Journal of Biomedical Informatics 80 (2018) 43–51

Contents lists available at ScienceDirect

Journal of Biomedical Informatics

journal homepage: www.elsevier.com/locate/yjbin

# Augmentation of the think aloud method with users' perspectives for the selection of a picture archiving and communication system

Misagh Zahiri Esfahani[a], Reza Khajouei[b,c,*], Mohammad Reza Baneshi[d]

[a] Health Management and Economics Research Center, Iran University of Medical Sciences, Tehran, Iran
[b] Medical Informatics Research Center, Institute for Futures Studies in Health, Kerman University of Medical Sciences, Kerman, Iran
[c] Department of Health Information Sciences, Faculty of Management and Medical Information Sciences, Kerman University of Medical Sciences, Kerman, Iran
[d] Modeling in Health Research Center, Institute for Futures Studies in Health, Kerman University of Medical Sciences, Kerman, Iran

ARTICLE INFO

*Keywords:*
Picture archiving and communication system
User-computer interface
Radiology information system
Think aloud
Usability evaluation
User interaction

ABSTRACT

*Objectives:* Users attitude toward a picture archiving and communication system (PACS) and their interaction with this system are among the most important factors that influence its acceptance. This study aimed to augment the user's interaction with the user's perspective to select a usable PACS among three systems available on the market.

*Methods:* We augmented the think aloud (TA) usability evaluation method with the Post-Study System Usability Questionnaire (PSSUQ) to compare user interaction problems of three PACS user interfaces. Four radiologists and four internist physicians participated in this study. Usability characteristics including efficiency, effectiveness, learnability, error, and satisfaction were used to assess the usability of each PACS.

*Results:* There was a significant difference in efficiency ($p = 0.01$), effectiveness ($p = 0.005$), learnability ($p = 0.001$), and satisfaction ($p = 0.009$). However, no significant difference in the number of errors ($p = 0.18$), mouse clicks and keystrokes ($p = 0.12$), and the number of usability problems ($p = 0.6$) were observed among the three PACS systems studied.

*Conclusions:* This study showed that applying the proposed approach to augment TA with the user's perspective addresses almost all of the theoretical aspects of usability and can be employed to select the most usable PACS.

## 1. Introduction

The shift from hard copy film-based imaging to digital imaging has significantly changed the workflow in radiology departments and medical institutions. The picture archiving and communication system (PACS) is one of the most important medical imaging technologies that have contributed to digital radiography [1,2]. PACS systems use an electronic archive for short- and long-term storage, retrieval, and management of medical images, a secure network for distribution, and workstations or mobile devices for the presentation of medical images produced by various medical hardware modalities, such as X-rays, CT scans, MRIs, and ultrasound machines [3].

Limited financial resources are the major challenge for the selection, installation, and maintenance of PACS systems in healthcare organizations [4,5]. While PACS systems are expensive, they are among the most important medical applications and are capable of bringing high returns on investment [6]. The selection and implementation of a PACS also faces other challenges and obstacles, including difficulty in selecting an appropriate PACS due to the lack of awareness of the best selection criteria [7,8]. Currently, the selection is done through a multidimensional comparison, such as price and technical functionalities between the commercially available PACS systems developed by different vendors. Although various aspects of the systems are reviewed via this approach, some necessary concerns might remain unaddressed. It is important to clarify whether the selected software is easy to use since the software-user interaction plays an important role in the application performance [9,10]. Although having different functionalities might seem promising and lead to a higher rank in comparison, users may find them confusing, difficult, and sometimes even impossible to use. User-software interaction that fails to meet the users' needs reduces the efficiency, satisfaction, productivity, and accuracy [11,12]. Several studies have reported rejected or withdrawn health information systems due to difficulty of use [13–15]. Therefore, appropriate considerations must be given to these perspectives in order to enhance the successful adoption of a PACS [16].

Previous studies on the usability of PACS systems [17–24] identified

* Corresponding author at: Medical Informatics Research Center, Institute for Futures Studies in Health, Kerman University of Medical Sciences, Haft-Bagh Highway, P.O. Box: 7616911320, Kerman, Iran.
*E-mail address:* r.khajouei@kmu.ac.ir (R. Khajouei).

https://doi.org/10.1016/j.jbi.2018.02.017
Received 25 July 2017; Received in revised form 15 February 2018; Accepted 26 February 2018
Available online 28 February 2018
1532-0464/ 

problems that can make the interaction time consuming, causing delays in tasks, dissatisfaction, and frustration, preventing users from enjoying all of the benefits and functionalities of the system, as well as leading to more errors and difficulties in performing clinical analyses [25]. Furthermore, a recent review article showed the lack of studies on usability evaluations of PACS systems using formal evaluation methods involving the final users [25]. Some have either evaluated the user interfaces of a single PACS [19,21,22,24], or as a subsidiary part [23,26], but to the best of our knowledge only a few studies have specifically focused on and addressed the evaluation of the user interface of different PACS systems [27,28]. They have investigated the user's perspectives to compare PACS systems with no intention of applying this information for selection. Jorritsma et al. [18] investigated the user's perspectives and interactions in a comparative study for the selection of one of four PACS systems. The study was conducted on radiologists as a group of PACS users and used a webcam for collecting data in testing sessions and finally concentrated on the analysis of satisfaction and efficiency characteristics.

Recruiting different groups of users, employing a specialized tool as well as investigating other usability characteristics such as effectiveness, learnability, and errors when real users interact with the system can add to the existing knowledge and provide more insights into the design of a PACS user interface.

At the time of this study, Kerman University of Medical Sciences was in the process of selecting and purchasing a PACS. Three demonstration PACS systems were provisionally installed in three different medical centers for the purpose of comparison and evaluation. Since the users' interaction with a PACS and their view after the first interaction with the system have a major impact on the success and adoption of the PACS, this study was designed to compare users' interaction problems as well as perspectives about the three PACS systems in terms of different usability characteristics.

## 2. Methods

### 2.1. Design and procedure

Three common PACS applications, hereafter called A, B, and C, from Iranian PACS vendors were evaluated in this comparative study. One vendor supplied an imported PACS and the other two provided their own developed software.

The think aloud (TA) method and the Post-Study System Usability Questionnaire (PSSUQ) were used to study the user's interaction and perspective, respectively. We augmented these two methods to be able to measure all the characteristics contributing to a usability test.

TA is the gold standard of usability evaluation [29], concentrating on a user's cognition while interacting with a system. In this method, users are asked to verbalize their feelings, thoughts, and whatever else comes to mind while performing tasks on a series of predetermined scenarios. The task examples should be as realistic as possible and representative of end-user performances in daily life situations. During the session, there should be full audiotaping and/or video recording of the participants and, if possible, video recording of the computer screens to document all important information. Usability problems are detected by evaluators from analyses of user behavior and expressions during interactions with the system [30,31]. The participants should be a sample of users representing the expected end users.

The PSSUQ consists of 19 items that were designed for immediate administration after usability testing [32]. The PSSUQ utilizes the 5-point Likert scale from "strongly agree" to "strongly disagree" and has two text fields for any comment and explanation by the participants. The PSSUQ was first translated into Persian and, to approve the cross-cultural comparison of translation, was re-translated into English by an expert and then its compliance with the original text was confirmed. The content validity of the PSSUQ was confirmed by one radiologist and three medical informatics experts. The reliability was determined using Cronbach's alpha ($\alpha = 0.96$).

To define the measurement criteria, a coding framework was developed according to five usability characteristics and based on the International Organization for Standardization (ISO) and Nielsen's definitions [33–35] to recognize the specific user-computer interaction problems in detail. According to the ISO, usability is defined as "the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of use." Nielsen put forward five usability attributes: learnability, efficiency, memorability, errors, and satisfaction [35]. Combining ISO and Nielsen usability attributes yields the following six criteria: efficiency, effectiveness, learnability, memorability, errors, and satisfaction. Since the participants in this study used each system only once and there was no need to remember the options for a next session, we did not consider memorability in our evaluation. The remaining five attributes composed our coding framework.

We used the TA method to measure effectiveness, learnability, errors, and efficiency characteristics, and the PSSUQ was used to measure satisfaction.

### 2.2. Participants

Since five to eight participants are considered sufficient to perform TA [36,37], we recruited eight participants from two user groups. The participants included four radiologists and four internist physicians from a university hospital in Kerman, Iran. Since users mostly interact with a PACS via PACS viewers, the focus of this usability study was on evaluating the users' interaction with the PACS viewers.

Three evaluators acted as facilitators of the testing sessions and analyzed the results.

### 2.3. Testing protocol and data collection

TA sessions were held in the physicians' actual workplace. After training the participants with the TA method in 10-min sessions, they were given five scenarios containing seven to nine tasks. The participants were provided with TA instructions and the clinical information for each scenario. A radiologist and an internist physician, both with at least one year of experience working with PACS, were consulted for the design of scenarios. These scenarios contained different modalities, including two digital radiography (DR), two computed tomography (CT), and one magnetic resonance imaging (MRI). Both the radiologist and the internist physicians used the same scenarios except for the MRI containing study, which was run only by the radiologist, as it was a brain tumor case and irrelevant to the internist physicians. The scenarios were designed in a manner to examine different parts and functions of a PACS and covered the most common tasks that a clinician may use in a typical working application. Generally, the scenarios included the following tasks: lesion size measurement, densitometry (Hounsfield unit), contrast change and window level, zoom, magnification, and layout use (observation of images in different cuts and views).

The scenarios and tasks were offered to the radiologists and internists. Table 1 illustrates the scenarios, goals, and actions needed to complete the tasks.

A CD containing the medical images of patients and consistent with the scenarios was played for the physicians during each TA session. The ID number of each image matched the scenario number, allowing the physicians to perform the tasks via case-by-case patient selection.

The clinical images of the patients used in the scenarios were collected from the hospital database. All patient-identifying information was deleted from the PACS images to maintain confidentiality. The study was reviewed and approved by the ethics committee of Kerman University of Medical Sciences (IR.kmu.REC.1394.454).

Interaction with the system was done through a standard mouse and a keyboard. Capturing the video, audio, and all of the activities on the