Usability Evaluation of a Body Surface Potential Map Visualization System

Raymond R Bond, Dewar D Finlay, Chris D Nugent, George Moore

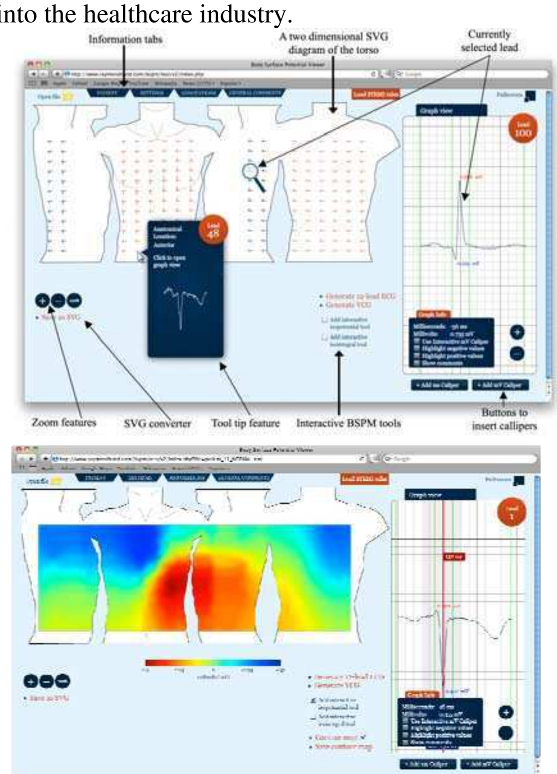
University of Ulster, Belfast, Northern Ireland, United Kingdom

Abstract

Although the Body Surface Potential Map (BSPM) is considered to be an improved technique for the detection of cardiac pathologies when compared to the 12-lead electrocardiogram (ECG), it has not been adopted for routine clinical practice. This is partly due to a lack of clinician-friendly tools for visualizing the large amount of data a BSPM contains. This work presents the details of a usability evaluation of a BSPM viewer. This evaluation was conducted at the 2010 Computing in Cardiology meeting. Seven delegates (7 males, age:39±11) were recruited to take part in the evaluation. The protocol required the participants to complete a pre-test questionnaire, attempt a series of tasks whilst 'thinking-aloud' and at the end of the session, complete a post-test questionnaire. Based on 3.5 hours of video and audio recordings (circa), 41 use errors were discovered and given a severity rating (mean:2.24±1.09, 1=cosmetic, 4=critical). Using a five-star scale, participants rated the responsiveness of the viewer (mode: 4/5), its usefulness (mode:3/4), learnability (mode:3/4), and its 'look and feel' (mode:5). In conclusion, a usability evaluation was carried out in a conference setting to assess and improve the interface of a BSPM visualization system.

1. Introduction

The Body Surface Potential Map (BSPM) is more sensitive in detecting cardiac abnormalities when compared to the 12-lead electrocardiogram (ECG) [1]. The BSPM has, however, not been widely adopted for a number of reasons. One reason is the lack of tools for managing and visualizing BSPMs [2]. In addressing this problem, the authors have developed a storage format referred to as XML-BSPM and a compatible BSPM viewer [3, 4]. The previously reported BSPM viewer has been enhanced as part of the work presented in this paper. The main interface of the viewer is shown in Figure 1a. The viewer entails interactive tools for exploring BSPM data. These include isopotential and isointegral mapping tools. In this enhanced version of the BSPM viewer, fully colored contour maps can be drawn (Figure 1b). The objective of developing the viewer is to be 'clinician-friendly'. In this regard, the viewer's usability has been evaluated in this paper. A usability evaluation, can be described as "a quality attribute that assesses how easy user interfaces are to use" [5]. Medical interfaces, in particular, should be usable given that in some cases counter-intuitive interfaces can cause fatal medical errors [6]. A usability evaluation can be used to uncover use errors prior to clinical practice. Once use errors have been identified, they can be 'designed-out' of the interface and usability can be 'designed-into' the interface. The National Institute for Standards and Technology (NIST) are seeking a standard to assess the usability of health IT systems [7]. The importance of usable medical interfaces has also been recognized by an international standard, namely the 2007 ISO/IEC 62366: Medical Devices-Application of Usability Engineering to Medical Devices [8]. This standard requires manufactures to evaluate the usability of medical interfaces prior to their introduction into the healthcare industry.



(a)

(b)

Figure 1. BSPM viewer (a) shows the interface of the viewer and (b) shows an example isopotential contour map generated at the time instant of the R peak.

2. Methods

The usability testing method

There are a number of usability evaluation methods, for example heuristic, cognitive walkthrough and usability testing, to name but a few [5]. The usability testing method was selected for use in this evaluation as it has seen successful utility in similar work [9]. This method involves recruiting potential users to attempt representative tasks using a given product. This allows the developers to record the use errors encountered by each of the participants. A usability test is traditionally carried out in a laboratory setting. This can, however, be problematic if a specialized group of experts are required. To get around this issue, the BSPM viewer was evaluated at the 2010 International Conference on Computing in Cardiology, which happens to attract experts in the subject area, i.e. electrocardiology and mapping [10].

Recruiting participants

The study was advertised in the conference delegate packs. Seven delegates (7 males, mean age:39±11 years old) subsequently volunteered to take part in the usability evaluation. A large number of participants are not necessary for this type of evaluation. This has been illustrated in previous studies where just five participants identified approximately 80% of use errors [5].

Table 1. Summary of participants (*n*=7) profile.

Demographic	Data
Gender	7 males
Age	Mean 39 ± 11 years old
Occupation	6 academics. 1 participant was based in industry and academia.
Primary expertise	2 Electrical Engineers, 2 Computer Scientists, 2 Medical Doctors and 1 Medical Physicist
First language	4 English, 2 German, 1 Dutch
Average level of computer literacy	Mean 4.7 ± 0.5 Mode:5
Average computer usage per week	Mean 34 ± 3 hours
Average Internet usage per week	Mean 19 ± 13 hours
Average level of ECG expertise	Mean 4 ± 0.4 Mode: 4 (where 1 is low and 5 is high)
Average level of BSPM expertise	Mean 3.6 ± 1 Mode: 4

Pre-test questionnaire

Subsequent to signing the voluntary consent form, each participant completed a pre-test questionnaire, which is available online [11]. This questionnaire consisted of ten questions that were used to profile the participants. The demographics obtained from this questionnaire have been summarized in Table 1.

Representative tasks

Following the pre-test questionnaire, each participant attempted three small tasks and four elaborate tasks (Table 2). All tasks were given to the participants in the same order. These tasks are 'representative' as the authors considered the tasks to be fundamental to the BSPM viewer.

Table 2. Representative tasks given to the participant.

Task no.	Task wording (n=7)
Small task 1	Find the patient's name.
Small task 2	What is the width of the QRS?
Small task 3	What is the width and height of the T wave in lead 75?
Elaborate task 1	Find the lead with the highest R wave.
Elaborate task 2	Which lead has the highest elevation at ST60 (J point + 60 milliseconds).
Elaborate task 3	Visualize the entire ST segment and export this image to a file.
Elaborate task 4	Leave a comment over the T wave in lead 78 for a colleague to read.

The hardware and software

A desktop computer, digital microphone and screen recording software were the only hardware and software items used to carry out this evaluation. The computer and the Web browser were used to access the BSPM viewer. The computer was located in a secluded area to avoid pre-conditioning other delegates passing by and to avoid unwanted disruption. Whilst the participants attempted each of the tasks, the microphone and screen recording software were used to record the participant's verbal feedback and their interactions with the BSPM viewer.

Think-aloud protocol

Each participant was asked to 'think-aloud' whilst attempting the tasks and a microphone was used to record the audio. The 'concurrent think-aloud' protocol helps gain insight into the user's working memory and to how they actually solve problems. Although it originated in the field of cognitive psychology, it was later used by Clayton Lewis [12] in 1982 for evaluating computer interfaces at IBM.

Post-test questionnaire

At the end of each session, each participant completed a post-test questionnaire which is available online [11]. This questionnaire was used to obtain a number of subjective opinions regarding the efficiency, usefulness, learnability and user satisfaction with the software.

3. Results

Using the microphone and screen recording software, approximately 3.5 hours of audio and video were obtained during the entire evaluation. Whilst watching these recordings, the use errors were transcribed into a report. Each use error in the report was assigned a severity rating, i.e. 1 (Cosmetic), 2 (Medium), 3 (Serious) and 4 (Critical). The severity ratings were assigned from a consensus involving two usability experts. The frequency (the number of users that encountered the same use error) partly influenced the severity ratings. Rating the severity of use errors is important as it prioritizes ensuing development work. From this evaluation, 41 unique use errors were identified in the BSPM viewer. The severity rating of these use errors is presented in Figure 2. The average severity rating is 2.24 (±1.09).

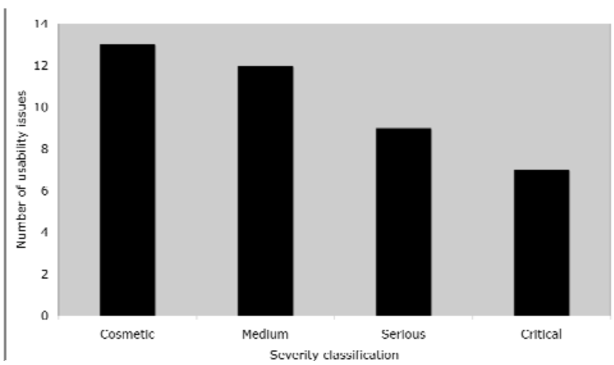


Figure 2. Severity rating of use errors identified in the BSPM viewer.

Table 3. Four examples of use errors.

Use error	Severity & freq.	Solution
User struggled to find the relevant feature to change the time instant of the isopotential tool.	*Severity*: Critical *Freq*: 3	Make the relevant feature more prominent.
The clicking hot spot area is not big enough to click on the leads.	*Severity*: Serious *Freq*: 2	Enlarge the hot spot areas for selecting a lead.
The isopotential time needle does not snap onto beat markers.	*Severity*: Medium *Freq*: 1	Add a 'snap to beat markers' option.
User was confused by the technical term 'SVG'.	*Severity*: Cosmetic *Freq*: 1	Replace the text 'Save as SVG' to 'Save as image'.

An excerpt of the use errors can be found in Table 3. This table presents a description of the use error, the number of times that particular error occurred (frequency), the severity rating and a potential solution. Alongside the identification of use errors, the efficiency, usefulness, learnability and the user satisfaction with the BSPM viewer were also evaluated. This was largely informed by the post-test questionnaire. Efficiency was measured by the task completion rates and the task completion times. The task completion times were calculated using time-stamps in the video recordings. The completion times of the elaborate tasks can be viewed in Table 4.

Table 4. The completion rates and times of the elaborate tasks carried out using the BSPM viewer.

Elaborate task	Completion rate (%)	Average completion time (mins)
1	100	2.59
2	100	0.54
3	100	1.10
4	100	1.35

All four elaborate tasks were completed by all participants without any assistance. The average completion time for elaborate task one was, however, excessive. It is unreasonable to expect a user to engage with any ECG application for three minutes just to locate the ECG lead with the highest R wave. This excessive completion time was likely due to a lack of information hierarchy. That is, the appropriate tool needed to complete the task was not easily accessible. Elaborate task two was very similar to the elaborate task one, in that they both require the same tool (i.e. the isopotential tool) to complete the task. Given that the user was acquainted with the appropriate tool (i.e. the isopotential tool) from the first task, the average completion time for the second task was only 54 seconds. A *t*-test was used to support the hypothesis that elaborate task one can be carried out much more efficiently once the user has knowledge of where to find the appropriate tool. The *t*-test was used to test the difference of the means of tasks one and two. This resulted in a statistical significance at the 5% level (p-value 0.017). The completion rates and times for tasks three and four were also considered to be satisfactory given that first time users carried them out. Efficiency was also evaluated using the post-test questionnaire. Participants rated the 'responsiveness' of the BSPM viewer between 3 and 5 (mode: 4/5, mean: 4.28) on a five point semantic differential scale (where 1 denotes low and 5 denotes high).

The usefulness of the viewer was evaluated by the post-test questionnaire. Participants were asked how useful the viewer would be in the ECG clinical domain (mode: 4, mean: 4). Usefulness was also assessed by a number of polar questions (Figure 3). All participants said a BSPM viewer is needed to further research in the ECG domain, six out of seven of participants said the viewer would be useful in education and six out of seven participants said they would recommend the BSPM viewer.

The learnability was assessed by a number of questions in the post-test questionnaire. Participants were asked to rate how easy the application was to learn. They gave a rating between 3 and 5 (mode:3/4, mean:3.7) on the semantic differential scale. Learnability was also assessed by several polar questions. Five out of seven participants said they could remember how to use the viewer if seen for a second time, four participants indicated that they could learn the application without formal training and five said they could learn all of the features in the BSPM viewer within a period of hours as opposed to days, weeks or months.

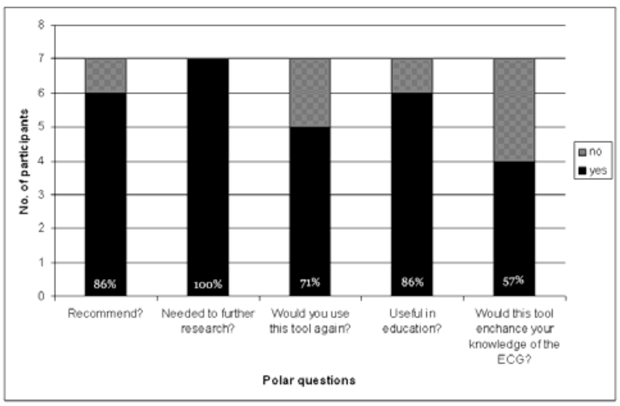


Figure 3. Results from a series of polar questions.

The satisfaction of the participants and their views in relation to the interface aesthetics were assessed in the post-test questionnaire. Although this parameter is subjective, user satisfaction has been associated with 'high control' of an application [5]. Participants rated the 'look and feel' of the BSPM viewer on the semantic differential scale (mode:5, mean:3.85). Moreover, four out of seven participants said the color scheme was good and six said the interface layout was intuitive.

4. Conclusion

The usability of medical interfaces can be a matter of life or death, and an evaluation process is essential to improve the usability of a medical interface. The BSPM viewer can now be described as a usable system. Although a BSPM viewer may in itself help streamline the BSPM into routine clinical practice, a 'usable' BSPM viewer will allow clinicians to use the viewer more effectively and efficiently and therefore provide better support for accurate diagnostic decisions.

This usability evaluation employed the 'usability testing method'. To the best of our knowledge, we are the first to use it in a conference setting. Given a usability evaluation framework has yet to be standardized [7], different methodologies (e.g. usability testing in a conference setting) need to be described, made available and compared with other methods.

References

[1] Ornato JP, Menown IBA, Peberdy MA, Kontos MC, Riddell JW, Higgins GL, et al. Body surface mapping vs 12-lead electrocardiography to detect ST-elevation myocardial infarction. Am J Emerg Med 2009;27:779-784.

[2] Medvegy M, Duray G, Pinter A, Preda I. Body surface potential mapping: historical background, present possibilities, diagnostic challenges. Annals of Noninvasive Electrocardiology 2002;7:139-151.

[3] Bond R, Finlay D, Nugent C, Moore G. XML-BSPM: an XML format for storing Body Surface Potential Map recordings. BMC Medical Informatics and Decision Making 2010;10:1-26.

[4] Bond RR, Finlay DD, Nugent CD, Moore G. A Web-based tool for processing and visualizing body surface potential maps. J Electrocardiol 2010;43:560-565.

[5] Nielsen J. Usability engineering. 1st ed. London: Academic Press; 1993.

[6] Koppel R, Metlay JP, Cohen A, Abaluck B, Localio AR, Kimmel SE, et al. Role of computerized physician order entry systems in facilitating medication errors. JAMA 2005;293:1197.

[7] BiomedMe. National Institute for Standards and Technology (NIST) to evaluate ease of use of Health IT systems. Available from: http://biomed-me.com/general/national-institute-for-standards-and-technology-nist-to-evaluate-ease-of-use-of-health-it-systems_6213.html

[8] ISO/IEC. IEC 62366:2007 Medical devices - Application of usability engineering to medical devices. Available from: http://www.iso.org/iso/catalogue_detail.htm?csnumber=38594

[9] Jaspers MWM. A comparison of usability methods for testing interactive health technologies: Methodological aspects and empirical evidence. Int J Med Inf 2009;78:340-353.

[10] CINC. Computing in Cardiology. Available from: http://www.cinc2010.org/

[11] Bond RR. Pre and post-test questionnaires. Available from: http://www.raymondbond.com/survey/

[12] Lewis C. Using the" thinking-aloud" method in cognitive interface design. IBM TJ Watson Research Center; 1982.

Address for correspondence.

Raymond Bond (Room 16J26),
University of Ulster (UUJ),
Shore Road,
Newtownabbey,
Co. Antrim,
BT370QB,
bond-r@email.ulster.ac.uk