

Short report

Development of a database for cervical cytology reports

Eduardo Dreizzen MD DSc

Professor, Department of Medical Informatics, Faculty of Medical Sciences, National University of La Plata (UNLP), Argentina

María Silvia Lima MD

Lecturer, Department of Medical Informatics and Department of Pathology, Faculty of Medical Sciences, National University of La Plata (UNLP), Argentina

Rosa Beatriz Kitagawa MD

Lecturer, Department of Medical Informatics and Department of Histology, Cytology and Embryology, Faculty of Medical Sciences, National University of La Plata (UNLP), Argentina

Félix José Corrons MD

Professor, Department of Pathology and Department of Histology, Cytology and Embryology, Faculty of Medical Sciences, National University of La Plata (UNLP), Argentina

Oswaldo Spinelli MD

Professor, Department of Medical Informatics, Faculty of Medical Sciences, National University of La Plata (UNLP), Argentina

ABSTRACT

Background Cervical cancer is an important cause of female mortality. Its incidence, however, has been drastically reduced by early detection using the Papanicolaou screening method. As a result of the widespread use of this method, a large amount of cytological information has to be stored, printed and handled efficiently. This paper describes a database system designed to manage information from studies of cervical cytology, which was developed at the Faculty of Medical Sciences, National University of La Plata, Argentina.

Methods Information needs were assessed by interviews with physicians and analysis of cytology reports. The database was developed using MS Access 2007.

Results The database is designed to collect, display, sort and print patient demographics, clinical data and cytological studies. Diagnoses can be stored either as international classifications or as free text. The database contains 45 tables and 50 forms for data entry, editing and display. This information can be exported to spreadsheets and statistical packages.

Conclusion This database gives users easy access to patients' cytological diagnoses and provides a useful tool for physicians and public health researchers.

Keywords: cervical smears, data collection, database management systems, computerised, medical informatics, medical records systems, software design

What this study adds

- This study shows the implementation of a database software for single users which is a useful tool for cervical cytology reporting.
- The database software can be easily distributed and installed; it requires only brief training, and can properly protect the confidentiality of patient data.
- The database gives the choice to use free text, a list of personal options or international classifications. This feature is very useful to speed up data entry and allows the use of the information for statistical and epidemiological studies.

Introduction

Uterine cervical cancer is the second most common neoplasm among women. Despite the big efforts made in its prevention, early diagnosis and the treatment of precursor lesions, it still is an important cause of death in the third world and is the fifth most common cause of death in women worldwide.¹ Cervical cytology screening is the standard method of early diagnosis of cervical cancer and has resulted in a drastic decrease in the incidence of the disease. As a result of the worldwide use of this method, a large amount of information has to be stored, handled and sent to physicians. The process of storing, modifying and transmitting this information is greatly simplified by the use of electronic medical records. The purpose of this paper is to describe the development and characteristics of a database designed to handle information resulting from uterine cervical cytology studies.

Development

The database program was built by members of the Faculty of Medical Sciences, National University of La Plata, Argentina, working at the Medical Informatics, Pathology and Histology Departments. MS Access 2007 was used in its development.

The usual steps in relational database design were followed.^{2,3}

- identification of users' needs, which was performed by interviews and analysis of cytology reports
- relational data model definition
- building of tables, primary and foreign keys, indexes and relationships
- user interface design
- interface programming using Visual Basic for Applications⁴
- debugging
- users' training.

Classifications for cervical cytology

Clinical coding and classification processes transform natural language descriptions in clinical text into data that can subsequently be used for clinical care and research.⁵ Classifications used in standardised reports are designed to improve clinical understanding of complex pathology data, and this understanding pro-

vides guidance for clinical management.⁶ Several standardised classifications for cervical cytology have been proposed; however, none of them reaches an absolute consensus.^{1,7-9} The Bethesda System, which was introduced for the first time in 1988 and revised in 1991 and 2001, includes specific statements about specimen adequacy, general categorisation, interpretation and results. In addition, it introduced a standardised approach for reporting whether an individual specimen is adequate for evaluation.¹⁰ The Bethesda System 2001 was developed through a process that involved committee review of the literature and solicitation of expert opinions.^{10,11} It includes revisions in statements of adequacy, general categorisation and interpretation, and the results of epithelial cell abnormalities. Although the Bethesda System is in widespread use, in many cases other classifications, either personal or proposed by local expert committees, are preferred. These issues have been taken into account in the development of the database.

Application

General characteristics

We developed a relational database containing 45 tables, 50 forms and 11 000 programming lines. The database consists of two files: a front-end containing the user interface, forms, reports and Visual Basic for Applications (VBA) code, and a back-end, containing the data. SQL language, which is integrated into the Access interface, is used as the source of forms and reports.

The interface was implemented using forms that allow easy entry and editing of information. The programs displays detailed lists, summaries and reports that can be used to perform statistical and epidemiological studies. The diagnoses can be written as international classifications such as Papanicolaou classification, Dysplasia-Carcinoma In Situ Classification (CIS), Cervical Intraepithelial Neoplasia Terminology (CIN System), World Health Organization and International Society of Gynecological Pathologists Terminology (WHO/ISGYP) and Bethesda System for reporting cervical/vaginal cytological diagnosis;^{1,7-9} they can also be entered as personal codes. Printed reports can be designed according to users' preferences.

Interface design

We used VBA programming language to automate the verification tasks and to display help messages.⁴ The

Información personal: Acevedo, Vanesa ID #73

Nombre: Vanesa
 Apellido*: Acevedo
 Apellido de casada:
 Tipo de Documento: DNI
 Documento Nº: 26946213
 Fecha de nacimiento: 12/12/1979
 Edad: 30
 Estado Civil: Casada
 Nro. de ID: 73

Deshacer cambios Cerrar

Figure 1 Forms for entering and editing patients' personal information. At the bottom of each form, the left-hand button cancels the last changes or deletes the current record and the right-hand button closes the form and saves the record. Required fields are marked with an asterisk. The title bar contains the patient's name and the type of information shown. The grey fields contain information that cannot be edited from inside the form

forms have similar characteristics across the application (Figures 1 and 2). Each form has two buttons at the bottom right-hand corner: one used to undo the last entry or delete the record and the other is used to save the data and close the form. Data are saved automatically when the user moves to another record or closes the form. Before saving the record, the application checks for the right data format, mandatory fields and duplicated fields. User confirmation is required in the case of data modification. The interface includes combo boxes and list boxes which allow the user to choose among prearranged options. These controls enhance the speed of data entry and reduce the probability of mistakes. The combo and list boxes can be easily modified by the user to add new options or modify existing ones. Forms are opened

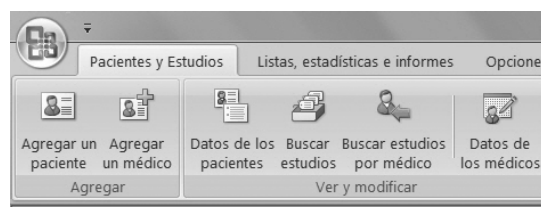


Figure 3 Tabs and groups from the application's ribbon

through a custom ribbon which replaces the MS Access 2007 ribbon and has tabs and controls specially designed for the program (Figure 3).

Patients have an unique ID number that is automatically generated when their information is saved for the first time. This ID number ensures a unique identification across the database and cannot be modified by the users. The database allows entering duplicate names; in this case an alert appears asking for confirmation before saving the record (Figure 4).

Personal and medical information is accessed through forms containing lists that allow searches to be made either by name, patient ID or study ID (Figure 5). In order to protect confidentiality, the database is encrypted and password protected, according to the programming procedures used in Access 2007 for accdb and accde files.⁴ The database generates lists and summaries that can be ordered and filtered according to different criteria such as date, physician name, patient name and diagnosis. These lists may be exported to word processors, spreadsheets and statistical packages, for statistic or epidemiological studies.

Detailed characteristics

The database can be used to enter, edit, search, order and display the following information:

Soriano, Karina ID #65 - Protocolo #93.860 - Diagnosticado el 30/07/2009

General Citología Conclusiones Equivalencias diagnósticas

Células superficiales: %
 Células intermedias: %
 Células parabasales: %
 Células Epiteliales Malpighianas: Presencia aislada de células intermedias con displasia severa.
 Células Epiteliales Cilíndricas: Células endocervicales con metaplasia escamosa.
 Otros Elementos:
 Flora: Inespecifica.
 Exudado:

Plegamiento: [dropdown]
 Citólisis: [dropdown]
 Trofismo: Hipotrófico

Ver... Borrar
 Ver... Borrar
 Ver... Borrar
 Ver... Borrar

Vista de impresión Deshacer Cerrar

Figure 2 Form for editing cytological data. At the bottom, the third button from the right can be used to preview the cytological report. Several fields can be filled either as free text or using a list of user-defined options

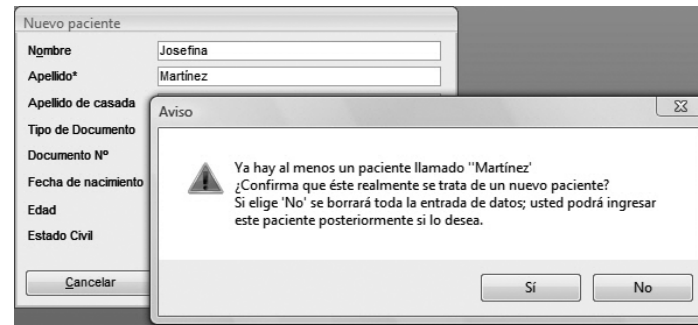


Figure 4 Warning for duplicate names. When an attempt is made to enter a patient's name that is already stored in the database, a warning message pops up and the user can either store the new patient's data, or undo the entry

Nombre e ID	Médico	Fecha	Protocolo	Papanicolaou	Conclusión diagnóstica
Acevedo, Vanesa ID #73	GJ	20/11/2009	94829	II	Sugestivo de infección por HPV.
Astor, Ramona ID #59	TS	14/10/2009	94.459	II	Cambios celulares reactivos asociados a inflamación
Astudillo, Fidelina ID #74	TS	14/10/2009	94429	I	Citología dentro de límites normales.
Baldasarre, Etelvina ID #66	NL	01/09/2009	94.126	II	Cambios celulares reactivos asociados a inflamación
Bartolomé, Samanta ID #57	NL	14/10/2009	94.436	II	Cambios celulares reactivos asociados a inflamación
Borges, Sonia ID #56	NL	16/10/2009	94.435	II	Células epiteliales escamosas con coilocitosis y binucleación
Erdosain, Romina ID #61	VN	12/11/2009	94.676	II	Células epiteliales escamosas con coilocitosis y binucleación
Giménez, Elsa ID #58	TS	14/10/2009	94.438	I	Citología dentro de límites normales.
Llorente, Victoria ID #64	NL	31/08/2009	94.103	II	Cambios celulares reactivos asociados a inflamación
Martínez, Gisela ID #52	NL	13/10/2009	94.434	II	Cambios celulares reactivos asociados a inflamación
Martínez, Luisa ID #70	NL	15/09/2009	94220	II	Cambios celulares reactivos asociados a inflamación
Méndez, Alicia ID #63	TS	31/08/2009	94.104	II	Cambios celulares reactivos asociados a inflamación
Miracoli, Alejandra ID #62	NL	15/09/2009	94.252	II	Cambios celulares por atrofia.
Orozco, Dora ID #72	TS	16/11/2009	94746	II	Cambios celulares reactivos asociados a inflamación
Pérez, Lorena ID #75	TS	23/07/2009	93741	IV	Lesión epitelial escamosa de alto grado (HSIL).
Soriano, Karina ID #65	TS	30/07/2009	93.860	IV	Lesión epitelial escamosa de alto grado (HSIL).

Figure 5 Search form for cytological studies

- patient information: name, date of birth, age, address, phone, insurance
- medical information: gynaecological history, medications, biopsies
- specimen type and adequacy
- general categorisation
- organisms and other non-neoplastic findings
- epithelial squamous and glandular cell abnormalities
- comments and recommendations
- cytological diagnoses may be expressed both as international classifications (Papanicolaou classification, Dysplasia–CIS, CIN System, WHO/ISGYP or Bethesda System) or free text
- printed reports may be designed according to users' preferences.

Discussion

Principal features

This study shows the development of a database for cervical cytology reports. The database software can be

easily distributed and installed; it requires only brief training and can properly protect the confidentiality of patient data.

The ability to report diagnoses either as free text or as cytological classifications is an important feature of the database. When using free text, users can freely write their diagnoses and then store them for reuse. Alternatively, they can choose from several predefined diagnoses from different classifications (Papanicolaou classification, Dysplasia–CIS, CIN System, WHO/ISGYP or Bethesda System) (Figure 6). Users can also create and edit personal classifications. In the case of options such as organisms, non-neoplastic findings, specimen adequacy and general categorisation, the program enables free text writing that can be stored for reuse.

After a testing period, the database is currently being used for reporting cytological findings at the Pathology Department of the Alejandro Korn Hospital, which is a large public institution that annually processes 3500 cytological studies from inpatients, outpatients and peripheral health units. The software is being used by members of the Pathology Department for processing information from cytological

Figure 6 Form used for choosing cytological classifications and diagnoses

studies and sending the results to physicians as printed reports.

Unlike these systems, the database described in this work is mainly intended for single users and represents a resource of low cost and ease of use.

Cervical cancer screening and electronic medical records

Many different agents have been implicated in the aetiology of the cervical cancer, but current evidence indicates that certain types of human papillomavirus (HPV) are the most important aetiological factor for the development of cervical cancer. In 1995, the International Agency for Research on Cancer categorised HPV 16, HPV 18 and other HPV types as carcinogenic factors.¹² Since the pioneering studies carried out by Papanicolaou¹³ and their application on cervical cancer early diagnosis, the cytology study is the standard for the secondary prevention of the cervical cancer and has been used by national public health organisations for the last 40 years. Since the development of cytology based cervical cancer screening a drastic reduction in the rates of cervical cancer has been observed. It is currently estimated that systematic screening can reduce death rates from cervical cancer by 70% or more.¹⁴ The strength of the cervical screening method is based on a large experience in its use, low cost, high specificity and the fact that the lesions can be easily manageable.

As a result of the widespread use of cervical screening, a large amount of cytological information has to be stored, printed and handled efficiently. Such information is usually written on reports and sent to the attending physicians. Electronic medical records have some clear advantages over paper storage of medical information: they are not static; they improve the legibility of clinical notes; they can be presented at multiple locations; data can be retrieved and sorted on the basis of user-defined criteria; and data can be used in statistical and epidemiological studies without transcripts.^{15–17}

Some information systems used by several institutions to handle cervical cytology and colposcopic image reports have been reported in the literature.^{18–23}

System limitations

It must be noted that the software is not a multi-user application, and it is not designed to be used over a network; therefore it may have a limited usefulness in an institution with several different workstations being used for cytology reporting.

In a future line of work we plan to add the ability to upload microscopy images to the database which will be used for clinical and teaching purposes.

Conclusions

In summary, the final product of this development is a database system that provides a useful tool for handling and reporting information from studies of cervical cytology using the accessibility and convenience of computers and software.

REFERENCES

- 1 Zeferino LC and Derchain SF. Cervical cancer in the developing world. *Best Practice & Research Clinical Obstetrics & Gynaecology* 2006;20(3):339–54.
- 2 Riordan R. *Designing Relational Database Systems*. Redmond: Microsoft Press, 1999.
- 3 Groh MR, Stockman JC, Powell G, Prague CN, Irwin MR and Reardon J. *Access™ 2007 Bible*. Indianapolis, IN: Wiley, 2007.
- 4 Hennig T, Cooper R, Griffith G and Stein A. *Access™ 2007 VBA Programmer's Reference*. Indianapolis, IN: Wiley, 2007.
- 5 Stanfill MH, Williams M, Fenton SH, Jenders RA and Hersh WR. A systematic literature review of automated clinical coding and classification systems. *Journal of the*

- American Medical Informatics Association* 2010;17(6): 646–51.
- 6 Raab SS. Human papillomavirus reporting: impact on Bethesda cytology reports. *Archives of Pathology & Laboratory Medicine* 2003;127(8):969–72.
 - 7 Lacruz Pelea C. Nomenclatura de las lesiones cervicales (de Papanicolaou a Bethesda 2001). *Revista Española de Patología* 2003;36(1):5–10.
 - 8 Kocjan G, Priollet BC, Desai M et al. BSCC, Bethesda or other? Terminology in cervical cytology European panel discussion. *Cytopathology* 2005;16(3):113–19.
 - 9 Herbert A, Bergeron C, Wiener H et al. European guidelines for quality assurance in cervical cancer screening: recommendations for cervical cytology terminology. *Cytopathology* 2007;18(4):213–19.
 - 10 Apgar BS, Zoschnick L and Wright TC Jr. The 2001 Bethesda System terminology. *American Family Physician* 2003;68(10):1992–8.
 - 11 Solomon D, Davey D, Kurman R et al. The 2001 Bethesda System: terminology for reporting results of cervical cytology. *Journal of the American Medical Association* 2002;287(16):2114–19.
 - 12 International Agency for Research in Cancer (IARC). *IARC Monographs- Classifications – Group 1*. <http://monographs.iarc.fr/ENG/Classification/crthgr01.php>. (accessed 4 May 2010).
 - 13 Papanicolaou GN. Cytologic studies in diagnosis of carcinoma. *Journal of the International College of Surgeons* 1954;21(4):419–26.
 - 14 Kitchener HC, Castle PE and Cox JT. Chapter 7: Achievements and limitations of cervical cytology screening. *Vaccine* 2006;24(S3):63–70.
 - 15 Fraser HS, Biondich P, Moodley D, Choi S, Mamlin BW and Szolovits P. Implementing electronic medical record systems in developing countries. *Informatics in Primary Care* 2005;13(2):83–95.
 - 16 Haux R, Knaup P and Leiner F. On educating about medical data management – the other side of the electronic health record. *Methods of Information in Medicine* 2007;46(1):74–9.
 - 17 Kalogiropoulos NA, Baran J, Nimunkar AJ and Webster JG. Electronic medical record systems for developing countries: review. *Conference Proceedings of the IEEE Engineering in Medicine & Biology Society* 2009 September;1730–3.
 - 18 Terpos AA, Efstratiadou M, Symiakaki H et al. Building up a computerized follow-up register and information system for cervical cytology. *International Journal of Biomedical Computing* 1995;39(3):277–85.
 - 19 Golabek JK, Brueckner PJ and Seidenfeld AM. CytoBase: an electronic medical record for cervical cytology. *Proceedings of the AMIA Annual Fall Symposium* 1997:37–41.
 - 20 Puig-Tintoré LM, Galceran J, Torné A, Grau A and Puig-Soler LM. Sistema integrado de bases de datos clínicos e imágenes colposcópicas, en la prevención del cáncer del tracto genital inferior de la mujer. XVIII Congreso Anual de la Sociedad Española de Ingeniería Biomédica, 2000; 291–4.
 - 21 Bucchi L, Pierri C, Caprara L, Cortecchia S, De Lillo M and Bondi A. Computerized monitoring for integrated cervical screening. Rationale, methods and indicators of participation. *Pathologica* 2003;95(1):9–21.
 - 22 Matsopoulos GK, Kouloulis V, Asvestas P, Mouravliansky N, Delibasis K and Demetriades D. MITIS: a www-based medical system for managing and processing gynecological–obstetrical–radiological data. *Computer Methods and Programs in Biomedicine* 2004;76(1):53–71.
 - 23 Raab SS, Grzybicki DM, Zarbo RJ, Meier FA, Geyer SJ and Jensen C. Anatomic pathology databases and patient safety. *Archives of Pathology & Laboratory Medicine* 2005;129(10):1246–51.

ACKNOWLEDGEMENTS

The project was fully sponsored by the National University of La Plata (UNLP) with grant number M128/08.

CONFLICTS OF INTEREST

None.

ADDRESS FOR CORRESPONDENCE

Eduardo Dreizzen
 Área de Informática Médica
 Facultad de Ciencias Médicas
 Universidad Nacional de La Plata
 60 y 120
 (1900) La Plata
 Argentina
 Email: edr@med.unlp.edu.ar

Accepted September 2011