

JANEIRO / MARÇO 2011

ARTIGOS ORIGINAIS

Protótipo de um sistema de gerenciamento de protocolos de câncer colorretal

Management system prototype of colorectal cancer protocols

HUEI DIANA LEE¹, LUIZ HENRIQUE DUTRA DA COSTA², CARLOS ANDRES FERRERO³, CLÁUDIO SADDY RODRIGUES COY⁴, JOÃO JOSÉ FAGUNDES⁵, RENATO BOBSIN MACHADO⁶, FENG CHUNG WU⁷

¹Professora Doutora e Coordenadora Geral do Laboratório de Bioinformática (LABI), Universidade Estadual do Oeste do Paraná (UNIOESTE), Foz do Iguaçu, PR, Brasil. ²Membro do Laboratório de Bioinformática (LABI), Universidade Estadual do Oeste do Paraná (UNIOESTE), Foz do Iguaçu, PR, Brasil. ³Professor Mestre e Pesquisador do Laboratório de Bioinformática (LABI), Universidade Estadual do Oeste do Paraná (UNIOESTE), Foz do Iguaçu, PR, Brasil. ⁴Professor Doutor do Departamento de Cirurgia (D.M.A.D.), Serviço de Coloproctologia da Universidade Estadual de Campinas (UNICAMP), Campinas, SP, Brasil. ⁵Professor Doutor do Departamento de Cirurgia da Faculdade de Ciências Médicas da UNICAMP, Campinas, SP, Brasil. ⁶Professor e Coordenador da Área Computacional do Laboratório de Bioinformática (LABI), Universidade Estadual do Oeste do Paraná (UNIOESTE), Foz do Iguaçu, PR, Brasil. ⁷Professor Participante e Pesquisador do Serviço de Coloproctologia da Faculdade de Ciências Médicas da UNICAMP, Campinas, SP, Brasil. Professor Doutor da UNIOESTE e Coordenador da Área Médica do Laboratório de Bioinformática (LABI), Universidade Estadual do Oeste do Paraná (UNIOESTE), Foz do Iguaçu, PR, Brasil.

LEE HD; COSTA LHDD; FERRERO CA; COY CSR; FAGUNDES JJ; MACHADO RB; WU FC. Protótipo de um Sistema de Gerenciamento de Protocolos de Câncer Colorretal. *Rev bras Coloproct*, 2011;31(1): 1-7.

RESUMO: **Objetivo:** desenvolvimento de um protótipo de sistema de cadastro e controle de protocolos de Câncer Colorretal, com a finalidade de armazenar um conjunto abrangente de dados de forma estruturada para posterior aplicação de métodos de análise inteligente de dados. **Material e Método:** o protótipo foi construído em cinco etapas: análise do domínio de Câncer Colorretal, definição de requisitos básicos (operacionais e informacionais), projeto do sistema, construção do sistema usando tecnologias *open source* e avaliação do sistema com o apoio de especialistas. **Resultados e Discussão:** durante o desenvolvimento, o protótipo foi avaliado por especialistas quanto à estrutura do conjunto de dados, à coerência de organização das informações, à validade das

Laboratório de Bioinformática (LABI), Universidade Estadual do Oeste do Paraná (UNIOESTE).
Serviço de Coloproctologia (SC) da Faculdade de Ciências Médicas (FCM) da Universidade Estadual de Campinas (UNICAMP).
Conflito de interesse: nada a declarar.
Fonte de financiamento: Universidade Estadual do Oeste do Paraná (UNIOESTE).
Autorização para publicação na Revista Brasileira de Coloproctologia: Profa. Huei Diana Lee

Recebido em: 24/06/2010

Aprovado em: 14/04/2011

funcionalidades implementadas e ao atendimento aos requisitos básicos definidos. **Conclusão:** de acordo com os especialistas usuários, o desenvolvimento do protótipo completou a primeira etapa do projeto de desenvolvimento de um sistema em nível de 100% de satisfação e, atualmente, está sendo aplicado em cadastros de dados reais. Na próxima etapa será consolidado e implantado o sistema completo, atendendo a requisitos de segurança com suporte a multiusuários.

Palavras-chave: neoplasias colorretais; coleta de dados; análise de dados; sistemas de informação; mineração de dados.

INTRODUÇÃO

O Câncer Colorretal é um dos que mais atinge a população brasileira. Segundo a estimativa do Instituto Nacional do Câncer – INCA, do Ministério da Saúde, no ano de 2010, sem considerar os tumores de pele não melanoma, o câncer de cólon e reto é, em homens, no Brasil, o terceiro mais frequente nas regiões Sul e Sudeste. Na região Centro-Oeste ocupa a quarta posição e, nas regiões Nordeste e Norte, ocupa a quinta posição. Para as mulheres, é o segundo câncer de maior incidência na região Sul e Sudeste, o terceiro nas regiões Centro-Oeste e Nordeste, e o quinto na região Norte ¹.

Diversos trabalhos têm demonstrado avanços no diagnóstico do Câncer Colorretal, como o de SANTOS², no qual se relata um declínio na mortalidade devido a essa enfermidade nos últimos anos nos Estados Unidos. ALTENBURG et al³ trata do sucesso de programas de prevenção, demonstrando que a colonoscopia apresenta-se como o melhor exame preventivo, e também a importância de realização de triagem de pacientes por meio da pesquisa de sangue oculto nas fezes.

No entanto, a abrangência do conhecimento sobre a relação dos hábitos do paciente, seu histórico e sua predisposição genética é limitada dentro de cada escopo, sendo que poucos estudos são realizados correlacionando as diversas variáveis que contribuem para o desenvolvimento deste tipo de câncer.

Para possibilitar a determinação de novos fatores de predisposição ao Câncer Colorretal, assim como avaliar diferentes métodos de tratamento, seria importante a realização de uma análise mais completa das informações de pacientes acometidos dessa neoplasia maligna. Com os avanços tecnológicos, principalmente na área computacional, tornou-se possível o armazenamento de grandes e complexos conjuntos de dados, assim como foram desenvolvidas técnicas computacionais para a análise inteligente desses con-

juntos e, conseqüentemente, a descoberta de novos conhecimentos embutidos nesses dados^{4,5}.

No entanto, para que essas técnicas de análise inteligente de dados sejam corretamente aplicadas, é necessário que procedimentos relacionados, por exemplo, à integridade dos dados, assim como o armazenamento desses dados, sejam realizados de forma adequada.

Assim sendo, este trabalho foi desenvolvido com o objetivo de auxiliar no armazenamento seguro, confiável e acessível de um grande conjunto de dados de pacientes acometidos de Câncer Colorretal, para posteriormente possibilitar a aplicação de processos computacionais de análise inteligente de dados e de descoberta de conhecimento, utilizando, por exemplo, métodos de Inteligência Artificial.

Após a construção de uma base de dados íntegra e estruturada do domínio, a aplicação de técnicas de análise inteligente de dados pode ser feita de maneira mais direta, evitando, assim, a realização de determinadas conversões, padronizações e uniformizações e auxiliando positivamente o processo sob o fator custo vs. tempo. Além disso, pode-se armazenar um conjunto mais abrangente de informações relacionadas a cada paciente, o que permitirá a realização de análises mais completas.

Neste trabalho é proposta a construção de um protótipo de um sistema de cadastro e controle de protocolos de Câncer Colorretal, doravante denominado de Protótipo.

O Protótipo apresentado neste trabalho compõe a primeira parte do projeto de desenvolvimento deste sistema. Na segunda parte será realizada a construção e a implantação do sistema completo. Esse projeto constitui uma parceria do Laboratório de Bioinformática (LABI) da Universidade Estadual do Oeste do Paraná (UNIOESTE), Campus de Foz do Iguaçu, com o Serviço de Coloproctologia (SC) da Faculdade de Ciências Médicas (FCM) da Universidade Estadual de Campinas (UNICAMP).

MATERIAL E MÉTODO

O processo de construção do Protótipo, baseado em diretrizes de Engenharia e Qualidade de Software, foi dividido em cinco etapas:

1. Análise do domínio de Câncer Colorretal;
2. Definição dos requisitos básicos do Protótipo;
3. Projeto do Protótipo;
4. Construção do Protótipo;
5. Avaliação do Protótipo.

Na etapa (1) foi realizada uma análise do domínio de Câncer Colorretal, por meio de estudo da literatura do domínio e de reuniões periódicas com médicos especialistas da área. Nessa fase, foram analisadas todas as informações que podem ter, potencialmente, relação com Câncer Colorretal^{6,7,8,9,10,11,12}.

Na etapa (2) foi realizada a definição dos requisitos básicos do Protótipo, os quais podem ser divididos em requisitos operacionais e requisitos informacionais. Os requisitos operacionais definidos para o Protótipo foram:

- Integridade: as informações gerenciadas pelo Protótipo devem manter-se íntegras ao longo do tempo;
- Facilidade de uso: a interface do Protótipo deve apresentar seus elementos de modo que o uso por médicos e outros profissionais da área de saúde ocorra de maneira eficaz;
- Coerência e concordância: as estruturas interna e externa das informações gerenciadas pelo Protótipo devem ser coerentes e estar em concordância com o contexto do domínio de Câncer Colorretal.

Os requisitos informacionais referem-se a quais informações devem ser gerenciadas pelo Protótipo. Essas informações foram definidas com base na análise das informações potencialmente relacionadas com Câncer Colorretal, definidas na etapa (1).

Na etapa (3) foi realizado o projeto, no qual foram definidas as tecnologias que seriam utilizadas

para a implementação do Protótipo. De modo a permitir acesso remoto, foi definido que o Protótipo funcionaria com o processamento centralizado em um servidor, e a interface de apresentação desse Protótipo seria acessada via Internet por um navegador em computadores remotos. Por meio desses procedimentos, um servidor centralizado seria capaz de fornecer acesso ao Protótipo a vários usuários separados geograficamente por meio da Internet. Adicionalmente, optou-se por desenvolver o Protótipo utilizando apenas ferramentas livres (*open source*), que são:

1. Implementação: linguagem de programação Java^A, a qual é uma linguagem livre, amplamente utilizada para desenvolvimento de aplicações para a *Web*;
2. Apresentação: linguagem de marcação xHTML^B, a qual é recomendada para a apresentação de páginas da *Web* pela “World Wide Web Consortium”^C(W3C), que é uma entidade que define os padrões da *Web*;
3. Modelo arquitetural: *framework* JBoss Seam^{®D}, o qual é um *framework* livre, que apresenta um modelo arquitetural específico para o desenvolvimento de sistemas para a *Web*;
4. Servidor de aplicação: JBoss Application Server^{®E}, o qual é o mais recomendado para utilização em conjunto com o *framework* JBoss Seam[®];
5. Banco de dados: sistema gerenciador de banco de dados PostgreSQL^{®F}, o qual é livre, completo em termos de segurança, integridade e confiabilidade dos dados.

Nessa mesma etapa, as informações gerenciadas pelo sistema foram agrupadas. Essas informações foram divididas em:

- Informações gerais do paciente;
- Histórico do paciente;
- Hábitos sociais;
- Hábitos alimentares;
- Dados complementares.

^A <http://java.sun.com/>

^B <http://www.w3.org/TR/xhtml1/>

^C <http://www.w3.org/>

^D <http://seamframework.org/>

^E <http://www.jboss.org/jbossas>

^F <http://www.postgresql.org/>



Figura 1. Modelo de dados em alto nível.

Essas informações, num total de mais de 100 campos de dados que poderão ser preenchidos, são relacionadas aos principais fatores ligados à etiologia dessa neoplasia e também à evolução do Câncer Colorretal. Foi construído um modelo de dados, apresentado em alto nível na Figura 1, para representar o conjunto de informações com base em diretrizes de modelagem de dados¹³.

Na etapa (4) foi realizada a construção do Protótipo por meio da codificação do Protótipo. Paralelamente, iniciou-se a etapa (5) com a avaliação do Protótipo. Durante a construção foram realizados testes de verificação, os quais consistem em avaliar se as funções implementadas no Protótipo estavam funcionando adequadamente, assim como testes de validação, os quais consistem na validação das funções do Protótipo de acordo com o definido nas etapas (2) e (3), conforme padrões de teste de Engenharia de Software¹⁴ e no modelo arquitetural utilizado¹⁵.

Após a conclusão da construção, o Protótipo foi avaliado por especialistas quanto às características básicas esperadas, definidas nas etapas (2) e (3), ao conjunto de informações efetivamente gerenciadas pelo Protótipo, à concordância e à coerência da interface do Protótipo em relação ao conjunto de informações.

Após essa avaliação interna realizada pelos pesquisadores do LABI, o Protótipo foi disponibilizado aos parceiros do projeto da UNICAMP para avaliação e identificação de mudanças e melhoria necessárias e desejáveis.

RESULTADOS

Primeiramente foi avaliado se as características básicas do Protótipo foram corretamente atendidas. Verificou-se que as informações armazenadas no banco de dados se mantiveram íntegras ao longo dos testes, não apresentando adulterações, perdas ou armazenamento indevido. A interface também foi avaliada por especialistas, os quais analisaram a distribuição das informações pela interface, os métodos de entrada e saída de dados no Protótipo, assim como a facilidade em preencher as informações.

O Protótipo apresentou uma estrutura consistente das informações, agrupando-as corretamente por afinidade. Esses agrupamentos podem ser parcialmente visualizados nas telas do Protótipo, presentes nas Figuras 2,



Figura 2. Tela exibindo informações do paciente.

Figura 3. Tela exibindo o cadastro de dados de histórico familiar.

Figura 4. Tela exibindo o cadastro dos dados iniciais de uma ocorrência.

Nome	Idade Atual	Sexo	Raça ou Cor	Naturalidade	Diabético
Paciente 1	52	Feminino	Parda	Brasileira	Sim
Paciente 2	43	Masculino	Preta	Argentina	Não
Paciente 3	32	Feminino	Branca	Brasileira	Sim

Figura 5. Tela exibindo a lista de pacientes cadastrados.

3, 4 e 5. Também foi avaliado o conjunto de informações gerenciadas pelo Protótipo. Observou-se que todos os requisitos definidos nas etapas (2) e (3) do desenvolvimento do Protótipo foram avaliados como satisfatórios.

DISCUSSÃO

A construção de um sistema para cadastro e controle de protocolos de pacientes que tenham sido acometidos de Câncer Colorretal possibilitará que médicos e profissionais de saúde tenham um meio de fácil acesso e de uso simplificado para gerenciar informações de seus pacientes.

Além de ser útil para o armazenamento e a visualização das informações, a base de dados que será construída ao longo do tempo fornecerá um conjunto abrangente de dados sobre pacientes que tenham sido acometidos de Câncer Colorretal. Esse conjunto de dados, por sua vez, poderá ser utilizado para análise por meio de métodos de análise inteligente de dados e extração de conhecimento, utilizando, por exemplo, métodos de Inteligência Artificial, como Árvores de Decisão e Redes Neurais^{4,5}.

Com um conjunto abrangente de dados em um formato estruturado e padronizado, a união de diversas bases de dados de diferentes unidades do sistema pode ser realizada por um processo simples, sem a necessidade de padronizar, converter ou uniformizar os dados. Desse modo, a compilação de um grande conjunto de dados, obtidos de diferentes clínicas e hospitais, pode ser analisada conjuntamente por meio de métodos de análise inteligente de dados.

Para alcançar esse objetivo, a primeira etapa foi a construção de um protótipo desse sistema. Esse Protótipo é uma versão menos robusta de um sistema funcional, apresentando a interface completa para a manipulação das informações que serão gerenciadas pelo sistema. A construção desse Protótipo tem o propósito de facilitar o processo da definição do modelo de armazenamento dos dados, assim como o processo de adequação da interface quanto à sua utilização pelos seus usuários finais: médicos, pesquisadores e profissionais da área de saúde.

O Protótipo foi desenvolvido em cinco etapas. Na etapa (1) foi realizada uma análise do domínio de Câncer Colorretal, com o propósito de identificar todos os aspectos relacionados a essa neoplasia. Na etapa (2), foram definidos os requisitos básicos do Protótipo com base no propósito do desenvolvimento do Protótipo que foi citado no parágrafo anterior.

Na etapa (3) foi realizado o projeto do Protótipo, utilizando a aplicação de padrões de Engenharia e Qualidade de Software. O projeto consiste na especificação da estrutura do Protótipo, assim como a estrutura de todos os seus componentes. Na etapa (4) foi realizada a construção do Protótipo, por meio da codificação, utilizando as linguagens e ferramentas definidas de acordo com as necessidades e as características identificadas. Paralelamente a essa etapa, iniciou-se a etapa (5), na qual foram realizados testes de verificação, a fim de avaliar se cada nova funcionalidade implementada é executada da maneira esperada, e testes de validação, a fim de validar a implementação em relação aos requisitos operacionais e informacionais definidos na etapa (2), assim como em relação à estrutura definida na etapa (3).

Devido ao nível de complexidade e quantidade de informações gerenciadas pelo sistema, a facilidade de uso da interface foi avaliada como satisfatória por profissionais da área de saúde. Os avaliadores criaram listas de sugestões de mudanças e melhorias desejáveis que tornariam o Protótipo ainda mais completo. Essas listas serão utilizadas na próxima fase do projeto, que será a construção de um sistema completo de gerenciamento de protocolos de Câncer Colorretal.

Após as avaliações, o Protótipo foi disponibilizado a pesquisadores parceiros para ser utilizado para a inserção de dados reais, os quais serão integralmente importados para o sistema completo.

CONCLUSÃO

O Protótipo de um sistema de gerenciamento de protocolos de Câncer Colorretal apresentado neste trabalho possui um grande potencial de aplicação, tanto para fins de análise dos dados armazenados por meio de métodos de análise inteligente de dados, assim como uma ferramenta para controle de informações de seus pacientes que pode ser utilizada em hospitais e clínicas.

A forma estruturada em que os dados são armazenados é de grande importância para a aplicação de técnicas de análise inteligente de dados e extração de conhecimento. Graças à padronização no armazenamento, bases de diferentes locais, que utilizam o sistema, podem ser combinadas com facilidade.

Na próxima fase do projeto será realizado o desenvolvimento do sistema completo, utilizando o Pro-

tótipo como referência. Nessa fase, serão implantados outros requisitos necessários, como criptografia, segurança, acesso multiusuário concorrente, refinamento e aprimoramento da interface, bem como refinamento e consolidação do modelo de dados.

AGRADECIMENTOS

À Universidade Estadual do Oeste do Paraná (UNIOESTE) e à Universidade Estadual de Campinas (UNICAMP).

ABSTRACT: Objective: to develop a prototype of a Colorectal Cancer records management system, aiming to store an embracing structured dataset for later application of intelligent data analysis methods. **Material and Method:** the prototype was developed in five stages: Colorectal Cancer domain analysis, definition of basic requirements (operational and informational), system design, system construction using open source technologies and system evaluation supported by domain experts. **Results and Discussion:** during development, the prototype dataset structure, coherence of information structure, validity of functions implemented and the attendance to the basic requirements define, were evaluated by domain experts. **Conclusion:** according to the experts evaluation, the prototype development completed the first stage of the project of developing the system with a 100% satisfaction rate and is now being used into recording real data. In the next stage it will be finished and deployed the complete system, attending to security requirements and multiuser support.

Keywords: colorectal neoplasms; data collection; data analysis; information systems; data mining.

REFERÊNCIAS

1. Instituto Nacional do Câncer. Estimativa/2010 – Incidência de câncer no Brasil. [acesso em 2010 Mar 3]. Disponível em: <http://www1.inca.gov.br/estimativa/2010/estimativa20091201.pdf>
2. Santos Jr JCM. Câncer ano-reto-cólico: aspectos atuais IV – câncer de cólon – fatores clínicos, epidemiológicos e preventivos. Rev Bras Coloproct [periódico na Internet]. 2008 Set [citado 2010 Jun 11]; 28(3):378-385. Disponível em: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0101-98802008000300019&lng=pt. doi: 10.1590/S0101-98802008000300019
3. Altenburg FL, Biondo-Simões MLP, Bahten LCV. A pesquisa de sangue oculto nas fezes associada a um questionário de sinais e sintomas na prevenção do câncer colo retal. Rev bras coloproctol. [periódico na Internet]. 2009 Mar [citado 2010 Jun 11]; 29(1):57-62. Disponível em: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0101-98802009000100008&lng=pt. doi: 10.1590/S0101-98802009000100008
4. Han J, Kamber M. Data mining: concepts and techniques. San Francisco: Morgan Kaufman; 2006.
5. Witten IH, Frank E. Data mining: practical machine learning tools and techniques. San Francisco: Morgan Kaufman; 2005.
6. Chlebowski RT, Wactawski-Wende J, Ritenbaugh C, Hubbell FA, Ascensao J, Rodabough RJ, et al. Estrogen plus progestin and colorectal cancer in postmenopausal women. N Engl J Med 2004, 350(10):991-1004.
7. Cordeiro F, Yamaguchi NH, Habr-Gama A, Cutait R, Reinan, RJ, Abramoff R, et al. Diagnóstico, estadiamento e tratamento cirúrgico e multidisciplinar do câncer colorretal. Associação Médica Brasileira e Conselho Federal de Medicina; 2005.
8. Larsson SC, Orsini N, Wolk A. Diabetes mellitus and risk of colorectal cancer: a meta-analysis. J Natl Cancer Inst 2005, 97(22):1679-87.
9. Neves FJD. Mortalidade por cancer de cólon e reto e perfil de consumo alimentar em capitais brasileiras [dissertação]. Rio de Janeiro (RJ), Brasil : Escola Nacional de Saúde Pública; 1995.
10. Quilici FA. Colonoscopia. São Paulo (SP), Brasil: Editora Lemos; 2000.
11. Rocha JJR. Coloproctologia: princípios e prática. São Paulo: Atheneu; 2005.
12. Santos Jr. Júlio CM. Câncer ano-reto-cólico: aspectos atuais II – câncer colorretal – fatores de riscos e prevenção. Rev bras. colo-proctol. [periódico na Internet]. 2007 Dec [citado 2010 Jun 11] ; 27(4):459-473. Disponível em: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0101-98802007000400016&lng=en. doi: 10.1590/S0101-98802007000400016.
13. Elmasri R, Navathe SB. Sistemas de Banco de Dados. São Paulo: Pearson Addison-Wesley; 2005.
14. Pressman RS. Engenharia de Software. São Paulo: McGraw-Hill; 2006.
15. Allen D. Seam in action. Greenwich: Manning Publications Co.; 2008.

Endereço para correspondência:

HUEI DIANA LEE

Laboratório de Bioinformática (LABI)

Universidade Estadual do Oeste do Paraná (UNIOESTE)

Parque Tecnológico Itaipu (PTI)

Av. Tancredo Neves, 6731

Caixa Postal: 39 - CEP: 85856-970 – Foz do Iguaçu (PR), Brasil

E-mail: hueidianalee@gmail.com