Original Article

Prototype system to manage data on coloproctology surgery

HUEI DIANA LEE[1], WILSON JUNG[2], ADRIELI CRISTINA DA SILVA[3], LUIZ HENRIQUE DUTRA DA COSTA[2], BIANCA ESPINDOLA[4], CLÁUDIO SADDY RODRIGUES COY[5], JOÃO JOSÉ FAGUNDES[5], FENG CHUNG WU[6]

[1]Professor and Doctor at the Computer Science course and in the Post Graduation Program of Dynamic Systems and Energy engineering (PGESDE) and general coordinator at the bioinformatics laboratory (LABI) of the Universidade Estadual do Oeste do Paraná (UNIOESTE) – Foz do Iguaçu (PR), Brazil; Visiting professor in the Post Graduation Program of Surgical Sciences at the School of Medical Sciences of Universidade Estadual de Campinas (UNICAMP) – Campinas (SP), Brazil. [2]Student at PGESDE; Member of the LABI at UNIOESTE – Foz do Iguaçu (PR), Brazil. [3]Member of the LABI at UNIOESTE – Foz do Iguaçu (PR), Brazil. [4]Student in the Post Graduation Program of UNICAMP – Campinas (SP), Brazil. [5]Professor and Doctor of the Department of Surgery, the Service of Coloproctology and the Post Graduation Program of Surgical Sciences at the School of Medical Sciences of UNICAMP – Campinas (SP), Brazil. [6]Professor participating in the Post Graduation Program of Surgical Sciences of UNICAMP – Campinas (SP), Brazil; Professor and Doctor of UNIOESTE and coordinator of the Medical Department of LABI at UNIOESTE – Foz do Iguaçu (PR), Brazil.

LEE1 HD, JUNG W, SILVA AC, COSTA LHD, ESPINDOLA B, COY CSR, FAGUNDES JJ, WU FC. Prototype system to manage data on coloproctology surgery. **J Coloproctol**, 2011;31(4): 351-361.

ABSTRACT: Objective: To develop a prototype system to manage data on coloproctology surgery, aiming at Data Quality (DQ) and the adoption of a DQ monitoring process, which is nonexistent in most biomedical systems. Methods: The construction of the prototype was separated into five steps: analysis of an existing system (legacy), the analysis of requirements and specifications for the new prototype, the development of the model, definition of technologies and the development of a prototype. Results: The analysis of the legacy system revealed several limitations and inconsistencies, which can result in problems concerning the DQ. Therefore, actions to prevent these problems are already being executed at the step of developing the prototype, such as the creation of interactive and more elaborate interfaces, the use of validation mechanisms on data fields and the proposal of a process to monitor inconsistencies and incompleteness in patients' data. Conclusion: The adoption of DQ mechanisms on system development results in building a reliable and consistent database, to assist tasks such as management, scientific research and future intelligent data analysis methods. Future work includes subjective evaluations of DQ indicating the adequacy of the prototype for the users' needs.

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

INTRODUCTION

Technology has been increasingly contributing by assisting the different activities in various fields, pushing the development of human knowledge. Among the different technological contributions, information systems have an important role concerning data management, since they feed databases with information, which is the main raw material for scientific research, besides being essential for decision making.

According to this point of view, in the health field it is necessary to manage information on different aspects of a patients for the performance of a surgical procedure; more specifically, in coloproctology, this specialty comprises various diseases that can be treated with surgery, such as cancer, adenomatous polypo-

Study carried out at the Bioinformatics Laboratory (LABI) of Universidade Estadual do Oeste do Paraná (UNIOESTE) – Foz do Iguaçu (PR), Brazil. Coloproctology Service of the School of Medical Sciences at Universidade Estadual de Campinas (UNICAMP) – Campinas (SP), Brazil.
Financing source: Aid by the scholarship financing programs: National Council for Scientific and Technological Development (CNPq) and Itaupu Technological Park Foundation (FPTI-BR).
Conflict of interest: nothing to declare.

Submitted on: 08/08/2011
Approved on: 09/09/2011

sis, Chron's disease, among others[1-6]. Thus, an information system may significantly contribute with the management of data concerning surgery, that is, pre-, inter and postoperative periods, enabling the construction of a solid and organized database.

In order to assist in a more complete analysis of the great volume of data created by the information systems, computer processes such as Data Mining (DM) can be used, since they aim to discover knowledge or interesting standards, which are not difficult to obtain when conducted by the non-automatic data analysis[7]. However, so that these data can be used, both for decision making and for the application of DM processes, it is important that the data that are present in the information systems be a faithful representative of the facts[8]. Low-quality data can harm any process that uses them, and the adoption of corrective measurements leads to high costs for the institutions[9].

Problems with data quality are not only characterized by data inconsistency, but also by how useful the information is[(1)]. Therefore, Data Quality (DQ) can be defined as the information that has less inconsistencies and that is adequate to the purposes of the users[10].

The variety of possible DQ problems is huge. However, the problems have characteristics that allow their classification as to perspective, user and data, and to the depending or not on the context[10], according to the following requirements:

- Data perspective/regardless of the context: problems that may be present in any database, such as spelling mistakes, duplicated data and incorrect values;
- Data perspective/depending on the context: problems that violate the specifications of the business, that is, rights and restrictions of a specific domain that define the norms of the system;
- User perspective/depending on the context: problems concerning data processing, such as inaccessible, insecure data and those that are difficult to recover;
- User perspective/regardless of the context: problems that indicate that data are not adequate to the purposes of the user, such as incomplete data, those that are irrelevant to the activities and difficult to understand.

In order to assist the evaluation of DQ, there are different dimensions in literature to represent the characteristics of data[11-15]. Among the most cited dimensions, the *accuracy* characterizes the faithful representation of real world facts by data obtained by the information systems. *Completeness*, on the other hand, characterizes the complete representation of the states of an entity in the real world, which are necessary for the user's needs. There is also the *consistency* dimension, which shows the presence of redundant data or the consistency between two related elements, such as address and postal code. Concerning data timing, the *punctuality* dimension shows the data that are up-to-date and available to meet the users' needs at a proper time. As to the level of credibility and truth of the facts, it is characterized by the *credibility* and the *relevance* dimension, which characterize the level of applicability and assistance that the data can offer to meet the users' needs.

Thus, the evaluation of DQ uses the existing dimensions associated with metrics, which represents the problems with DQ to quantify the level of quality of the specific characteristics of data.

Nowadays, the Coloproctology service at the School of Medical Sciences (FCM) of *Universidade Estadual de Campinas* (UNICAMP) uses a system developed with *Microsoft Acess*[(2)] to register the surgeries performed in this service. However, this system, which was later called the legacy system, presents limitations and inconsistencies, thus not meeting all the users' needs, which impacts directly on the quality of the stored data.

The objective of this study is to present the development of a prototype to manage data on coloproctogy surgery that offers a proper structure to the processes that use it; it also aims to create a vast data base with DQ. Thus, these data can be used for the administrative control and also for academic research; the latter is performed by the use of DM processes by health professionals, together with informatics professionals. It is interesting to point out that the proposal of this system includes a process to monitor DQ, which is still uncommon in biomedical systems, which aim to minimize the presence of wrong information in databases, thus performing the validation of fields in the

---

[1]*In this study, the terms data and information are indistinctively used.*
[2] *http://office.microsoft.com/*

interface and the emission of alerts and reports on inconsistencies and lack of patients' data[16].

This study is part of the project of Intelligent Data Analysis, which is developed in a partnership between the Bioinformatics Laboratory (LABI) and *Universidade Estadual do Oeste do Paraná* (UNIOESTE), Foz do Iguaçu, the Coloproctology Service of FCM, in UNICAMP, The Laboratory of Computer Intelligence (LABIC), of *Universidade de São Paulo* (USP), São Carlos, and the Interdisciplinary Group in Data Mining and Applications (GIMDA), of *Universidade Federal do ABC* (UFABC), São Paulo.

METHODS

In order to develop the prototype, consolidated methods of software engineering, such as prototype and the incremental development[17;18].

The prototype model, as represented in Figure 1, enables the development of a system with the construction of prototypes in continuous interactions with users, thus obtaining the complete system[17].

The incremental model also adopts the interactive philosophy of prototype. A set of functions is added to each version of the system, which are called increments, starting with the essential elements until the system is complete, that is, after the performance of all increments[18].

The project of this system was organized in five main steps:

1. Analysis of the legacy system;
2. The gathering of requirements for the prototype;
3. Development of models and prototype screens;
4. Definition of Technologies to be used;
5. Development of the prototype.

In step (1), the legacy system was analyzed, thus mapping the structure of the system to identify the set of previously collected information, the existing limitations and the possible flaws that can have a negative interference in DQ. Figure 2 represents a high level vision of the legacy system. In this figure, it is possible to observe that the item *patient* is related with the other items of the legacy system. However, one of the limitations in the system of is that it accepts the registration of only one surgery and a set of preoperative examinations per patient.

Based on the analysis of step (1), step (2) was characterized by the definition, concluded by experts in the medical and informatics fields, of the basic necessary requirements for the new system:

- Control of access to the system: only authorized people can access the patients' data;
- Organization of the information: the system should have a friendly interface and provide information according to the classic elements of medical propaedeutics[19];
- Mobility and easy access: the system should provide mobility to the health professionals, so that the access does not depend on the geographic position or on the computer that is being used;
- Data quality: the system should have mechanisms to assist the monitoring of the quality of data inserted into it.

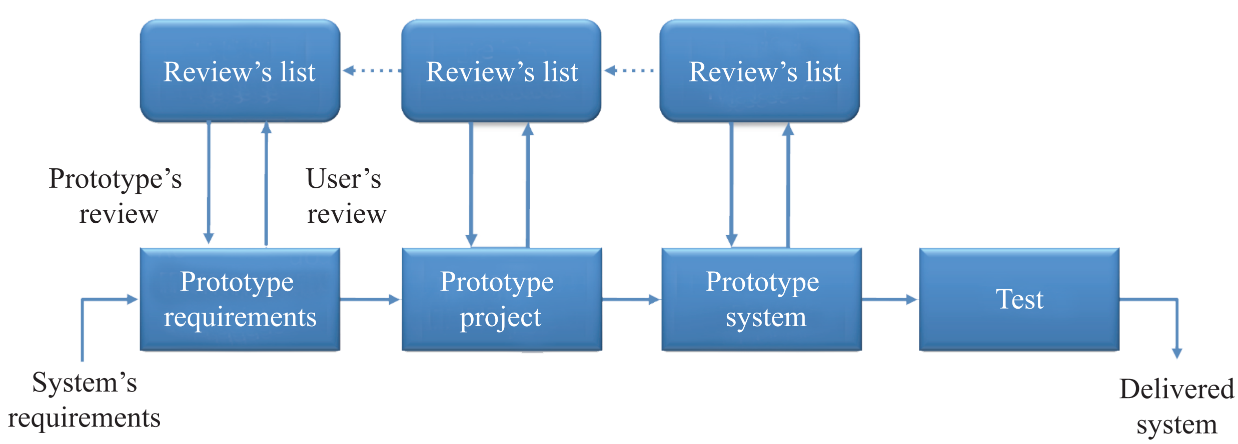

Figure 1. *Prototype Model*[modified from 17].

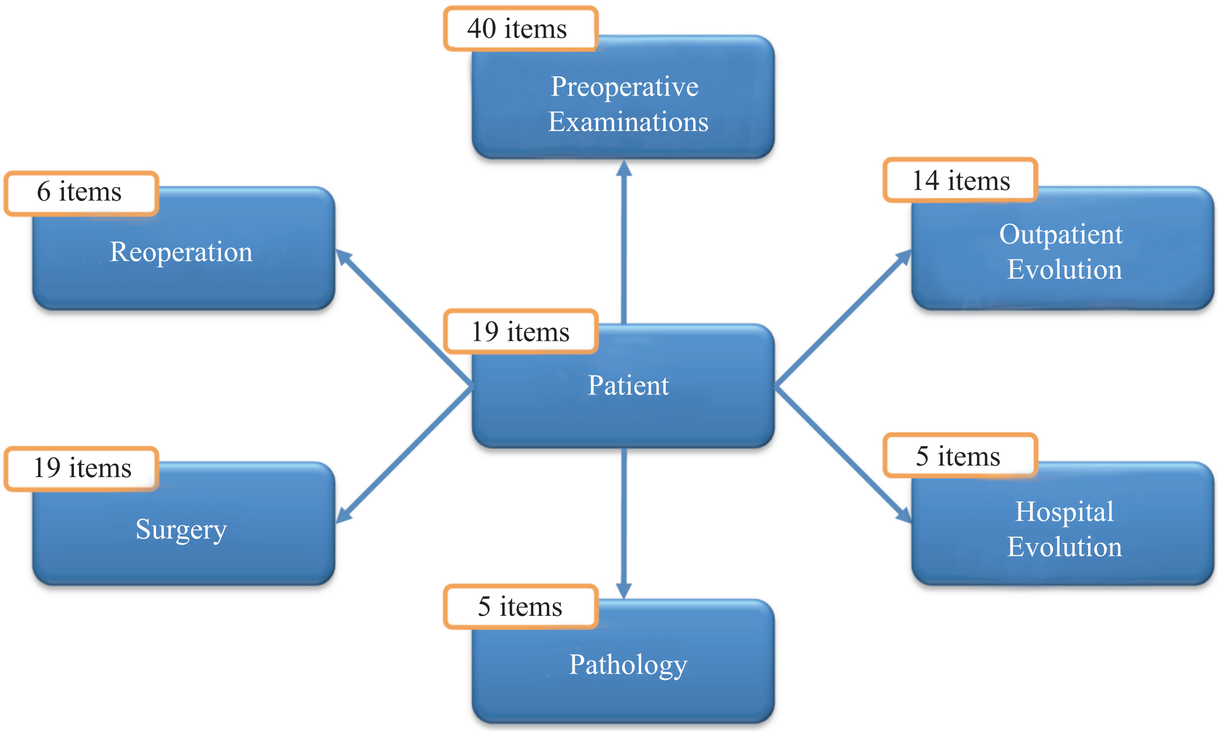

Figure 2. *Model of high level data in the legacy system.*

Besides the presented requirements, additional items were defined that were not present in the legacy system. The proposed prototype is composed of approximately 446 items (Figure 3) in comparison with the 108 items of the legacy system (Figure 2).

On step (3), the models were created to represent the entities of the new system and its relations, as demonstrated in Figure 3, providing a high level view of the prototype.

In this step, the screen prototype was developed, which enabled the previous definition of the structure of the system, its functions and the distribution of information. This step consisted of continuous interactions among experts in the medical and informatics field, so that all the requirements of the prototypes could be met.

Thus, the information concerning the patients that is managed by the prototype is organized into three sections:

- Identifying the patient;
- Occurrence;
- Family history.

The *identification* section of the patient is comprised of data that allow the identification of a patient, as well as other information concerning his or her origin (nationality), religion and data on blood type and RH factor. Afterwards, the *occurrence* section is comprised of data related to one surgery, so it is possible to register as many occurrences as necessary for each patient. The data in this section are organized in different subsections, such as:

- Personal history: data on diseases, such as *diabetes mellitus*, hypertension and neoplasm, social and eating habits of the patient;
- Diagnosis: data on coloproctological disease;
- Preoperative examinations: data on laboratory, imaging and clinical examinations in the preoperative period;
- Surgery: data on performed surgical procedures;
- Postoperative: data on hospital and outpatient follow-up related to the respective occurrence.

As mentioned, it is possible to register many occurrences for each patient, and each of them can be identified by date and time by the responsible doctor.

Finally, the *Family history* section consists of data that identify the family characteristics of the person, such as the number of brothers, children, if there is has a twin sibling or if the person is adopted.

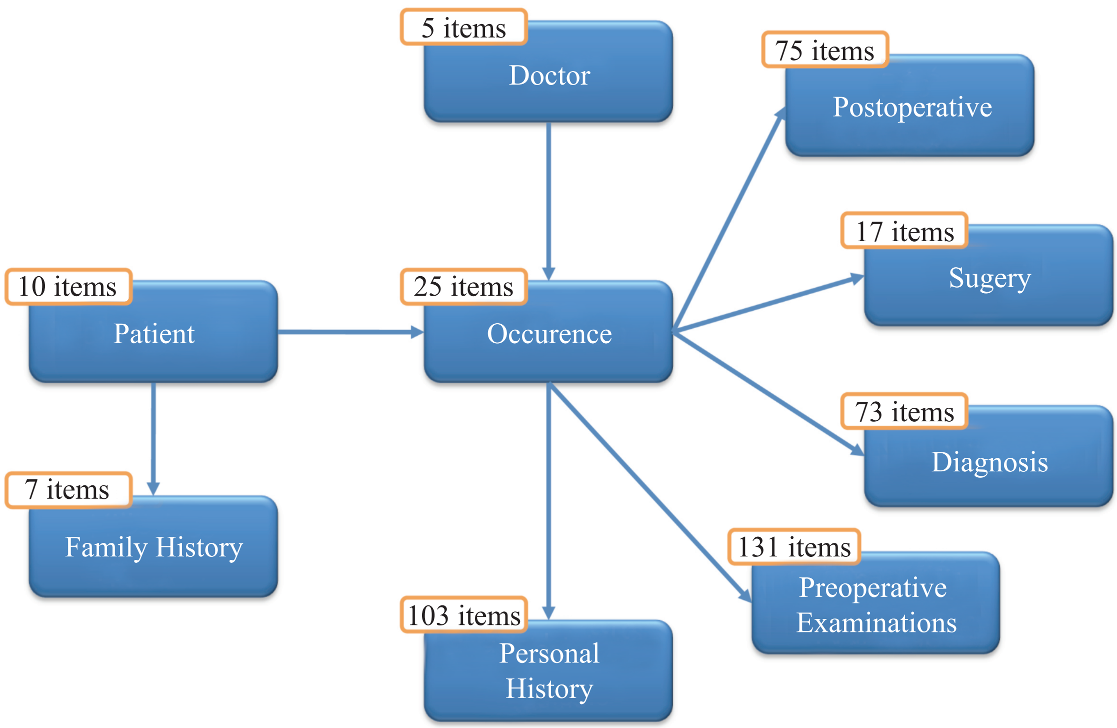

Figure 3. *Model of high level data of the prototype.*

In step (4), the technologies used to develop the project and to make th prototype work were defined. These opensource technologies were:

1. Implementation: The Ruby programming language[3], which is a programming language focused on simplicity and productivity, and the framework for web development *Ruby on Rails*[4], provides the fast development of applications;
2. Application server: Apache application server[5], widely used for web applications, together with the plugin Passenger[6], which allows the installation of Ruby applications in Apache servers.
3. Database: system to manage the database (SGBD) PostgreSQL[7], which guarantees the integrity of the stored data.

In step (5), the prototype was developed with the help of Technologies presented in step (4) according to the models established in step (3). The development of the prototype was followed-up by experts in the medical field.

[3] *http://www.ruby-lang.org/*
[4] *http://rubyonrails.org/*
[5] *http://www.apache.org/*
[6] *http://www.modrails.com/*
[7] *http://www.postgresql.org/*

RESULTS

This paper led to study the viability of adopting DQ measurements during the development of a biomedical system, which is not so common in existing systems and significantly contributes with the creation and the maintenance of a quality database[20-22].

From a previous bibliographic analysis, it was possible to define characteristics that contribute with the DQ, and also that helps to put them into practice during the process to develop a prototype to manage data on coloproctological surgery.

The system modelling was conducted by experts in the medical and informatics fields, resulting in the structure of the system represented in Figure 3; this enables the organized and loyal representation of the processes and information concerning the patients and their occurrences. In Figures 4, 5, 6 and 7, examples of prototype screens developed according to the requirements and models defined in steps (2) and (3).

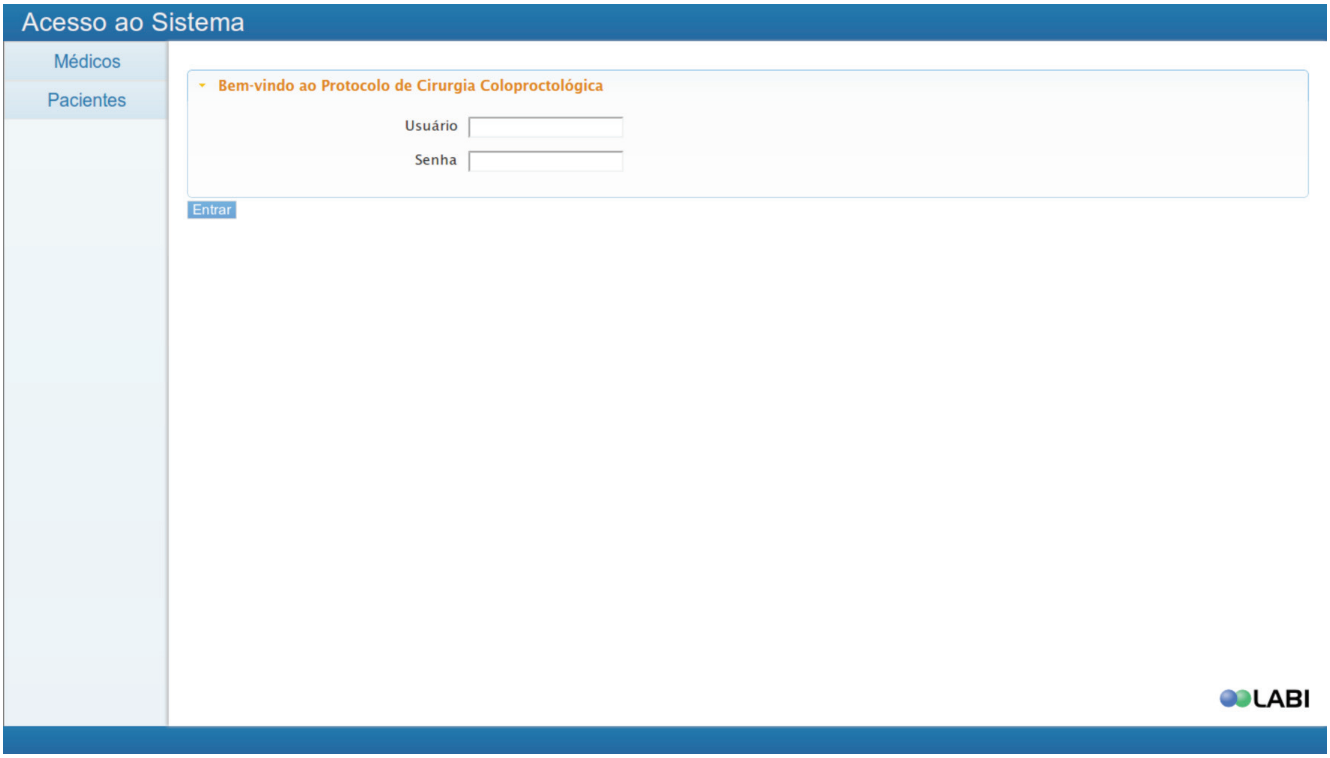

Figure 4. *Screen to access the system.*

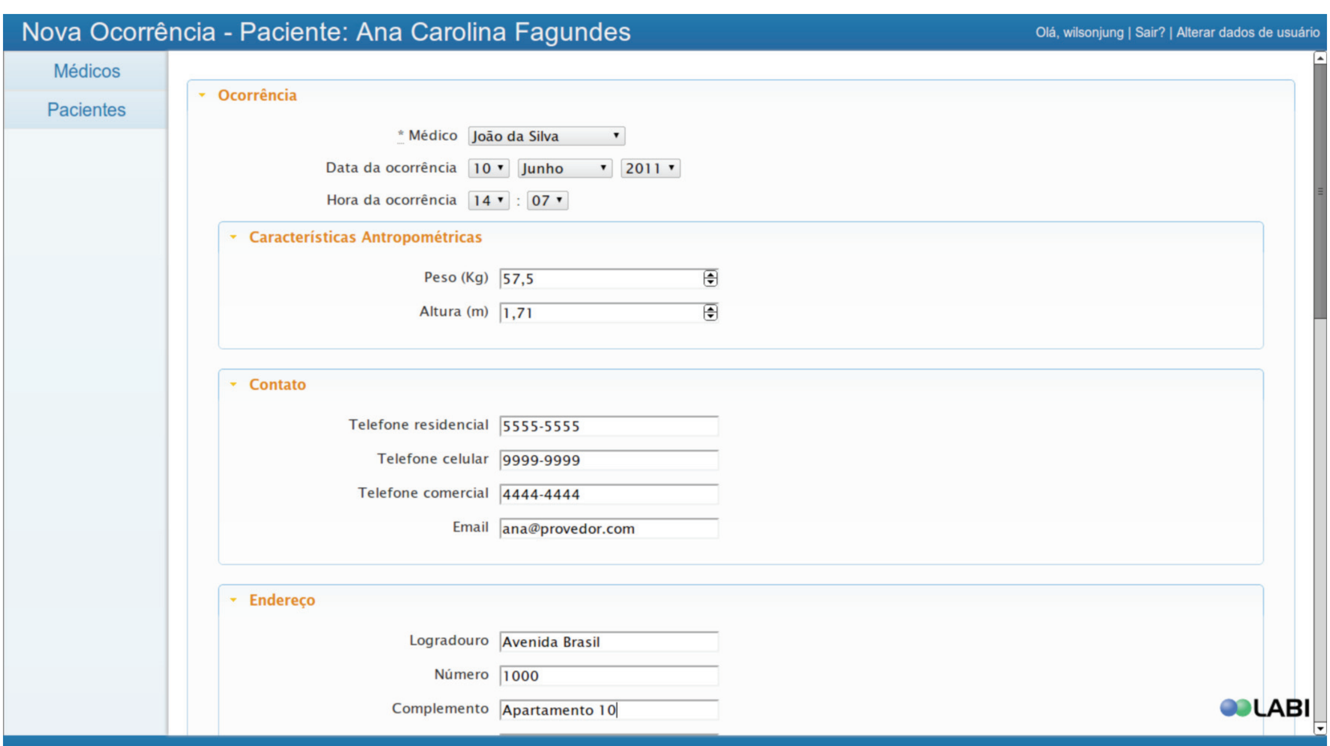

Figure 5. *Screen for the initial registration of data.*

Figure 6. *Screen to consult patients' data.*

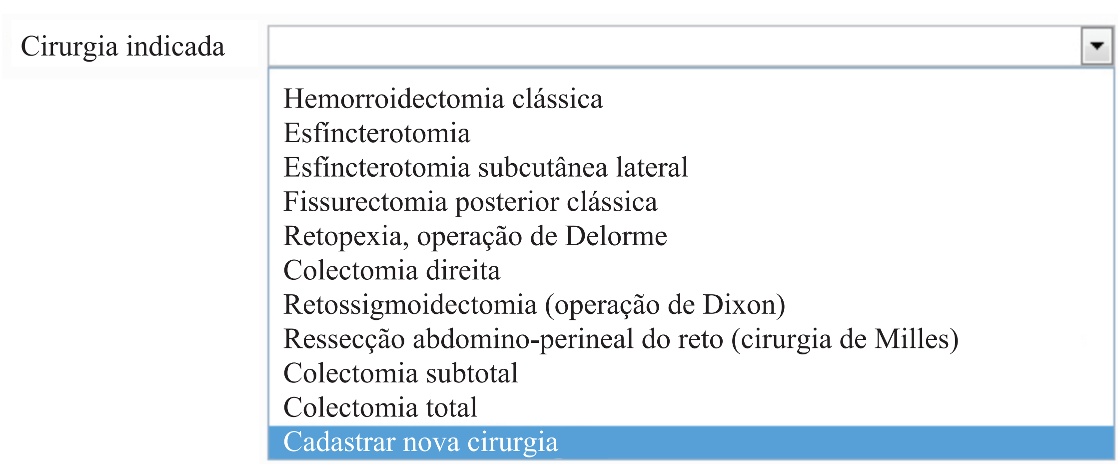

Figure 7. *List of possible values for the field "cirurgia indicada" (Indicated surgery).*

Figure 4 shows the result of the implementation of mechanisms to control the access, so that the system can only be accessed by authorized people, thus assuring the privacy of the patients' data. The users can only be registered by people with administrative privileges.

After registering in the system, the user can access the list of patients and doctors, according to the privileges given to his account by the administrator, by the fields showing on the left side of the screen. For example, in Figure 5, the screen to register the information concerning the occurrences of one patient is presented. It is important to say that these data are fictitious, used only to show the functions of the system. The style adopted for this screen is also used in the other sectios of the system.

After finishing the registration, it can be consulted in the screen represented in Figure 6, which is similar to the registration screens; however, its only purpose is to show the information.

In the fields in which predefined values can be inserted, as represented in Figure 7, there are mechanisms that enable the user to select the value or to register a new value, in case the wanted item is not listed.

DISCUSSION

The analysis of the legacy system conducted in step (1) showed many limitations and inconsistencies that can lead to problems in DQ. The mentioned system was developed as an emergency solution to register data concerning the performed coloproctological surgeries. Thus, its modeling does not properly reflects the needs of the clinical and surgical processes represented by the data in the patients' medical records.

The data model of the legacy system, presented in Figure 2, enables to verify the relationship between the entities in this system. In this model, a patient can be only submitted to one surgery, and only one set of preoperative examinations can be registered. The item *Reoperation* registers patients' recurrences, but it does not allow the registration of results of preoperative examinations related to the recurrence. Much important information that characterize the clinical picture of the patients do not have specific fields for registration, thus, it is necessary that professionals register these data in the observation fields. The data in these observation fields are in natural language, which is easy to understand; however, it is more difficult to apply the processes of analysis, like DM, besides being more prone to typos. Also, the legacy system does not have a friendly interface or security mechanisms that prevent non authorized people from accessing data; its structure is not adequate to store a great volume of data, and does not provide the possibility for the remote access to the system.

Figure 3 represents the relation of the items in the proposed prototype, and it is possible to observe that, unlike the legacy system, the information on a surgery is gathered in the item *Occurrence,* which is also related to a doctor. Besides, the additional items enable a complete representation of the characteristics and conditions associated with the occurrences of patients, and also contribute with future applications of processes, such as DM.

It is also important to point out that DQ positively influences the application of the DM process[23], resulting in more reliable models and reducing the time of the data preprocessing step, which is the most difficult in the process[7].

Because of the presented issues and the importance of DQ for the processes that use them, steps (2) and (3) resulted in the project of a prototype to manage data on coloproctological surgery. As a guide to properly build this prototype, the Electronic Health Records Certificate (*Manual de Certificação para Sistemas de Registro Eletrônico em Saúde)*, approved by the resolution 1821/07 of the Federal Council of Medicine[24], presents several guidelines for the development of the project.

Initially, it was defined that the prototype would be a web system, in order to make update and maintenance of the system easier, since they would be centralized in the server. Another advantage of choosing a web system is the mobility for the health professional, since he or she can access the system regardless of their geographical location or computer – only a web navigator would be necessary. The development of a prototype was also accompanied by medical experts, who suggested some improvements in the implemented functions, characterizing the interaction process that occurs with the prototype and the incremental development used as software process models.

The programming language used to develop the prototype, as described in step (4), is the Ruby language, accompanied by the framework for the web development of Ruby on Rails. A programming language that allows the logic of the program to be developed in an understandable way, that would be later translated into computer instructions. A framework, on the other hand, can be defined as a set of library codes and mechanisms to enable the development of activities they are destined to[25].

The Ruby on Rails framework particularly allows the automation of different tasks during creation, installation and maintenance projects of a web system, enabling the fast development of applications[26]. It was built to solve 80% of the problems that are present in the web development, considering that the other 20% depend on the context of the application[25]. Another characteristic of this framework is the presence of mechanisms that assist the validation of graphic interface fields, establishing standards for the data to be inserted by the users[25]. Besides, this framework provides mechanisms to develop test routines to assess the proper implementation of the functions of the system.

Another important aspect of the project is the concern about DA in the step of prototype development [step (5)]. Thus, the functions offered by the Ruby on Rails framework also contribute with this purpose. In literature, a great part of the solutions for DQ problems are addressed to corrective methods. These methods are important because they imply high costs to the institutions. Thus, it is better to reach balance between preventive and corrective measurements[9].

As a preventive measurement of DQ, the project of the system also aims the implementation of processes to support the constant monitoring of the quality level of the system's data, thus enabling the adoption of corrective measurements, if necessary, before the problems are spread and result in serious conflicts[20]. The monitoring process also aims to help the health professional to identify data inconsistencies, indicating records with problems. Thus, the professional can count on more reliability for decision making based on data.

Another impact factor for DQ is the presence of pleasant and organized graphic interfaces in the system, so the users can easily access the system. Poor graphic interfaces may result in frustrating experiences for the users, who may leave blank or incorrect fields[27]. Because of this, the information will be hierarchically organized in the new system, providing an environment that is more structured to insert data. Also, mechanisms that allow to hide non used data were adopted, resulting in a cleaner and more pleasant interface, and also making the location of specific information easier.

The prototype presented in this study was designed to get a larger volume of structured information to represent the patients' characteristics. When the set of values for the fields is known, mechanisms that present the relation of possible values for that field were used, thus allowing the user to choose the proper value, as presented in Figure 7. Such mechanisms help to reduce inconsistencies and typos, contributing with DQ. If the desired value is not in the presented list, it is possible to register a new value, that will be later included in the list.

It is important to say that the proposed prototype presents some special functions, aiming to facilitate and stimulate the use of the system. For instance, we could cite the assistance of the semiautomatic filling, with the confirmation of personal history by a health professional after the first occurrence. These data are recovered from the base of previous occurrences, so the user does not need to reinsert the information that is already in the system – it is only necessary to update the information in case there is any change.

To complement the DQ solution, another important function is the registration of operations in the system by a determined user, thus helping to identify the cause of problems, if they occur.

CONCLUSION

Considering the impact that DQ problems can cause, the adoption of measurements that help to prevent these problems is very relevant, since it increases the reliability of the system and has a positive impact on the processes that use it. Despite the careful planning to develop the new system and its characteristics, the users also need to be aware and carefully insert the data.

It is important to point out that the process of DQ monitoring enables the adoption of corrective measurements that are necessary in a proper time, before any problems can be spread, thus avoiding the origin of serious conflicts. This differential does not exist in most biomedical systems, helping to build a structured data base that can be used for simple consultations and for scientific research; for example, methods for an intelligent analysis in DM, as well as statistical analysis, can be used.

In the next step, the prototype will be improved, and the data of the legacy system will migrate and be adjusted to the new system, thus minimizing the impact and becoming more pleasant to the users. Afterwards, a subjective DQ assessment of the built prototype will be performed, leading to the analysis of adapting to the needs of the users, as well as making adjustments, if necessary.

ACKNOWLEDGEMENTS

To the National Council for Scientific and Technological Development (CNPq) and Itaipu Technological Park Foundation (FPTI-BR), for the support by the scolarship financing.

RESUMO: Objetivo: Desenvolvimento de um protótipo para gerenciamento de dados de cirurgia coloproctológica, visando à Qualidade de Dados (QD), e a adoção de um processo de monitoramento da QD, inexistente na maioria dos sistemas biomédicos. Métodos: A construção do protótipo foi dividida em cinco etapas: análise de um sistema existente (legado), levantamento dos requisitos para o novo protótipo, elaboração de modelos, definição das tecnologias e desenvolvimento do protótipo. Resultados: A análise do sistema legado revelou diversas limitações e inconsistências que podem resultar em problemas de QD. Sendo assim, medidas para prevenir esses problemas estão sendo adotadas, já na etapa do desenvolvimento do protótipo, como a criação de interfaces mais elaboradas e interativas, a utilização de mecanismos de validação dos campos de dados e a proposta de um processo para monitoramento das inconsistências e incompletudes dos dados dos pacientes. Conclusão: A adoção de medidas de QD no desenvolvimento de sistemas resulta na construção de uma base de dados confiável e consistente, contribuindo com as tarefas de gerenciamento, pesquisas científicas e futuras aplicações de métodos de análise inteligente de dados. Trabalhos futuros incluem avaliações subjetivas de QD que indiquem o nível de adequação do protótipo às necessidades dos usuários.

Palavras-chave: cirurgia colorretal; coleta de dados; análise de dados; sistemas de informação; mineração de dados.

REFERENCES

1. Valarini R, Campos FGCM. Resultado do registro nacional brasileiro em vídeo-cirurgia colorretal em 2007. J Coloproctol 2008;28(2):145-55.
2. Martins JF, Rocha JG, Miranda EF, Sartor MC, Steckert JS, Steckert AS, et al. Análise da prevalência de entidades coloproctológicas nos pacientes idosos do serviço de coloproctologia de um Hospital Universitário. J Coloproctol 2009;29(2):145-57.
3. Torres Neto JR, Souza JRMCA, Santiago RR, Prudente ACL. Cirurgias colorretais no Hospital Universitário da Universidade Federal de Sergipe: três anos de criação do serviço de coloproctologia (série histórica). J Coloproctol 2008;28(1):77-83.
4. Prudente ACL, Torres Neto JR, Santiago RR, Mariano DR, Vieira Filho MC. Cirurgias proctológicas em 3 Anos de serviço de coloproctologia: série histórica. J Coloproctol 2009;29(1):71-6.
5. Queiroz FL, Côrtes MGW, Rocha NP, Alves AC, Freitas AHA, Lacerda Filho A, et al. Resultados do registro de cirurgias colorretais videolaparoscópicas realizadas no estado de Minas Gerais - Brasil de 1996 a 2009. J Coloproctol 2010;30(1):61-7.
6. Oliveira RG, Faria FF, Lima JRACB, Rodrigues FG, Andrade MMA, Gomes DMBM, et al. Cirurgias êntero-colorretais – abordagem cirúrgica de 129 pacientes do SUS no programa de pós-graduação Sensu Lato em coloproctologia. J Coloproctol 2010;30(3):333-43.
7. Han J, Kamber M. Data mining: concepts and techniques. 2nd ed. San Francisco: Elsevier; 2006.
8. Orr K. Data quality and systems theory. Communications of the ACM 1998;41(2):66-71.
9. Guerra-García C, Caballero I, Piattini M. A systematic literature review of how to introduce data quality requirements into a software product development. In: 5th International Conference on Evaluation of Novel Approaches to Software Engineering. Athens, Greece: 2010.
10. Ge M, Helfert M. A Review of information quality research. In: Proceedings of the 12th International Conference on Information Quality. Cambridge: MIT; 2007. p. 76-91.
11. Wand Y, Wang RY. Anchoring data quality dimensions in ontological foundations. Commun ACM 1996;39(11):86-95.
12. Lee YW, Strong DM, Kahn BK, Wang RY. AIMQ: a methodology for information quality assessment. Information & Management 2002;40(2):133-46.
13. Pipino LL, Lee YW, Wang RY. Data quality assessment. Communications of the ACM 2002;45(4):211-8.
14. Lee YW, Pipino LL, Funk JD, Wang RY. Journey to data quality. Cambridge: MIT Press; 2006.
15. Batini C, Scannapieco M. Data quality: concepts, methodologies and techniques. Berlin: Springer-Verlag; 2006.
16. Jung W, Lee HD, Silva AC, Costa LHD, Espindola B, Coy CSR, et al. Adoção de medidas de qualidade de dados para o desenvolvimento de sistemas biomédicos. In: X Conferência Brasileira de Dinâmica, Controle e Aplicações - Dincon 2011. Águas de Lindoia, Brasil; 2011.
17. Pfleeger SL. Engenharia de software: teoria e prática. 2nd ed. São Paulo: Prentice Hall; 2004.
18. Pressman RS. Engenharia de software. 5th ed. Rio de Janeiro: McGrawHill; 2002.
19. Porto CC. Exame clínico. Rio de Janeiro: Guanabara Koogan; 1982.
20. Salvador VFM, Britto M, Moura Junior LA, Almeida Junior JR. Qualidade de dados para gestão de conhecimento na área de saúde. In: Anais do X Congresso Brasileiro de Informática em Saúde. Florianópolis, Brasil; 2006.
21. Fadel FDQ, Malucelli A, Bastos LC. Prontuário eletrônico para pacientes de hanseníase via web. In: X Congresso Brasileiro de Informática em Saúde. Florianópolis, Brasil; 2006. p. 1462-7.
22. Herzberg S, Rahbar K, Stegger L, Schäfers M, Dugas M. Concept and implementation of a computer-based reminder system to increase completeness in clinical documentation. Int J Med Inform 2011;80(5):351-8.
23. Blake R, Mangiameli P. The effects and interactions of data quality and problem complexity on classification. JDIQ 2011;2(2):8.

24. Conselho Federal de Medicina. Resolução CFM Nº 1.821/07 [Internet]. 2007. Available from: http://www.portalmedico.org.br/resolucoes/CFM/2007/1821_2007.htm
25. Carneiro Junior C, Al Barazi R. Beginning rails 3. Berkely: Apress; 2010.
26. Ruby S, Thomas D, Hansson DH. Agile web development with rails. 4th ed. Raleigh: Pragmatic Bookshelf; 2011.
27. Maydanchik A. Data quality assessment. Bradley Beach: Technics Publications, LLC; 2007.

Correspondence to:
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), Brazil
E-mail: labi.unioeste@gmail.com