Development of an oral health database for the management of clinical records

Desenvolvimento de um banco de dados de saúde bucal para administração de registros clínicos

Luciano Costa Blomberg [a]
Eduardo Gonçalves Mota [b]
José Antônio Poli de Figueiredo [b]
Duncan Dubugras A. Ruiz [a]

[a] Graduate Program in Computational Science, Pontifical Catholic University of Rio Grande do Sul, Porto Alegre, RS, Brazil
[b] Graduate Program in Dentistry, Pontifical Catholic University of Rio Grande do Sul, Porto Alegre, RS, Brazil

Abstract

Purpose: The aim of this article was to present the conceptual model and development of an oral health database for storage, organization, and recovery of data from dental clinical records of a university dental school.

Method: This is a joint project of the Graduate Programs in Dentistry and in Computational Science of the Pontifical Catholic University of Rio Grande do Sul (PUCRS). During a period of three months in 2008, semi-structured interviews were conducted with the dental faculty of five clinical areas of the PUCRS Dental School (FO-PUCRS). Collected data were used to build the database language and architecture.

Results: Conceptual, logical, and physical database models were developed, as well as the generation of a script in SQL (Structured Query Language). After generating the script, the tool Oracle SQL Developer (Oracle Corporation, Redwood Shores, CA, USA) was used to import and create the FO-PUCRS OHDB (Oral Health Database) implemented in the Oracle 10g, a relational database management system (RDBMS).

Conclusion: The FO-PUCRS OHDB represents an important advancement towards automating the institutional operational demands and will contribute to organize the clinical and social information of the patient population for administrative, educational, and research purposes.

Key words: Oral health; data modeling; database; clinical records

Resumo

Objetivo: Este artigo teve como objetivo documentar o trabalho realizado em conjunto pelos Programas de Pós-Graduação em Odontologia (PPGO), e Ciência da Computação (PPGCC) da Pontifícia Universidade Católica do Rio Grande do Sul (PUCRS), visando à construção de um ambiente de armazenamento, organização e recuperação de dados provenientes de prontuários odontológicos.

Método: Durante 3 meses em 2008 foram realizadas entrevistas semiestruturadas com docentes de cinco áreas de atendimento odontológico da Faculdade de Odontologia (FO-PUCRS) para análise e coleta de requisitos.

Resultados: Com base nas informações obtidas neste processo, foram desenvolvidas as modelagens conceitual, lógica e física do banco de dados, bem como a geração de um *script* em SQL (Structured Query Language). Após a geração do *script*, foi utilizada a ferramenta Oracle SQL Developer (Oracle Corporation, Redwood Shores, CA, USA) para sua importação e criação do BDSB (Banco de Dados em Saúde Bucal) implementado no SGBD (Sistema Gerenciador de Banco de Dados) relacional Oracle 10g.

Conclusão: A construção do BDSB junto ao PPGO-PUCRS representa um importante avanço em direção a automatização das demandas operacionais da organização, mas principalmente a exploração de uma fonte valiosa de conhecimento, rica em informações clínicas e sociais.

Palavras-chave: Saúde bucal; modelagem de dados; base de dados

Correspondence:
Eduardo Gonçalves Mota
Av. Ipiranga, 6681
Porto Alegre, RS – Brazil
90619-900
E-mail: eduardo.mota@pucrs.br

Received: April 14, 2008
Accepted: June 21, 2009

Introduction

The Dental School of the Pontifical Catholic University of Rio Grande do Sul (FO-PUCRS) is located in Porto Alegre city, in the South region of Brazil. It has over 50 years of experience in teaching, research, and extension activities through the dental undergraduate and graduate programs. The *stricto sensu* graduate program in Dentistry comprises the fields of Oral Surgery & Maxillofacial Traumatology, Restorative Dentistry, Endodontics, Oral Medicine, Dental Materials, Orthodontics & Facial Orthopedics, and Prosthodontics. Thus the daily activity at the dental school produces a large amount of clinical data, laboratory and image exams (radiography, tomography, and echography), and dental charts, which makes the manual data analysis an extremely exhausting process.

Automated databases derived from this type of environment may constitute a rich source of new knowledge, which can be used for a comprehensive analysis of the relationship between oral pathologies and the socio-demographic and clinical profile of patients. For example, according to Baldani et al. (1), social and socioeconomic indicators of income, housing, educational level, access to dental services and water fluoridation were obtained from the governmental database of DATASUS (Computational department of SUS) and IBGE (Brazilian Geographic and Statistical Institute) for the analysis of the association of these variables with the DMF-T index (decayed, missing and filled teeth) in the State of Paraná. Similarly, the study of Fernandes and Peres (2) explored the association of dental indicators (basically first appointment, annual mean of the 0-14 year-old population receiving public dental services, and the proportion of permanent tooth extractions in relation to the total dental procedures) with social, economic, and municipal service indicators, such as the human development index, child development index, and the population of the state of Santa Catarina. Cunha and Dias (3) presented a case study involving the exploration of a clinical database managed by the Microsoft Access 2000 database (Microsoft Corporation, Redmond, WA, USA), in which the data mining algorithms were applied (C4.5 and Apriori) (4) to identify the patient's profile and the treatment procedures adopted.

In relation to the design building of oral health databases, the studies by Taylor (5-7) showed the definition of logical data structure and development of a computational application for the management of clinical information of 6-16 year-old children in the United Kingdom. The construction of a dental database can add a series of operational and research benefits to the institution. For instance, the manual process for management of patients' historical data (appointments, diagnosis, clinical exams and treatment) is replaced by the automated process, which allows the integration of these data into the different areas of dental attendance. Simultaneously, the diversity of items of information and the treatment of texts, temporal data, and later inclusion of DICOM images (Digital Imaging Communications in Medicine) may offer new fields for research, such as the natural language processing and image recognition.

Therefore, this joint project of the Graduate Programs in Dentistry and in Computational Science of the Pontifical Catholic University of Rio Grande do Sul (PUCRS) aimed to document the building process of the OHDB (Oral Health Database) for the data storage, organization, and recovery of historical clinical records of the fields of General Dentistry, Oral Medicine, Dental Trauma, Cleft Palate Fissures, and Community Dentistry from the last 20 years.

Methods and Results

The method adopted to build the OHDB was based on the stages proposed by Elmasri and Navathe (8) (Fig. 1).

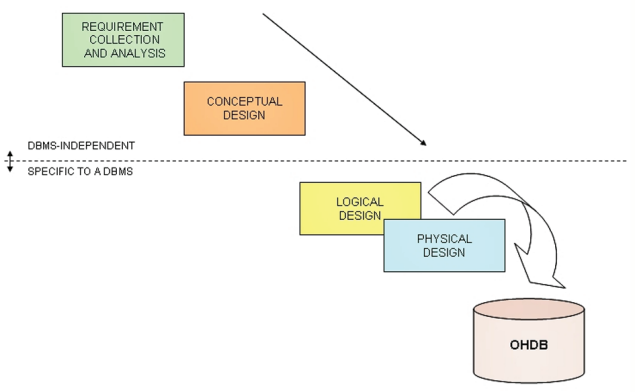

Fig. 1. Schematic diagram of the OHDB building design.

Preliminary requirement collection and analysis was carried out during a period of three months in 2008. Semi-structured interviews were used to collect information from the clinical faculty of FO-PUCRS in the following areas: the University Extension Center Vila Fátima (SUS Health Unit), São Lucas Hospital (Oral Medicine Services), CERLAP (Labiopalatal Fissure Rehabilitation Center), Endodontics, and Dental Trauma. The information retrieved from the interviews were: process of data collection and storage, existing clinical records, type of data generated (texts, numbers, historical data, and image sequences), and other components of the process (appointments, anamneses, clinical and complementary exams, diagnosis, and treatment). At the end of each interview, minutes were produced containing all the information collected and a class diagram in UML notation (Unified Modeling Language) (9).

After this first stage, the documentation generated for each area was analyzed to result in a high level abstract model using a class diagram in UML notation. This conceptual model was a concise description of data requirements (8) including detailed descriptions of the types of classes, attributes, associations, and restrictions. Once the conceptual model was elaborated, it was validated with the interviewed faculty, and adjustments were made to the class nomenclature and diagram attributes. The final diagram aimed to integrate all the areas and had 85 classes, 241 attributes, and 96 associations (Fig. 2).

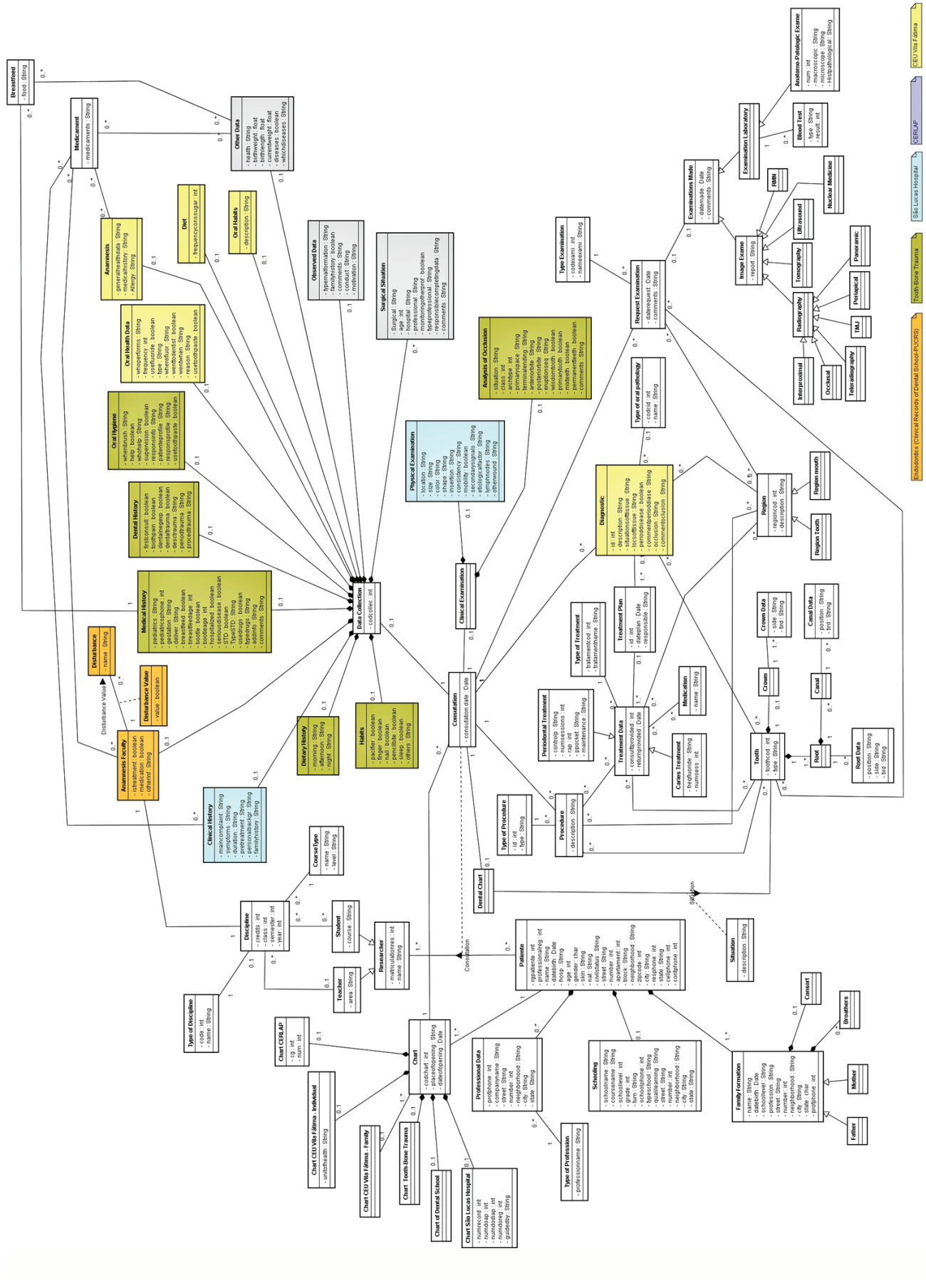

Fig. 2. Conceptual modeling of OHDB performed with the CASE Jude Community tool (Change Vision, Shinjuku, Tokyo, Japan).

The next stage was to develop the logical data model in the relational Oracle OHDB using the SQL language. In SQL, nomenclature conventions such as class, attribute, and association are replaced by tables, columns, and integrity restrictions. In the logic model, the data structures and their implementation details, such as optimization of the number of tables and implementation of primary and foreign keys were defined. Figure 3A shows the edition of the logic model built for this study, by means of the CASE tool (Computer Aided Software Engineering) DBDesigner 4 (Free Software Foundation, Boston, MA, USA).

Additionally, other modifications were included in the final model with regard to the conceptual model such as the inclusion of the column Patient_rg as a primary key in the Chart table (Fig. 3A). This modification concerns the type of relationship (1 - *, Merise cardinality notation) between the Chart and Patient tables; that is, a record of the Patient table may be connected to several (*) records in the Chart table (at least one), while the record of the Chart table must be connected to a single record of the Patient table (at least one).

At the end of this stage the exploration module of the DBDesigner 4 tool was used to generate a script in SQL language with the specificities of Oracle OHDB (Fig. 3B). Although previous studies (8,10,11) included the physical modeling as an independent stage in the database building design, several CASE tools deal with physical aspects, such as storage structures and access pathways (indexes) within the logic data modeling stage itself. For this reason, this study at first adopted the standard definition of the CASE tool DBDesigner 4 to compose the physical scheme of the OHDB in Oracle.

After generating the SQL script, the Oracle SQL Developer tool (Oracle Corporation, Redwood Shores, CA, USA) was used to import and generate the OHDB, as partially shown in Figure 3C. Some details with reference to the physical structure of the Chart table data are the definition of primary keys, required fields (information), and types of specific data of Oracle 10g relational OHDB.

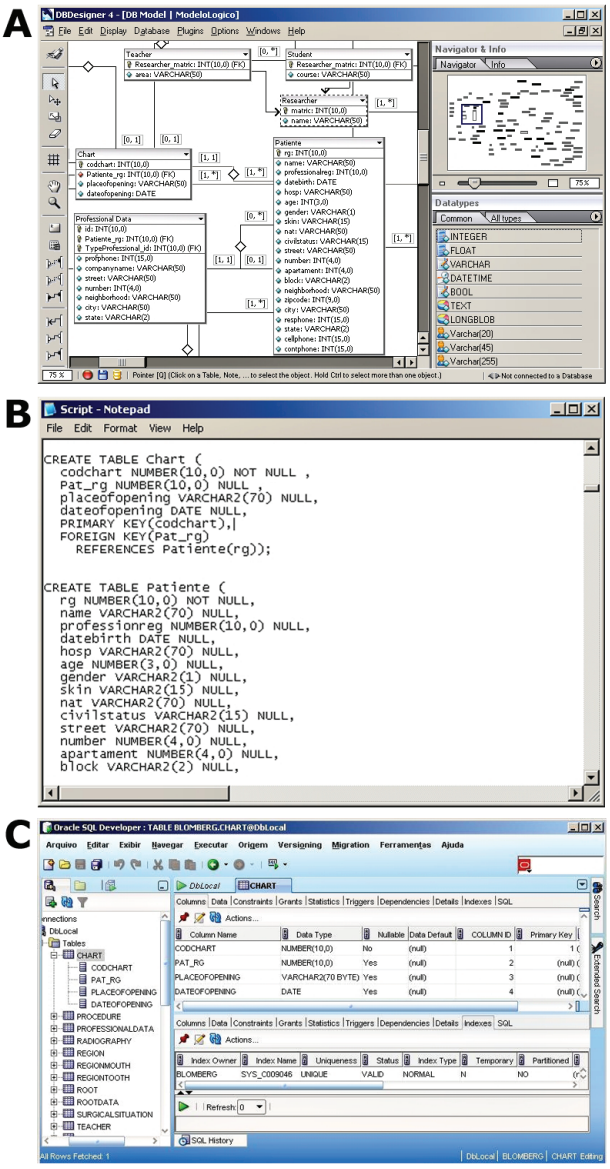

Fig. 3. **(A)** Logical modeling of OHDB performed in the DBDesigner 4 software. **(B)** SQL script generated in the DBDesigner 4 software. **(C)** Database implemented in the Oracle 10g OHDB.

Discussion

The science history shows periodical changes in knowledge production and dissemination of information, which requires effective resources to storage and recover data. Thus, the present project was developed due to the necessity to describe the process of building an oral health database, as database building design has been previously explored in several areas of knowledge but specific studies in the dental literature still are scarce.

Similar works in Dentistry were developed by Taylor (5-7), but those studies concentrated on the logic structure of data and not on the complete database building design (Table 1). The data explored in the present project are broader in scope than those explored by Taylor, since the public attended was composed of several age groups and not only school-age children.

Table 1. General considerations about the Oral Health Database Building Design.

	Taylor	Our approach
Data source	Data from clinical questionnaires stored in ASCII files	Data from clinical records
Areas of Attendance	Clinical	Clinical, Stomatology, Dento-Alveolar Trauma, Cleft Palate Fissures and Collective Health
Stages of the database design implemented	Logic	Requirement collection and analysis, conceptual, logic and physical*

* According to Elmasri and Navathe, 2006 (8).

Future studies intend to use the OHDB as a basis for the development of an analytical model for metric management and oral health indicators for the target population. Other objectives include the application of the automated data mining techniques to identify the clinical and social patient's profile related to the incidence of oral pathologies (dental caries, oral cancer, malocclusion, and periodontal diseases). Therefore, the creation of the OHDB following the proposed model represents an important contribution to education and research in a dental facility and school because of its broad scope (number of patients and age groups involved), clinical information, diversity of data (texts, numbers, temporal data), and areas of knowledge. Furthermore, the construction of this database will provide a series of benefits to the organization, such as automation of operational demands to optimize administrative systems.

Acknowledgments

This study was supported by the National Council of Scientific and Technological Development (CNPq) – research grant MCT/CNPq27/2007.

References

1. Baldani MH, Vasconcelos AG, Antunes JL. Associação do índice CPO-D com indicadores sócio-econômicos e de provisão de serviços odontológicos no Estado do Paraná, Brasil. Cad. Saúde Pública 2004;20:143-52.
2. Fernandes LS, Peres, MA. Associação entre atenção básica em saúde bucal e indicadores socioeconômicos municipais. Rev. Saúde Pública 2005;39:930-6.
3. Cunha GG, Dias CR. Aplicação de mineração de dados em uma base odontológica. Rev Eletrônica Fac Metodista de Granbery [serial on the Internet]. 2007 Jul [cited 2009 February 2] Available from: http://re.granbery.edu.br/index.php?
4. Tan P, Steinbach M, Kumar V. Introduction to data mining. Boston: Addison-Wesley; c2006.
5. Taylor D, Naguib RN, Boulton S. A dynamic clinical dental relational database. IEEE Trans Inf Technol Biomed 2004;8:298-305.
6. Taylor D, Naguib RN, Amin S, James A, Boulton S. Storage, querying and visualisation of clinical dental records. In: Engineering in Medicine and Biology Society. Proceedings of the 23rd Annual EMBS International Conference; 2001 Oct 25-28;Istanbul, Turkey. Istanbul: IEEE;2001. p.3718-20.
7. Taylor D, Naguib RN, Amin S, James A, Boulton S. Design considerations for a relational database schema to hold dental epidemiological data. In: Information Technology Applications in Biomedicine. Proceedings IEEE EMBS International Conference; 2000 Nov. 9-10; Arlington (VA). Arlington: IEEE;2000. p.342-6.
8. Elmasri R, Navathe SB. Sistemas de banco de dados. 4.ed. São Paulo: Pearson Education; 2006.
9. Rumbaugh J, Jacobson I, Booch G. The unified modeling language reference manual. Reading: Addison-Wesley; c1999.
10. Heuser CA. Projeto de banco de dados. 5.ed. Porto Alegre: Sagra-DC Luzzatto; 2004.
11. Atzeni P, Ceri S, Paraboschi S, Torlone R. Database systems: concepts, languages & architecture. London: McGraw-Hill; 2000.