User testing of an admission, discharge, transfer system: Usability evaluation

Majid Jangi1,2, Reza Khajouei3*, Mahmoud Tara2, Mohammad Reza Mazaheri Habibi4, Azade Kamel Ghalibaf2, Sara Zangouei5, Mostafa Mostafavi6

1Health Information Technology Research Center, Isfahan University of Medical Sciences, Isfahan, Iran
2Department of Medical Informatics, School of Medicine, Mashhad University of Medical Sciences, Mashhad, Iran
3Medical Informatics Research Center, Institute for Future Studies in Health, Kerman University of Medical Sciences, Kerman, Iran
4Department of Health Information Technology, Varastegan Institute for Medical Sciences, Mashhad, Iran
5Islamic Azad University of Mashhad, Mashhad, Iran
6Medical Informatics Research Center, Institute for Future Studies in Health, Kerman University of Medical Sciences, Kerman, Iran

ABSTRACT

Introduction: To improve the first step of the hospitalization procedure, appropriate interaction must be established between users and the admission, discharge, and transfer system. The aim of this study was to evaluate the usability of the ADT system in some of selected Iranian non-teaching hospitals.

Material and Methods: This study was cross-sectional research that has evaluated the usability of a selected ADT system using the think-aloud method by 11 medical record administrators. Users were asked to follow the provided scenario, then share and elaborate on what they saw, thought about, did, felt, and decided during their interaction with the system. Users’ feedbacks were collected and organized into four main categories for further processing.

Results: To evaluate the usability of an ADT system, four routine scenario tasks were followed by users and only 45.45% of them could implement all tasks. Overall, 36 independent problems were identified. All problems were related to the data entry categories that accounted for the largest share. The most important problems were related to the issues regarding “date of birth” field in this category which deals with the outpatient admission process.

Conclusion: The study of the usability testing method indicated that the ADT subsystem of non-teaching hospital has many problems in interact with real users with the system. It showed that more than half of the users could not completely and successfully perform the entire real-world scenario tasks. Furthermore, the most usability problems were found in data entry categories.

Keywords: Usability Evaluation Think Aloud ADT System Medical Informatics Hospital Information Systems

INTRODUCTION

Hospital information systems (HIS) are used to mechanize traditional and manual hospital processes. These systems are expected to help hospital staff increase their work speed and accuracy in carrying out assigned duties as well as reducing administrative and medical errors, which can improve the quality of underlying medical procedures [1]. One of the most basic and crucial HIS components is the admission-discharge-transfer (ADT) subsystems that play an important role in both introducing the patient to the system and finalizing the service. The main tasks of ADT are admitting outpatients and inpatients, discharge them, and transfer them as needed. The main subtasks of admitting outpatients and inpatients are as follows: completing forms of admission and recording patient information, assigning beds to the patient, and registering specific hospitalization data. To be performed correctly for the inpatient admission process, appropriate interaction must be established between the system and reception staff users of the ADT system. The quality of users and system interaction depends on several features, such as...
convenience and ease of use, known as “Usability” [2]. On the one hand, problems associated with usability strongly affect the user's ability to do their daily activities [3]. In addition to developing appropriate interaction, usability improvement can drop the development costs and system support. According to Karat [4], for any investment in the usability and design stages of the system, there was a 10 to 100 fold return. System usability problems cause users to incorrectly connect to the system. Accordingly, due to the sensitivity of information and usability errors, inappropriate use of health care systems is closely related to errors and would have a negative impact on the patients' health improvement [3, 5]. Regarding the enormous impacts of usability of information systems on user satisfaction and user-system interaction, evaluation results can be useful in the improvement of design or redesign of the system interface [6]. Also, usability evaluation in the user-centered design for information systems is an important step [7]. Three approaches to evaluate the usability of information systems are included as follows: inspection-based, user-based, and model-based [8]. The most important is the user-based approach [9]. Using which, sample activities of real users are observed during their activities and reactions are recorded, and analyzed by the researchers [2]. During recording, some factors as the time required completing a task, task completion rate, and number and types of errors are measured [8]. Based on this approach, several recommended methods such as think-aloud assessment have been introduced as the gold standard for usability studies [10].

In Iran, only a few studies have evaluated the usability of HIS, including ADT systems [11-16]. Most of them have been focused on indirect evaluation methods such as user surveys and expert opinions. However, thereal participation of users and their real interaction issues had not been covered yet. In this study, by applying the think-aloud method, we aimed to evaluate the usability of the ADT system in the selected non-teaching hospital which uses information systems. Having considered this method, we mostly focused on problems and issues which our real users have met while interacting with the system [17].

MATERIAL AND METHODS

This was a cross-sectional descriptive study that evaluated the selected ADT system based on the think-aloud evaluation method. The study was carried out on the ADT system of the HIS, which has been used in fifty non-teaching hospitals in Iran.

Participants

Eleven system users participated who were medical records administrators (undergraduate students of medical records disciplinary in the final year of their education program). They had the experience of working with other ADT systems. The point is that the participants have not worked with the ADT of this study so far. Ten out of 11 participants were women and who were categorized by age into 20 to 26 years.

Partners

In this study, four trained medical informatics specialists who participated in the usability evaluation were used as facilitators. They did not interfere with the evaluation process, and only if participants halted the thinking aloud process, they briefly reminded and urged them to loudly express their thoughts [18]. In evaluating each participant, there was a facilitator to evaluate each participant's interaction. Four evaluation sessions were performed simultaneously in the rooms equipped. A software engineer was also present alongside the team to fix possible technical problems during the evaluation process.

System and setting

The evaluation was carried out in rooms similar to real hospital working environments with adequate lighting. Each room was equipped with a table, two chairs, and a computer system equipped with a standard keyboard and a mouse. The computer was processing the evaluation system. The client version of the ADT system was installed on the system as well. We also used Camtasia Studio 8 software to record the user-system interaction including mouse clicks and movements, keystrokes, and any typing by the user. This software was active in the background and did not disturb the user's work. In addition, a microphone and a digital camera were employed to record the users’ voices, gestures, and behaviors.

Required materials in evaluation

The scenario was created based on main subtasks of admitting that cover all functionality of the system and included as follows: outpatient admissions, completion of the inpatient admission form, selection of the inpatient's bed, and information entry regarding the inpatient’s attendant. The scenario was developed for users to perform the think-aloud method. It included several ADT functions for an imaginary patient with specific clinical and demographic characteristics (Fig. 1). We tested the scenario with five real users of the ADT system in a hospital and the average time for conducting this scenario was between five to eight minutes. The success or failure of users was investigated when each task was completed. A specific form was also designed for the facilitators’ report to allow them to record their comments during the evaluation, the start and end time of the process, as well as the specific problems of their concern (Fig 2).
Evaluation process

The think-aloud method aimed was to collect data on the users' cognitive interaction with the system. In this study, users were asked to express whatever they did, saw, thought, and felt about their interaction, and also the decisions they possibly have made during a particular case [18].

The evaluation process was carried out in some main stages: In the first stage, users and facilitators were trained by chief researchers. Explanations were given to users on how to express their thoughts, feeling, etc. They were also assured that the aim of the evaluation was research and all of their expressions will be kept confidential [18]. Training of the think-aloud evaluation method was carried out in a joint meeting for 30 minutes. Also, before starting each assessment, every user has retrained from the think-aloud evaluation method for fifteen minutes. A fifteen-minute instruction was also given to every facilitator on how to manage a think-aloud session, including their duties, and the importance of non-interference with the users' interaction. In the second stage, users' expressions were organized and pre-processed by facilitators for further analysis. During the process, the users' recorded voice along with images and movies from users' interaction and also the users' gestures were tagged and saved in separate folders for further analysis. In addition, the facilitators' reported forms were also converted into electronic forms and placed in designated folders. Finally, the evaluation process was finished.

Analysis

Having analyzed the evaluation results, two researchers independently reviewed the videos and sounds recorded from the users and facilitator's notes. In order to identify usability problems, the researchers were attended to:

1. The interaction video of the users' activities (such as input data and mouse movements),
2. The users recorded verbal comments and talks during the process
3. The recorded videos of users' gestures, and
4. The announcement of points by the facilitators.

Based on the conflict resolution strategy, in cases where there was a discrepancy between problems for two independent researchers, the third researcher reviewed differences and gave his opinion.

To classify problems, the method used by Haake et al. [19] was applied. According to this method, problems were divided into four main categories:

1. Layout and order of system components (e.g. user did not find a specific key or link on the page)
2. Terminology (e.g., user did not understand some options or description contained in the guide).
3. Data entry (e.g., user did not know how to enter information in a field using a specific task).
4. Comprehensiveness (e.g., user felt that the system guide was not comprehensive).

Besides, based on the analysis of recorded videos and users’ voices, items listed in Table 1 were measured. Total and average numbers of collected data were calculated in independent evaluation sessions. Data were analyzed by using Microsoft Excel 2010 software.

Table 1: List of measured items in the evaluation

<table>
<thead>
<tr>
<th>Row</th>
<th>Factor</th>
<th>Unit of measurement</th>
<th>Total (n=11)</th>
<th>Average Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Length of time needed to implement the scenario</td>
<td>Minutes</td>
<td>160:05</td>
<td>14:36</td>
</tr>
<tr>
<td>2</td>
<td>Number of times to implement the scenario (success/failure)</td>
<td>Number</td>
<td>5</td>
<td>0.45</td>
</tr>
<tr>
<td>3</td>
<td>Number of times problems caused stoppage in implementing the scenario</td>
<td>Number</td>
<td>36</td>
<td>3.27</td>
</tr>
<tr>
<td>4</td>
<td>Number of times problems were resolved by the user and the process continued</td>
<td>Number</td>
<td>15</td>
<td>1.36</td>
</tr>
<tr>
<td>5</td>
<td>Number of times problems were resolved by the users with the help of the facilitator and the process continued</td>
<td>Number</td>
<td>8</td>
<td>0.73</td>
</tr>
<tr>
<td>6</td>
<td>Number of times problems were not resolved and were abandoned by the user and the process continued, if possible, without considering them.</td>
<td>Number</td>
<td>13</td>
<td>1.18</td>
</tr>
<tr>
<td>7</td>
<td>Length of time to resolve problems (problems that caused stoppage in implementing the scenario)</td>
<td>Minutes</td>
<td>33:18</td>
<td>3:01</td>
</tr>
<tr>
<td>8</td>
<td>Number of times the system guide was used</td>
<td>Number</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>9</td>
<td>Number of times help was given by the facilitator to the user</td>
<td>Number</td>
<td>50</td>
<td>4.5</td>
</tr>
<tr>
<td>10</td>
<td>Length of time used by the facilitator in helping the user</td>
<td>Minutes</td>
<td>15:30</td>
<td>1:42</td>
</tr>
<tr>
<td>11</td>
<td>Number of times users exhibited nervous status while doing tasks</td>
<td>Number</td>
<td>8</td>
<td>0.72</td>
</tr>
<tr>
<td>12</td>
<td>Number of times the phrase “I wish” were used</td>
<td>Number</td>
<td>25</td>
<td>2.27</td>
</tr>
<tr>
<td>13</td>
<td>Number of technical problems</td>
<td>Number</td>
<td>18</td>
<td>1.63</td>
</tr>
</tbody>
</table>

**RESULTS**

By analyzing the 11 recordings, interesting results were taken. Our first finding was how much the task was fully accomplished. From the four given tasks inserted into the scenario, 100% of users completed the first task (outpatient admissions). Also, 81.1% of users completed the second (completion of the inpatient admission form) task and 54.54% of them implemented the third (selection of the inpatient's bed) task successfully, and only 45.45% of them implemented the fourth (information entry regarding the inpatient's attendant) task. On average, 54.55% of users didn’t complete the scenario (all tasks) successfully.

Some of these users spent a lot of time on the process. So that on average, the required time for each assessment was 14 minutes and 36 seconds.

Not all of the spending time was for the scenario, part of it was devoted to solving problems. For example, on average time, users spent three minutes and one second to solve problems, which was equivalent to 21% of the duration of each assessment. Based on the method used by Haake et al. [19], Table 2 presents obtained data related to usability problems.

Table 2: Categories of identified problems based on Haake et al. [19]

<table>
<thead>
<tr>
<th>Category</th>
<th>Number of problems</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Layout and order of system components</td>
<td>13</td>
<td>36</td>
</tr>
<tr>
<td>Terminology</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>Data entry</td>
<td>21</td>
<td>58</td>
</tr>
<tr>
<td>Comprehensiveness</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>Cumulative categories</td>
<td>36</td>
<td>100</td>
</tr>
</tbody>
</table>

As Table 2 shows, 36 problems were totally identified. Among them, eight cases (22% of the problems) were resolved by facilitators avoiding the process to be ceased. 13 cases (36% of the problems) were not resolved and were abandoned by users, so the process was continued without the completion of the assigned task. For all the remaining cases, problems were resolved by users without any intervention of facilitators.

In classifying the problems based on the four mentioned groups, problems related to the data entry category accounted for the largest percentage of problems. Most problems in the data entry categories were related to the date of birth field, which is in the outpatient admission form. Incorrect entry of the date in the field due to lack of user awareness of its format caused 22% of the problems in this field, such that 8 out of 11 people had difficulties with this field.

The second problem in the data entry category was the referrer physician field that accounted for 8% of problems; three users had difficulty with this field. Several other problems of low incidence were also observed during the study. For instance, two users (6%) were not finding the save button (Fig 3).
By removing duplicate cases, out of 36 registered problems, only 21 unique problems were identified (Fig 4).

Results of 11 evaluations were collected based on determining measurements (Table 1) and in all of them, users did not use the system guide after facing problems. Furthermore, 72.72% of users became upset while interacting with the system, 54.54% of users used the word "if" 15 times while working with the system. Also, 83.33% of users used the expression "I wish the date format was pre-specified". As a result of evaluations, a list of "wish sentences" that were usually expressed by users is shown in (Table 3). By considering their suggestions, the system would do finitely improve.

During the evaluation processes, the system encountered many technical difficulties (18 times) which the technical expert fixed them.

**DISCUSSION**

This study, evaluated the usability of the ADT system, by applying the standard usability evaluation methods. More than half of users could not complete all of the given scenario-based tasks. In a similar experience in 2012, Verheul et al. [20] stated that none of their users were able to perform all of the given tasks completely.

<table>
<thead>
<tr>
<th>Row</th>
<th>Sentence</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>I wish the fields were in a menu or list.</td>
</tr>
<tr>
<td>2</td>
<td>I wish the date format were pre-specified.</td>
</tr>
<tr>
<td>3</td>
<td>I wish the introduction of the clinic option were in a better position.</td>
</tr>
<tr>
<td>4</td>
<td>I wish the free beds for inpatient allocation were more obvious.</td>
</tr>
<tr>
<td>5</td>
<td>I wish it had larger fonts.</td>
</tr>
<tr>
<td>6</td>
<td>I wish the error messages were in the middle of the screen.</td>
</tr>
<tr>
<td>7</td>
<td>I wish all the error messages were in a single language.</td>
</tr>
</tbody>
</table>

In addition to the uncompleted tasks, most usability problems were found in the data entry fields. In fact, problems of the data entry category were related to the user-system interaction in order to enter the required information. In the study carried out by...
Haake et al. [19] in 2004, they found that most problems were encountered by inexperienced users who were at the stage of data entering and word perception. Assuming the possible sufficiency of education and work experience level of our users in general, we found that unfamiliarity of users with the system was the main reason for the most problems in this category. Therefore, it seems that designing and providing appropriate training courses will absolutely improve many of the problems we faced during the data entry process.

Despite the unfamiliarity issue, the inconsistency of the Persian language with the default language of the program was another usability problem. For example, one problem was about switching language between Persian and English. It usually caused typing errors and consequently wasted the users’ time (Fig 5).

Another major problem regarding the usability of the ADT system was the date format. 73% of users faced difficulty putting in the date of birth. The structural difference between the Persian language and the system languages is the main reason for that. For example, in the Persian language, the date can be written in eight different formats. As a result, the unawareness of users regarding the system’s acceptable date format, common errors, and the correction process was not only time-consuming but also frustrating. However, by working with the system continuously, the number of users who had problems with the date field was dropped to around 50%.

According to table 2, 36 problems were identified, and users could resolve nearly 50% of them and continue the process. However, for 22% of problems, users asked for the facilitator’s help. This indicated that the problem had a very strong effect on the working process and the user could not resolve it alone.

Our measurement (in section: Required materials in evaluation) showed that the sufficient time for performing the process completely is between 5-8 minutes. The average time to carry out the process for users who completely performed the entire assessment process was 14 minutes and 9 seconds; this time included the time to fix the problem as well. It’s obviously clear that with continuous use of the system and training sessions, users will have fewer problems to resolve and therefore the three minutes and one second of problem resolution will be eliminated.

63.63% of users expressed their negative feeling with facial and verbal expressions in different situations of the system, 25% used the phrase “What a slow system” to express their anger. Of course, the speed of the system was a technical issue and was related to the system architecture, and not related to usability. Some of them said “Oh! The date field again!” when they reached this field, while more than 50% of users used the expression “I wish” during the evaluation process. This shows the users’ willingness to provide suggestions in order to improve the system. It seems that early evaluation of the system (at the system design stage) may tend to more satisfaction of the users and efficiency of the system [2, 17, 21].

Another finding was that users did not use the system guide on a regular basis. Khajouei et al. [12] in their study on an emergency reception information subsystem reported that the guide had not been defined in the relevant subsystem. Meanwhile, in the evaluated system in our study, the guide was designed to be accessible on the home page of the system, but users - even when faced with a problem - did not use the guide at all.

The Study showed that correcting problems identified by the HIS evaluation will significantly
improve system usability and user satisfaction. The improvement will increase the investment return rate over time [22]. In the systematic review carried out by Ahmadian et al. [21] they stated that few studies have been conducted to assess Iran’s HIS. Despite the fact that without system usability evaluations, reduced efficiency will result [23].

Another important finding was that the evaluated HIS was installed in 50 non-teaching hospitals across the country, while most of the evaluations have been conducted on systems used in teaching hospitals in Iran and other countries. Ahmadian et al. [21] have proposed that evaluating HIS must be performed on systems used in non-teaching hospitals because maybe results of the evaluation in teaching or non-teaching hospitals will be varied.

Although user-based evaluation and the think-aloud method, which was used in this study, were the gold approach for evaluations [10], the evaluation is based on users’ attributes and their experience and knowledge. Therefore, some problems may not be found [24]. Consequently, in order to improve HIS, our study results have been cited in other related studies as a way to aid clinical software developers to design systems with the real user model.

It is essential for clinical software programmers to design systems based on real users’ mental models [5, 20, 25-27]. The majority of problems in this study were the data entry category emphasis that suitable user interface and ease of data entry were very important for ADT systems. Hence, designers should pay special attention to it in designing such systems.

Limitations

One limitation of this study was the lack of evaluation for users in the hospital while they were working the real situation of their daily jobs. There are many conditions in the real job environments that might affect both timings as well as problem occurrence. We tried to make a similar environment but we knew that we could not mimic all the context. Although in research carried out by Kaikkonen et al. [28] the same problems were found in both environments (laboratory and field testing). Even though there were differences in the frequency of problem findings between the contexts.

In this study, only 11 people participated in the evaluation. By employing more users, the evaluation result will be more realistic and generalized. However, previous studies [6, 9, 24] have shown that in the think-aloud method, 3 to 10 users are sufficient to find problems cost-efficiently.

This evaluation was conducted on the system that is used in 50 non-teaching hospitals. The results cannot be completely generalized to all non-teaching ADT systems that have been produced by other companies. However, due to the similarities that exist between ADT systems in hospital information systems, these results can be used to improve other systems [12].

CONCLUSION

Our study showed that the non-teaching hospital ADT subsystem has many problems in the interaction of real users with the system. Data entry was the most frequent problem that can be resolved with better entry forms design. Researchers have suggested the use of one of the user-based evaluation methods such as the think-aloud method to increase the reliability of the results in the study of other hospital systems. The method that has been employed in this study can be used as a guide in this regard.

ACKNOWLEDGEMENTS

The authors wish to thank all employees of the Department of Health Information Technology and Medical Records, School of Paramedical Sciences, Mashhad University of Medical Sciences particularly Dr Kimiafar and Dr Sarbaz for their sincere assistance.

AUTHOR’S CONTRIBUTION

All authors contributed to the literature review, design, data collection and analysis, drafting the manuscript, read and approved the final manuscript.

CONFLICTS OF INTEREST

The authors declare no conflicts of interest regarding the publication of this study.

FINANCIAL DISCLOSURE

No financial interests related to the material of this manuscript have been declared.

REFERENCES


between laboratory and comparison. Interacting with


