Article

Health Informatics Journal

Health Informatics Journal
2015, Vol. 21(1) 36–45

DOI: 10.1177/1460458213496661
jhi.sagepub.com

Brain Tumor Database, a free relational database for collection and analysis of brain tumor patient information

Maurizio Bergamino
University of Genoa, Italy

David J Hamilton
George Mason University, USA

Lara Castelletti, Laura Barletta and Lucio Castellan
San Martino University Hospital, Italy

Abstract

In this study, we describe the development and utilization of a relational database designed to manage the clinical and radiological data of patients with brain tumors. The Brain Tumor Database was implemented using MySQL v.5.0, while the graphical user interface was created using PHP and HTML, thus making it easily accessible through a web browser. This web-based approach allows for multiple institutions to potentially access the database. The BT Database can record brain tumor patient information (e.g. clinical features, anatomical attributes, and radiological characteristics) and be used for clinical and research purposes. Analytic tools to automatically generate statistics and different plots are provided. The BT Database is a free and powerful user-friendly tool with a wide range of possible clinical and research applications in neurology and neurosurgery. The BT Database graphical user interface source code and manual are freely available at http://tumorsdatabase.altervista.org.

Keywords

brain tumors, MySQL, patient information, relational database, web-based approach

Corresponding author:
Maurizio Bergamino, Department of Neuroscience, Rehabilitation, Ophthalmology, Genetics, Maternal and Child Health, University of Genoa, Via De Toni, 5—16132 Genoa, Italy.
Email: maurizio.bergamino@gmail.com

Introduction

A tumor is any mass caused by abnormal and uncontrolled growth of cells. Tumors in the brain are categorized according to different factors, including the type of cells involved, their locations, and how quickly they grow.[1] Like each tumor, brain tumors (BTs) originate when a normal cell begins to grow abnormally in the brain and replicates too quickly. Impairment of cognitive functions (e.g. language, memory, visual-spatial perception, time-space orientation, attention, and executive function) occurs in almost all patients with BT and eventually compromises their independence.[2,3]

Among desired capabilities in medicine, ease of data collection and analysis of patients undergoing medical examinations in a hospital, such as patients with BT, is certainly one of the most important. Use of clinical databases is an easy method to collect and display data regarding these patients. Today, databases are valuable and important tools that permit the collection and retrieval of relevant patient data.[4–6] For instance, personal information about patients with BT, together with information like imagery, surgery, diagnostic, and therapy, can be inserted into a database in order to simplify the job of a physician from a clinical and statistical point of view. Standards assuring information accuracy and easy access to stored data must coexist with criteria that protect an individual's identifiable health information. In the organization of a database, the support of informatics for medical applications is fundamental, especially when the number of patients has grown together with the number of different diagnostic techniques and treatments.

In this study, the design and application of a free relational database to manage data from patients with BT is described. Additionally, we discuss efforts to balance simplicity of access with the ability to share and retrieve BT patient information, together with personal health information (PHI) preservation.

Methods

Database development and installation

MySQL v5.0[7,8] was the relational database used to develop the BT Database (BT-DB). PHP v.5.2.9[9] and HTML[10] programming languages were used to create the graphical user interface (GUI). MySQL is a popular open source relational database management system (RDBMS) and can be downloaded from http://www.mysql.com. PHP is a freely available (http://php.net/index.php) general-purpose scripting language specifically designed for the web that has extensive capabilities to interact with RDBMSs. GUI, implemented using PHP and HTML, simplifies user interaction with the database by making it accessible via generally available web browsers. A compressed file containing tutorials, installation instructions, and the required software is available, together with the database, on the BT-DB home page (http://tumorsdatabase.altervista.org).

Database access

PHI contained in the database must be protected with security safeguards against unauthorized access. Therefore, it is necessary to implement different modalities to access the database with different purposes. Additionally, easy access to the database must always be possible in emergency situations.

Users are provided with a user name and password with a specific level of privileges. Along these lines, four different privilege level profiles were implemented: Simple Consulter, Manager, Administrator, and Emergency Access.

1. Simple Consulter: the main purpose of a Simple Consulter is to retrieve data for statistical use and research. In fact, this user can utilize the tumor and search engines of the database in order to retrieve clinical data, but cannot have access to PHI. The user name and password that allow access with this modality are of limited duration and can thus be restricted to the time necessary for a single institutional review board-approved research project. This modality is important to share data across multiple institutions without disclosing PHI.
2. Manager: these users are allowed to interrogate and insert personal and clinical data in the database. In addition, Managers have access to all personal information across patients and may use the BT-DB for both clinical and statistical analysis.
3. Administrator: this user is allowed to access the database in a manner similar to the Manager and can also access the administration tool. The Administrator can perform BT-DB backups, change user name s/passwords of other users, and add new user profiles.
4. Emergency Access: in this modality, exclusive use by physicians in emergency situations is provided. With an emergency password, the physician can override database security for access to all data, including PHI.

Each access of the BT-DB is monitored. The database automatically populates a text log file containing historical records of transactions by recording user name, date, time, and IP address of the computer accessing BT-DB.

Description of BT-DB

BT-DB is free and can be installed on computers with different operating systems (OS) (e.g. Windows, Linux, and Mac). Using a specific user name and password, it is possible to access the main menu of BT-DB, which provides different options like insertion of new patient data, administrator tool, search engine, statistics, and tumor engine. The main features and functions provided by BT-DB are shown in Table 1.

Data entry and patient query

To minimize the risk of data entry errors, we implemented database constraints that require relations to satisfy certain properties, such as mandatory data and the use of check boxes, radio buttons, and pulldown menu in every webpage. Additionally, the user interface code reduces the risk of inaccurate data entry related to human error (e.g. typing errors). Figure 1 shows the screenshot of new patient data entry. In this page, the user can insert the personal information of a patient (e.g. first name, last name, date of birth, sex, and telephone number) and information on clinical presentations. In this page, some data are mandatory (date of insertion, first name, last name, and date of birth) and must be typed when the patient information is first entered. When the user attempts to enter the first name, last name and date of birth of a patient whose data already exists in the database, the user will be warned that the patient data is already stored. When new patient data is correctly inserted into the database, the user will be able to retrieve all patient and clinical data using the database query page (Figure 2(a)).

In the query page, the user can visualize and update the patient's personal information and clinical presentations. Also, the user can insert, modify, and remove different imaging data (e.g. Computerized Tomography (CT), Magnetic Resonance (MR) Imaging, MR spectroscopy, MR Imaging Blood Oxygen Level Dependent (BOLD), MR Diffusion Tensor Imaging (DTI), medical surgery, histological diagnosis, and medical therapy).

Table 1. Main features and functions provided by Brain Tumor Database (BT-DB).

Patient personal data	Last Name		
	Name		
	Date of birth		
	Sex		
	Address		
	Telephone		
	Department		
	Date of death		
Clinical Presentation	Date of first clinical sign		
	Clinical sign	Sensory deficit	
		Motor deficit	
		Epilepsy	
		Behavioral disorder	
		Headache	
Imaging	CT scan	Extra-axial	
		Intra-axial	
		Doubtful	
		Contrast enhancement	Homogeneous
			Inhomogeneous
			Ring
		Site	
	Morphological Magnetic Resonance	Extra-axial	
		Intra-axial	
		T2/FLAIR	
		FLAIR 3D	Tumor volume
		Diffusion	Hyper intense
			Hypo intense
		ADC	Reduced
			Increased
		Contrast enhancement	Homogeneous
			Inhomogeneous
			Ring
	Perfusion Magnetic Resonance	r-CBV value	
	Permeability Magnetic Resonance	Permeability value	
	Spectroscopy Magnetic Resonance	TE 135 ms/TE 30 ms	
		Naa/Cr value	
		Cho/Cr value	
		Lipids/Lactates	
		Myo Inositol	
	MR Blood Oxygen Level Dependent (BOLD)	Motor test	Site
		Sensory test	Site
		Language test	Broca activation
			Wernicke activation

(continued)

Table 1. (Continued)

Patient personal data	Last Name		
	RM Diffusion Tensor Imaging (DTI)	Fractional anisotropy value	
		Corticospinal tract	Infiltrated
			Compressed
		Fascicle arcuate	Infiltrated
			Compressed
		Superior longitudinal fascicle	Infiltrated
			Compressed
		Inferior fronto-occipital fascicle	Infiltrated
			Compressed
		Optic pathway	Infiltrated
			Compressed
Medical Surgery	Biopsy		
	Macroscopic total tumor resection		
	Partial tumor resection		
	Gliadel		
Histological diagnosis	Tumors engine	Name	
		ICD-O Code	
		WHO classification	
		Definition	
		Link Wikipedia	
		Note	
Medical therapy	Therapy	Conformational radiation therapy	
		Radiosurgery	
	Chemotherapy	Temozolomide	# cycles
		PC (V)	# cycles
		Fotemustine/ bevactizumab	# cycles
	Supportive therapy		

CT: computed tomography; FLAIR: fluid-attenuated inversion recovery; r-CBV: relative cerebral blood volume; NAA/Cr: *N*-acetylaspartate/creatinine; TE: echo times; ADC: apparent diffusion coefficient; Cho/Cr: choline/creatinine ratio; ICD-O: International Classification of Diseases for Oncology CODE. PC(V) : procarbazine, lomustine, and vincristine (PCV) chemotherapy

A PDF document containing a specific patient's information, including the Digital Imaging and Communications in Medicine (DICOM) images and clinical reports, can be created and managed by this page. Some examples of data insertion, such as CT, medical surgery, and medical therapy, are shown in Figure 2(b) to (d). It is useful to emphasize that the entry of variables is standardized across all database pages through the use of radio buttons, check boxes, and pulldown menus.

On the histological diagnosis page (Figure 3), the user may search for a tumor by name or by the International Classification of Diseases for Oncology (ICD-O) CODE, and retrieve essential

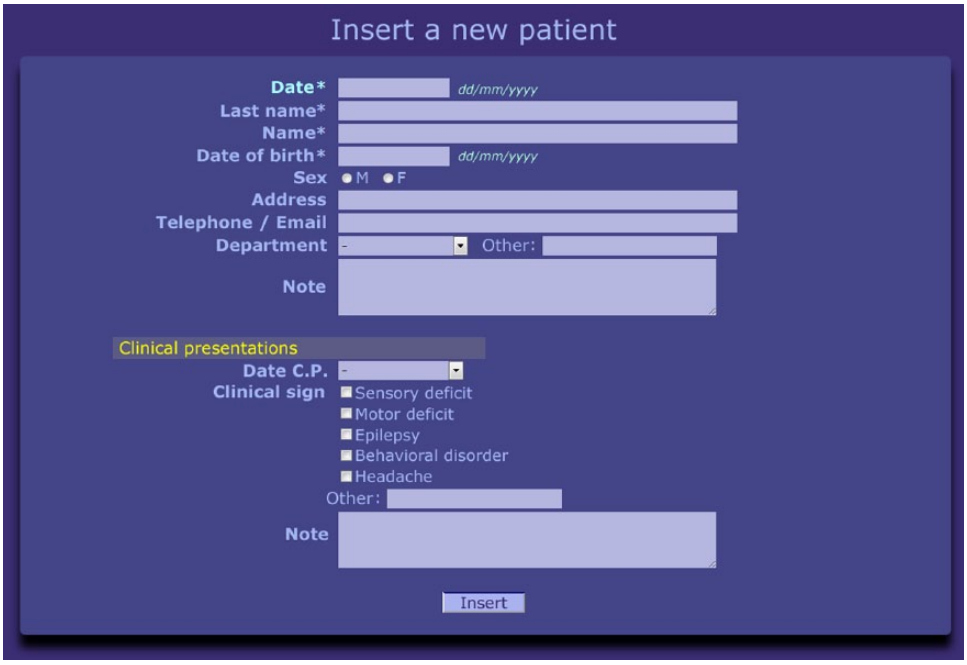

Figure 1. New patient screenshot. The user can insert personal data for a new patient and the clinical presentations. * are required fields and must be typed when the patient is entered for the first time.

information such as the tumor's ICD-O CODE, World Health Organization (WHO) classification, tumor's definition, and Wikipedia link (http://www.wikipedia.org).

Database search engine and statistics

A user has the ability to search for patient information leveraging two different approaches: by first name or last name in the BT-DB home page or by using the Database Search Engine (DSE) (Figure 4(a)). Using DSE, the user can investigate, with different filters, content across all fields included in the database to gather information toward answering clinical or technical questions. Different combinations of variables can also be inserted in the search fields by employing Boolean logic. When the researcher has retrieved data on patients, different plots can be visualized and statistics generated. Figure 4(b) to (d) show three different examples of these. In addition, all statistics can be exported as a *txt* file, in order to produce more accurate estimates using different statistical software packages.

Administrator menu

Users with administrator privileges can access the administrator menu. This tool enables the administrator to change or delete user names and passwords, add new users, assign different levels of access privileges, and create backups of all data stored in the database. The administrator generates the emergency password that enables the emergency access procedure. Moreover, the

Figure 2. (a) Patient's query screenshot. Different options are available to the user: (1) delete patient, (2) PDF document with all personal and clinical patient information, (3) file management, view, and modify, (4) personal information and date of death, (5) clinical presentation, (6) imaging clinical data (CT scan, Morphological Magnetic Resonance (MR) Imaging, Permeability MR, Perfusion MR, Spectroscopy MR, Diffusion Tensor Imaging, and MR Blood Oxygen Level Dependent), (7) medical surgery data, (8) Histological diagnosis data, and (9) medical therapy data; (b) options available for CT scan data (e.g. location and use of contrast enhancement); (c) options available for medical surgery (e.g. date of biopsy and date of total/partial tumor resection); and (d) options available for medical therapy (e.g. kind of therapy and kind of chemotherapy).

administrator is the only user with access to the log files recording monitored use of the database. These log files contain time and date of each operation, user name, and IP address of the accessing computer. It is a very important security safeguard for the database.

Tumor engine

An important instrument for physicians and researchers using BT-DB is the tumor engine tool. Here, it is possible to retrieve information about all tumors, by also viewing relevant articles on PubMed (http://www.ncbi.nlm.nih.gov/pubmed/). BT-DB is able to connect with remote servers to upload tumor information. These data can be retrieved for each tumor and the tumor definition ICD-O CODE and WHO 2010 classification included. ICD-O is a domain-specific extension of the International Statistical Classification of Diseases and Related Health Problems for tumor diseases, while the WHO classification is one component of a combination of criteria used in histological classification to predict outcome and response to therapy. Today, WHO grading is widely used, having incorporated or largely replaced other previously published grading systems.[11]

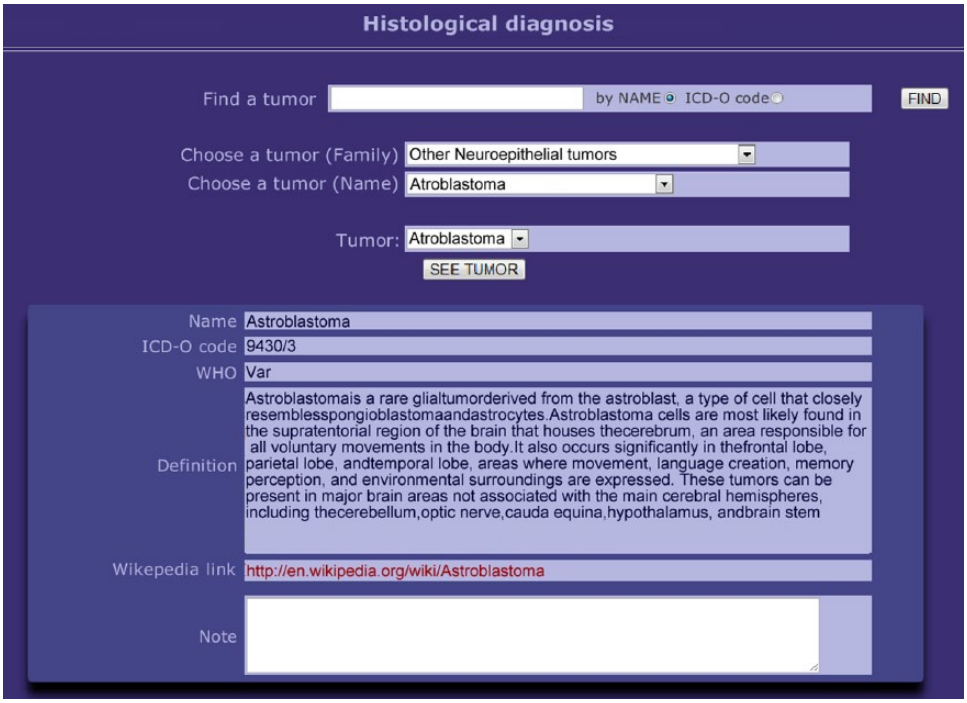

Figure 3. Screenshot of histological diagnosis. In this page, the user can look for specific tumor information. Name, ICD-O code, WHO classification, definition, and Wikipedia link for the selected tumor are generated from the database.

Conclusion

Today, databases are important tools in different fields, such as medicine, bioinformatics,[12] and neurosciences.[13] In this study, we describe the implementation of a database developed to collect and retrieve data on patients with BT. To build this database, we used MySQL, PHP, and HTML, and therefore, it can be easily accessed through any web browser. This web-based approach provides for potential use of the database by multiple institutions.

Since data stored in BT-DB contains PHI, we implemented a number of security safeguards controlling access to the database and organized different levels of access. Traceability of each access to the database is facilitated and should discourage unauthorized use. Being aware that these safeguards could impede access to the database in emergency situations, we also implemented an emergency access modality.

One of the problems encountered with databases is that they are useless if all the required data are not entered or if those data entered are not correct. Database designs should minimize the risk of data entry error. For this reason, in the BT-DB GUI, clinical and neuroradiological variables are, whenever possible, standardized and made available for selection with radio buttons or pulldown menus.

Since the field of diagnosis and treatment of BTs is evolving rapidly, the database is flexible in that the administrator can easily upgrade some main fields of the database. In addition,

Figure 4. (a) Screenshot of the database search engine. The user can retrieve information on patients by using name, surname, age, and date of birth or by using different combinations of variables that can be inserted into the search field, employing Boolean logic. BT-DB can show different statistics: for instance (b) patients' age statistics, (c) department of origin, and (d) date of patient hospitalization. Different information (such as mean, standard deviation, and median) and plots are generated.

complete customization of database structures is always possible, because BT-DB is a free and open source.

The search engine of the database enables retrieval of data stored in the database using a simple search GUI. The user can retrieve information across all fields included in the database to easily gather information and answer clinical or technical questions. Different combinations of variables can also be inserted into the search field by employing Boolean logic.

BT-DB can be used not only for clinical activity but also for research purposes. Different statistics (e.g. mean value, standard deviation, median, min, and max values) and different plots can be generated automatically from the database.

In conclusion, BT-DB is free and, as such, could become an important tool for clinical and research purposes helping physicians in their work. BT-DB GUI source code and manual are freely available to the general public and can be downloaded from http://tumorsdatabase.altervista.org.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

References

1. Andrew W and Walter MD. Brain tumors, http://kidshealth.org/parent/medical/brain/brn_tumors.html (2010).
2. Keime-Guibert F, Chinot O, Taillandier L, et al. Radiotherapy for glioblastoma in the elderly. *N Engl J Med* 2007; 356: 1527–1535.
3. Hottinger AF, Yoon H, DeAngelis LM, et al. Neurological outcome of long-term glioblastoma survivors. *J Neurooncol* 2009; 95: 301–305.
4. Bergamino M, Saitta L, Castellan L, et al. Cerebral aneurysms database, a new free database for intracranial aneurysms. *Neuroinformatics* 2011; 9: 427–430.
5. Sohn MJ and Byrne JV. A computerised relational database for auditing endovascular treatment of patients with intracranial aneurysms. *Neuroradiology* 2000; 42: 762–765.
6. Millar JS and Burke M. The UKNG database: a simple audit tool for interventional neuroradiology. *Neuroradiology* 2007; 49: 499–503.
7. Tahaghoghi MM and Williams H. *Learning MySQL*. 1st ed. Sebastopol (CA, USA): O'Reilly Media Incorporated, 2006, p.618.
8. Zawodny JD and Balling DJ. *High Performance MySQL Optimization, Backups, Replication, Load Balancing & More*. 1st ed. Sebastopol (CA, USA): O'Reilly Media Incorporated, 2004, p.304
9. Sklar D. *Learning PHP 5*. 1st ed. Sebastopol (CA, USA): O'Reilly Media Incorporated, 2004, p.370.
10. Shelly GB and Woods DM. HTML: *Complete Concepts and Techniques*. 5th ed. Boston (MA, USA): Course Technology Cengage Learning, 2004, p.384.
11. Louis DN, Ohgaki H, Wiestler OD and Cavanee WK. *WHO Classification of Tumours of the Central Nervous System*. 4th ed. Lyon: IARC, 2007, p.309.
12. Gallina AM, Bisignano P, Bergamino M, et al. PLI: a web-based tool for the comparison of protein-ligand interactions observed on PDB structures. *Bioinformatics* 2013; 29: 395–397.
13. Ascoli GA, Donohue DE and Halavi M. NeuroMorpho.Org: a central resource for neuronal morphologies. *J Neurosci* 2007; 27: 9247–9251.