

A Avaliação da Usabilidade de um Sistema Médico Inteligente: BCCT.core

Ana Paula Afonso
Informática, ISCAP – Instituto Superior de Contabilidade e Administração do Porto
apafonso@iscap.ipp.pt

Jaime S. Cardoso
Engenharia Informática, FEUP – Faculdade de Engenharia da Universidade do Porto
jaime.cardoso@inescporto.pt

Maria João Cardoso
FMUP – Faculdade de Medicina da Universidade do Porto
micard@med.up.pt

Manuel Pérez Cota
Departamento de Informática UV - Universidade de Vigo
mpcota@uvigo.es

RESUMO

A avaliação estética da mama é um requisito fundamental no tratamento conservador do cancro da mama.

Para ultrapassar as limitações de reprodutibilidade dos métodos subjectivos correntes, foi recentemente desenvolvido um sistema inovador, assistido por computador, denominado BCCT.core (**B**reast **C**ancer **C**onservative **T**reatment. **c**osmetic **r**esults), que prevê, automaticamente, o resultado estético do tratamento conservador do cancro da mama após intervenção [1].

Neste trabalho, analisamos alguns dos aspectos mais relevantes da interface do BCCT.core, relacionados com o estudo da interacção com o utilizador, através de um questionário dirigido a utilizadores reais.

O desempenho do sistema pôde assim ser avaliado por critérios claros, permitindo avaliar a percepção humana relativamente à interface. De uma forma geral pode concluir-se que a utilização de técnicas de IHC¹ na avaliação da funcionalidade e usabilidade do BCCT.core foram essenciais no sentido de o tornar mais usável, eficiente e com elevado grau de satisfação, neste caso particular, para o utilizador médico.

Palavras-chave: Usabilidade, IHC, Interfaces, QUIS.

1. INTRODUÇÃO

A omnipresença da ciência dos computadores tem impulsionado grandemente a investigação nas mais diversas áreas científicas. A medicina tem sido uma das áreas mais beneficiadas tornando o diagnóstico médico e a evolução dos tratamentos cada vez mais precisos e eficientes [2]. A avaliação do resultado estético do tratamento conservador do cancro da mama é um exemplo concreto desta aplicação.

Para ultrapassar as limitações de reprodutibilidade e praticabilidade dos métodos subjectivos correntes, foi recentemente desenvolvido um sistema inovador, assistido por computador para uso médico, denominado BCCT.core (**B**reast **C**ancer **C**onservative **T**reatment. **c**osmetic **r**esults). Este sistema avalia o resultado estético do tratamento conservador do cancro da mama após intervenção [1].

O BCCT.core introduz a medição automática de diversos índices relacionados com o resultado estético da intervenção cirúrgica, tornando a avaliação rápida, fácil e reprodutível. As medições são precedidas pela localização automática de alguns pontos de referência (mamilo, contorno mamários, sternum notch) na fotografia digital. As medidas de assimetria, diferença de cor entre mama tratada e não tratada, e visibilidade da cicatriz cirúrgica são posteriormente baseadas nestes pontos de referência. Todas estas medidas ficam imediatamente visíveis para o utilizador e são gravadas numa base de dados (Figura 1). Simultaneamente, o conjunto de medidas é automaticamente convertido numa avaliação global e objectiva do resultado estético, baseada num algoritmo de classificação automática. Este algoritmo foi treinado e optimizado para prever o resultado estético global numa escala de quatro classes (excelente, bom, razoável, mau). Esta

classificação é também disponibilizada automaticamente ao utilizador e gravada na base de dados. A avaliação objectiva fornecida pelo BCCT.core é um resumo do resultado estético global, que permite a comparação efectiva entre diferentes equipas de trabalho de quaisquer centros.

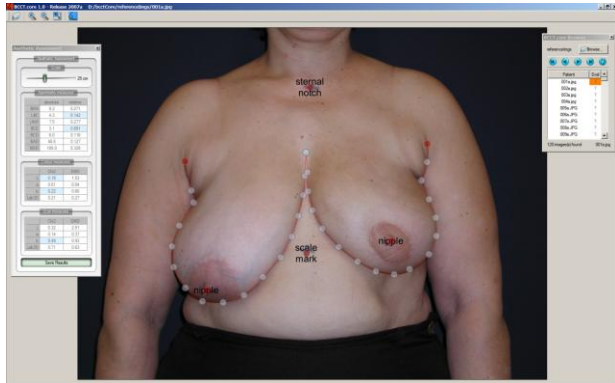


Figura 1. A interface do BCCT.core, versão 1.0.

Este sistema computacional é utilizado por profissionais médicos, com poucos conhecimentos na operação com sistemas deste tipo; como tal a interface é fundamental para a aceitação e sucesso do sistema.

Tendo como objectivo desenhar interfaces eficazes é fundamental envolver os utilizadores durante todo o processo, centrando nestes o desenho do sistema.

Neste trabalho analisamos a interacção do utilizador com o BCCT.core e avaliamos a funcionalidade e usabilidade, com base em técnicas de questionário a fim de identificar problemas específicos relacionados com o desempenho visando aumentar a satisfação do utilizador.

2. A INTERACÇÃO HOMEM-COMPUTADOR

A interface de um sistema computacional deve facilitar a interacção com o utilizador, tornando-a natural, fácil de usar e intuitiva.

A área da ciência dos computadores que estuda a interacção entre pessoas e sistemas computacionais denomina-se Interacção Homem-Computador (IHC), sendo uma área multidisciplinar que envolve as áreas da Ciência da Computação, Psicologia, Factores Humanos, Linguística, de entre outras. A IHC impulsiona a aplicação do conhecimento destas disciplinas no desenvolvimento e implementação de interfaces “amigáveis” (*user-friendly*), permitindo que o utilizador possa realizar as suas tarefas com

segurança, eficiência, eficácia e satisfação, ou seja, originando interfaces com elevado grau de usabilidade.

A norma internacional ISO 9241-11 define usabilidade como sendo um conceito que descreve até que ponto um sistema computacional pode ser usado pelos utilizadores de forma a atingir objectivos específicos como eficácia, eficiência e satisfação num dado contexto [3].

A **eficácia** permite que o utilizador atinja os objectivos iniciais de interacção, e é avaliada em termos de cumprimento de uma tarefa como também em termos de qualidade do resultado obtido.

Eficiência refere-se à quantidade de esforço e recursos necessários para se chegar a um determinado objectivo. Os desvios que o utilizador faz durante a interacção e a quantidade de erros cometidos podem servir para avaliar o nível de eficiência do sistema.

A terceira medida de usabilidade, a **satisfação**, é a mais difícil de medir e quantificar pois está relacionada com factores subjectivos. De uma maneira geral, satisfação refere-se ao nível de conforto que o utilizador sente ao utilizar a interface.

3. TÉCNICAS DE AVALIAÇÃO DA USABILIDADE

Como já referido, um dos grandes desafios no desenvolvimento de sistemas computacionais usáveis por pessoas é saber como fazer a transição entre o que pode ser feito — **funcionalidade** — e como deve ser feito — **usabilidade** — com o objectivo de satisfazer as necessidades do utilizador.

A necessidade de avaliação da qualidade da interacção homem-computador, ou mais concretamente da usabilidade do sistema, surge para verificar se a concepção do sistema coincide com as necessidades reais dos utilizadores.

A avaliação da usabilidade permite identificar problemas específicos na interface de um sistema [4].

De acordo com Dix et al. [5] e Rocha & Baranauskas [6] a avaliação tem como objectivos principais:

- **Avaliar a funcionalidade do sistema**, ou seja, medir a eficiência na execução das tarefas pelo utilizador;
- **Avaliar a satisfação do utilizador relativamente ao uso da interface**, ou seja, verificar o grau de facilidade de aprendizagem, flexibilidade e robustez;
- **Identificar problemas específicos do funcionamento geral do sistema**, com base na

verificação de aspectos de *design* da interface que contribuam para uma percepção incorrecta na execução das tarefas.

3.1 A escolha de um método de avaliação

Existe uma grande diversidade de métodos para avaliação de sistemas interactivos. De acordo com Dix et al. [5] existem pelo menos oito factores que devem ser considerados no momento da escolha de um método de avaliação, nomeadamente:

- O estágio de desenvolvimento do sistema;
- O estilo de avaliação;
- O nível de subjectividade ou objectividade da técnica;
- O tipo de medidas fornecidas;
- A informação fornecida;
- A imediaticidade da resposta;
- O nível de interferência implicado;
- Os recursos requeridos.

Existem técnicas baseadas em processos cognitivos e nos princípios de usabilidade, que como não requerem a intervenção directa do utilizador não prevêm a utilização real do sistema, limitando-se a preservar os aspectos relacionados com a usabilidade da interface e são por isso designados como métodos de inspecção da usabilidade [5]. Estas são de facto úteis na detecção de áreas da interface a serem refinadas, mas não substituem os métodos que recorrem aos utilizadores da própria aplicação.

Estas avaliações, centradas no utilizador, têm como premissa que a satisfação das necessidades do utilizador é um factor determinante para a aceitação e conseqüente sucesso do sistema. Estas abordagens são designadas por testes de usabilidade e baseiam-se na utilização de medidas quantitativas de desempenho.

Em geral, os testes de usabilidade são considerados como os métodos mais eficazes na detecção de problemas.

Os testes de usabilidade recorrem à utilização de questionários ou entrevistas, técnicas que envolvem directamente o utilizador questionando-o sobre o funcionamento do sistema.

Os dados obtidos através de entrevistas tendem a ser qualitativos enquanto os dados obtidos a partir de questionários tendem a ser quantitativos. Estes últimos têm a vantagem de permitir atingir um grande número de pessoas, conseguindo-se assim resultados estatísticos relevantes [7].

4. QUESTIONÁRIOS

A comunidade científica colocou à disposição vários questionários dedicados à avaliação da usabilidade [4], amplamente estudados, com propósitos diferentes, funcionando como linhas orientadoras em casos específicos.

Podemos referir, a título de exemplo, alguns dos mais divulgados como o QUIS [8], SUMI [9], WAMMI [10], SUS [11].

No caso particular do sistema em estudo — o BCCT.core — a sua divulgação, aos potenciais utilizadores, durante um workshop, permitiu obter, através do preenchimento de um questionário (informal), alguma informação acerca da aceitabilidade do software. A análise de resultados mostrou que a maioria dos participantes tinha interesse na sua aquisição, dadas as suas reais possibilidades em participar activamente na avaliação final do resultado cosmético. A interacção com o software mostrou ser positiva mas pouco conclusiva dada a inexperiência geral na sua utilização.

Após a aquisição e experimentação do software, por equipas médicas nacionais e internacionais, foi construído um novo questionário, dirigido especificamente a questões de usabilidade do sistema.

A avaliação consistiu, numa primeira fase, na análise dos questionários disponíveis, numa tentativa de selecção ou eventualmente na utilização de mais do que um em simultâneo.

De entre os vários modelos, optámos pelo QUIS – *Questionnaire for User Interaction Satisfaction* – desenvolvido pelo Human Computer Interaction Laboratory (HCIL) na Universidade de Maryland em 1997, após verificarmos que a organização de questões no QUIS, per si, permitia obter respostas às questões que pretendíamos estudar e era também o que mais se aproximava do contexto em questão.

Este questionário tem como finalidade estimar a satisfação subjectiva dos utilizadores focando aspectos específicos da interface. O QUIS tem demonstrado alta eficácia e fiabilidade sendo um questionário amplamente utilizado na avaliação de interfaces, tratando-se portanto de um instrumento bem testado [12].

O questionário QUIS original é bastante longo, de modo que visando reduzir o esforço dos utilizadores foi criado um novo baseado neste e dirigido especificamente à aplicação BCCT.core.

5. RESULTADOS

De acordo com Stone et al [13] em geral, não são necessários muitos participantes para empreender uma avaliação. Na prática, os participantes iniciais conseguem fornecer as informações requeridas pelos avaliadores e estima-se que cerca de cinco participantes são frequentemente suficientes.

Neste caso o sistema foi testado com dez utilizadores, sendo oito médicos, de diferentes países, e dois engenheiros informáticos que colaboraram na concepção da aplicação.

O questionário foi finalmente enviado aos utilizadores regulares do BCCT.core. As respostas obtidas foram compiladas numa folha de cálculo, onde foi possível obter resultados objectivos, clarificando assim as questões mais prementes em termos de usabilidade.

Analisadas as respostas, podemos resumir o resultado da avaliação nos seguintes aspectos:

- **Experiência com o sistema** — A maioria dos utilizadores usa o BCCT.core há mais de 6 meses, ou seja, estavam bastante familiarizados com a interface; despendem em média 1 a 4 horas por dia na sua utilização; e utilizam outros sistemas operativos além do Windows.
- **Experiência anterior** — Grande parte dos utilizadores conhece a maioria dos periféricos relacionados com equipamentos informáticos (em média 64%).
- **Reacção geral ao sistema** — A reacção geral ao software é muito positiva (em média 68%).
- **Funcionalidades** — De uma forma geral os utilizadores estão satisfeitos com as funcionalidades disponibilizadas, embora sentindo necessidade de cálculos adicionais, nomeadamente de médias e desvios padrões, entre outros.
- **Utilização do BCCT.Core** — A utilização revelou-se altamente positiva, já que 90% dos utilizadores mostraram intenções de continuar a usá-lo no futuro.
- **Tutoriais em linha** — Em média, 80% dos utilizadores consideraram que o tutorial é muito útil; o acesso e a navegação através do tutorial são simples; a informação é concisa e de leitura directa e as instruções para completar as tarefas são claras.
- **BCCT.core Video Demos** — As demonstrações foram consideradas como muito úteis por 83% dos utilizadores.

6. CONCLUSÕES

De uma forma geral pode concluir-se que a utilização de técnicas de interacção Homem-Computador na avaliação da funcionalidade e usabilidade do BCCT.core foram essenciais no sentido de o tornar mais usável, eficiente e com elevado grau de satisfação, neste caso particular, para o utilizador médico.

O desempenho do sistema pôde assim ser avaliado por critérios claros, apoiado numa base sólida que permite ter noção da percepção humana relativamente à interface, disponibilizando os meios necessários ao desenvolvimento de sistemas que apoiam efectivamente os utilizadores na realização de tarefas.

7. REFERÊNCIAS

- [1] Cardoso J. S., Cardoso M. J. Towards an intelligent medical system for the aesthetic evaluation of breast cancer conservative treatment. *Artificial Intelligence in Medicine*, pp. 115-126, vol. 40, 2007
- [2] Müller, H., Michoux, N., Bandon, D., and Geissbuhler, A. A review of content-based image retrieval systems in medical applications-clinical benefits and future directions. *International Journal of Medical Informatics* 73, 1 (2004), 1-23.
- [3] ISO 9241-11. Ergonomic requirements for office work with visual display terminals (VDTs) -- Part 11: Guidance on usability (1998)
- [4] Preece, J., Rogers, Y.; Sharp, H. *Interaction Design: Beyond Human-computer Interaction*. 2nd ed. 2007. John Wiley & Sons Inc.
- [5] Dix, A., Finlay, J., Abowd, G.D., and Beale, R. *Human-Computer Interaction*. 3 ed. Prentice Hall, Essex, 2004.
- [6] Rocha, H.V.M.; Baranauskas, C.C. Design e Avaliação de Interfaces Humano-Computador. NIED/UNICAMP, Campinas, SP, Brasil, 2003.
- [7] Preece, J., Rogers, Y., and Sharp, H. *Human-Computer Interaction*. Addison-Wesley, Wokingham, 1994.
- [8] Questionnaire for User Interaction Satisfaction. <http://lap.umd.edu/QUIS/>.
- [9] Software Usability Measurement Inventory. <http://sumi.ucc.ie>.
- [10]. Website Analysis and MeasurMent Inventory. <http://www.wammi.com>.
- [11] System Usability Scale. <http://www.mindd.com>.
- [12] Shneiderman, B. *Designing the User Interface – Strategies for effective Human Computer-Interaction*. 3rd ed. 1998. Addison Wesley Longman, Inc.

[13] Stone, D., Jarret, C., Woodroffe, M., and Minocha, S. *User Interface Design and Evaluation*. Morgan Kaufmann, San Francisco, 2005.

¹ IHC – Interação Homem-Computador