis included the covariates: age, sex and years of experience of the junior physicians. Model assumptions were tested using diagnostic plots. Results were expressed as estimates with standard errors (SE) and p-values. An odds ratio with confidence interval (CI) was calculated when applicable. A p-value of <0.05 was considered statistically significant. All calculations were made using the statistical software R version 3.3.2 (https://www.R-project.org).

**Results**

**Study population**

During the study period of 3 weeks, we screened 629 patients of which 320 were enrolled in the study (Figure 2). Exclusion rate was 49.1%; the main reasons for these 309 exclusions were “declined without reason statement” (43.4%), “symptoms did not allow
Three patients mentioned data safety concerns. Out of the 320 patient files, 241 were complete and therefore used for further analysis. Demographic characteristics of each study period are shown in Table 2. Of note, there was a significant difference in median age between included (median = 41 years) and excluded patients (median = 51 years) (p < 0.05).

Table 1. Study questionnaires

<table>
<thead>
<tr>
<th>Questions</th>
<th>Baseline (n = 86)</th>
<th>Intervention (n = 70)</th>
<th>Estimate (SE)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Have you obtained helpful information from another source other than your own history?</td>
<td>No: 51.2%, Yes: 48.8%</td>
<td>No: 34.3%, Yes: 65.7%</td>
<td>0.92 (0.36)</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>2. How good was the interaction between you and your patient?</td>
<td>8.2 (6.9 - 9.4)</td>
<td>8.6 (7.3 - 9.8)</td>
<td>0.24 (0.17)</td>
<td>0.18</td>
</tr>
<tr>
<td>3. How satisfied are you with how you cared for your patient?</td>
<td>7.7 (5.2 - 9.2)</td>
<td>8.2 (6.7 - 9.6)</td>
<td>0.26 (0.18)</td>
<td>0.16</td>
</tr>
<tr>
<td>4. How certain are you about your patient’s diagnosis?</td>
<td>8.4 (6.0 - 10.0)</td>
<td>9.1 (7.1 - 10.0)</td>
<td>0.25 (0.21)</td>
<td>0.23</td>
</tr>
<tr>
<td>5. Have you obtained all relevant information for the care of your patient from the application?</td>
<td>-</td>
<td>No: 50.0%, Yes: 48.6%, N/A: 1.4%</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>6. Could you integrate the information from the app in your workflow?</td>
<td>-</td>
<td>7.4 (4.4 - 8.9)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>7. Has the program reduced your workload?</td>
<td>-</td>
<td>No: 45.7%, Yes: 54.3%</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>8. Is the medical history obtained through the app useful to be used as patient documentation?</td>
<td>-</td>
<td>8.1 (7.1 - 9.1)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>9. All in all, how satisfied were you with your ED stay today?</td>
<td>8.4 (6.4 - 9.9)</td>
<td>7.7 (6.5 - 10.0)</td>
<td>-0.10 (0.25)</td>
<td>0.69</td>
</tr>
<tr>
<td>10. How good was the interaction with your doctor?</td>
<td>9.2 (7.5 - 10.0)</td>
<td>9.3 (7.4 - 10.0)</td>
<td>0.02 (0.21)</td>
<td>0.92</td>
</tr>
<tr>
<td>11. How satisfied are you with the care you received from your doctor?</td>
<td>8.9 (7.6 - 10.0)</td>
<td>9.2 (7.7 - 10.0)</td>
<td>0.14 (0.21)</td>
<td>0.52</td>
</tr>
<tr>
<td>12. Were you able to give information about all your essential health issues?</td>
<td>9.0 (7.3 - 10.0)</td>
<td>9.2 (8.0 - 10.0)</td>
<td>0.20 (0.21)</td>
<td>0.47</td>
</tr>
<tr>
<td>13. Do you think the program helped you discuss the essential points with your doctor?</td>
<td>-</td>
<td>7.5 (5.5 - 9.0)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>14. Did you give the program sensitive information, which you might not have told your doctor?</td>
<td>-</td>
<td>No: 80.0%, Yes: 12.9%, N/A: 7.1%</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>15. Would you give information regarding a sensitive topic to a tablet rather than to a doctor?</td>
<td>-</td>
<td>No: 62.9%, Yes: 21.4%, N/A: 15.7%</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>16. How did you experience today’s waiting time in the ED?</td>
<td>3.7 (0.8 - 6.7)</td>
<td>3.1 (1.2 - 6.7)</td>
<td>-0.04 (0.19)</td>
<td>0.83</td>
</tr>
<tr>
<td>17. Please estimate how much time your doctor needed to take your medical history.</td>
<td>10 min (7.0 - 15.0)</td>
<td>10 min (5.0 - 19.8)</td>
<td>0.14 (0.17)</td>
<td>0.42</td>
</tr>
<tr>
<td>18. Was the language of the program easy to understand?</td>
<td>-</td>
<td>No: 4.3%, Yes: 94.3%, N/A: 1.4%</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>19. Do you use a computer or other related electronic devices, e.g. tablets or smartphones?</td>
<td>-</td>
<td>No: 10.0%, Yes: 88.6%, N/A: 1.4%</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>20. In a typical day how long do you use a tablet or related electronic devices on a daily basis?</td>
<td>-</td>
<td>60 min (51 - 180)</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Numeric answers are expressed as median with first and third quartiles. Categorical variables are expressed as percentage over the total answers. Regression estimates are shown with standard errors and p-values whenever comparisons between groups were made; SE = standard error, min = minutes, N/A = not applicable, NP = no preference.
Endpoints

As primary endpoint, junior physicians (n=28) stated that they obtained helpful information from another source than their own medical histories in 46 cases (65.7%) in the intervention week, as opposed to 24 cases (34.3%) in the baseline week (p < 0.01) (Table 1), corresponding to an odds ratio of 2.5 (CI = 1.3-5.1) using an adjusted regression model.

Ratings of satisfaction showed to be high in the baseline and the intervention week, resulting in a ceiling effect (Table 1). Accordingly, there was no significant difference in satisfaction ratings between baseline and intervention.

Physicians rated their confidence in their diagnoses with a median of 8.4 (IQR 6.0; 10.0) in the baseline and with 9.1 (IQR 7.1; 10.0) in the intervention (p = 0.23; Table 1). During the intervention, physicians answered that they have been able to obtain all relevant information from patient-generated histories in 48.6% of the cases. Additionally, they rated the ability to integrate the information from the app in their workflow with a median of 7.4 (IQR 4.4; 8.9) (Table 1).

On the other hand, patients’ effectiveness ratings were not significantly different between comparison weeks. Patients rated the helpfulness of the app to discuss essential points with their physicians with a median of 7.5 (IQR 5.5; 9.0). Interestingly, 21.4% of the patients preferred disclosing information about sensitive topics to the app rather than talking to their physician. In this patient group 47% had a positive screening for high-risk alcohol consumption (i.e. AUDIT C Score≥2) and 29% reported drug abuse.

Regarding efficiency, junior physicians rated that the app reduced their workload in 54.3% of the cases during intervention. Additionally, the potential of the patient-generated histories to be used as EHR documentation was rated high with a median of 8.1 (IQR 7.1; 9.1) (Table 1). On the patient side, there was no change in efficiency between baseline and intervention.

Senior physicians rated the medical histories gathered by junior physicians as significantly more complete during the intervention (median = 10; IQR 8; 10; p < 0.01). When rating the importance of the information missing in the medical histories, no significant difference was found between weeks (p = 0.07; Table 1).

Importantly, the median SUS score for patients (n=141 completed questionnaires) using the app was 82.5 (IQR 65.0; 90.0), showing excellent usability for the app. The TAM score for physicians was a mean of 5.1 (SD = 1.1) on a 7-Point Likert-scale, showing a good acceptance of the app.

Comparison of medical histories taken by web-based software tool vs junior-physicians

85 pairs of independent patient and physician histories were rated for completeness by senior physicians during the run-in week. Percentages of histories with missing information in the categories “history of present illness”, “social history”, “personal history”, “medications”, “review of systems”, “family history”, “health habits” and “allergies” are shown in Figure 3. Results showed that the category “history of present illness” was perceived as significantly less accurate for medical histories taken by the app in comparison to those taken by physicians (information missing in 49.4% vs. 14.1%, p < 0.01), while “social history” was significantly more complete in the app in comparison to physicians’ histories (2.4% vs
Usability of a Web-based Software Tool for History Taking in the Emergency Department

8.2% incomplete histories, \( p < 0.01 \)). Although not significant, the app showed a trend towards higher completeness in the remaining categories except for "medications".

Discussion

The main results of our study are the confirmation of effectiveness, the evidence of increased efficiency, and an excellent usability of the investigated web-based software tool for medical history taking. Our data also confirmed previous studies showing the feasibility and the relative increase in standardised information gathered by the use of a history taking application.\(^{12,21,27-29}\)

The following results merit discussion: First, effectiveness was tested using the primary outcome of "obtaining helpful information". This outcome clearly focused on physicians' requirements. In light of the importance of the diagnostic process in the emergency setting, this should not be underestimated. Physicians must be effective at information gathering, as the consequence of incomplete or missing information may be a diagnostic error that has been linked to cause up to 10% of total mortality (30). It is therefore crucial to improve the diagnostic process. A recent report stated that "although health IT has the potential to improve diagnosis and reduce diagnostic errors, many experts are concerned that it currently is not effectively facilitating the diagnostic process and may even be contributing to errors".\(^{31}\) However, in our study, junior physicians felt significantly more confident about their diagnoses if supported by the patient-generated history. This effect might be explained by the more systematic approach of a web-based software tool for medical history taking, by reduced workload, and by ancillary information adding pieces to the diagnostic puzzle. Junior physicians also reported a high ability to integrate the app into their workflow, and specifically to use the patient-generated history as documentation in the EHR. Furthermore, senior physicians considered the patient-generated history in combination with the physician-generated history to be superior to either one alone.

Second, there was evidence for increased efficiency in the history taking process, as the majority of the junior physicians reported a decreased workload using the app. Of note, perceived waiting times were unchanged when patients used the app. Third, to our surprise, satisfaction did not decrease in patients. There was no trade-off between efficiency and satisfaction with the care patients received from their doctor. Fourth, usability scores showed excellent values taken from the patients’ perspective. Usability should not be measured using a single test only, but effectiveness, satisfaction, and efficiency need to be jointly appraised. All three dimensions, though not perfectly aligned, were rated higher in the intervention phase in comparison to the baseline.

Senior physicians stated that important information was missing when comparing the patient-generated with the junior physician-generated medical histories. Interestingly, “history of present illness” and “medications” were superior in the junior physicians’ histories. The results about the “medications” might be due to the fact that for patients who takes numerous medications (e.g. elderly), it could have been challenging to enter all manually in the app. These patients usually carry their medication list, which could be included in the EHR as an electronic copy. Thus, physicians would better document the medications compared to the app.

Nevertheless, patients' health status or restraint to use a tablet, as well as limitations regarding the completeness of the category “history of presenting illness” may hinder the use of this app. Indeed, there are several limitations to this study. Although, this is a single centre study, our population is comparable to other European urban EDs regarding case mix,\(^{22}\) hospitalisation rate (>30%) and immigrant population (>30%).\(^{6,7}\) Despite the moderate number of participants included in the study, the large effect size allows to confirm the hypothesis regarding the primary endpoint.

The number of exclusions was higher than in another recent study,\(^{32}\) possibly due to the use of a tablet, particularly daunting to the older population. Of note, median age of patients declining to participate was ten years higher as compared to the included population. Obviously, younger patients are more experienced in the use of tablets or other technology and may therefore be overrepresented in our study. Nevertheless, the oldest patient taking part in this study was 85 years old.

Even though a medically trained study team was present in three shifts around the clock, recruiting patients arriving late at night was challenging. One might speculate that night-shift patients are different from other emergency patients as to willingness to participate in a study. However, recruitment during night-time is rarely attempted in such study designs.

Finally, the application was available only in English and German. Surprisingly, we did not have to exclude more than 8% of the patients due to language barriers – our expectation being around 30% due to the large immigrant population in Basel, Switzerland. The limited amount of languages offered did lead to the exclusion of some patients, but the potential for future applications offering a selection of languages is apparent: to reach patients who do not speak one of the languages spoken by the physicians.
Usability of a Web-based Software Tool for History Taking in the Emergency Department

Taken together, a patient-generated medical history using a web-based software tool using a tablet resulted in excellent usability in an emergency population of lower acuity. However, its use in the current form is limited by shortfalls in the categories “history of present illness” and “medication”. Therefore, its potential remains in gathering ancillary information rather than substituting history taking by physicians. It is evident that physicians focus on the diagnostic process, such as gathering complete information in order to reduce diagnostic error, while patients may rather focus on conveying information deemed important and on receiving information facilitating the transition process. Thus, such apps could help save the physicians’ time and simultaneously benefit the patients’ need to convey all information they consider important.

Conflict of interests

The Authors declare that there is no conflict of interest. No funding was received from Sublimid.

References