

# How I Do it: A Practical Database Management System to Assist Clinical Research Teams with Data Collection, Organization, and Reporting

## Abstract

With the growing amount of research projects in the field of interventional radiology, storing and organizing the patients' data and information effectively is becoming more difficult. Existing hospital electronic medical record and archiving systems store patient information in the form of reports which put the physicians and researchers in a situation where they have to use time-consuming manual search through reports for suitable patients for different clinical researches. Spreadsheet programs such as Microsoft Excel® (Microsoft, Washington, USA) are often used as a data storage, which has limitations in both organization and the quality of the data. Higher risk of incorrect data entry and introducing duplicates are associated with data input and analysis without a database system. Furthermore, selection, as well as calculation of data, is time-consuming. This article aims to provide interventionists with an efficient data collection tool optimized for interventional oncology clinical research. The implementation of the current database and interface allows a much faster and more detailed retrospective analysis of patient cohorts. The database management system with an interface is a very time-efficient and robust tool that provides a significant edge over the manual retrieval of patient records by filtering data and assisting statistical analysis in a study-relevant fashion.

**Keywords:** Data collection, data organization, database, interventional radiology research

## Introduction

With the growing amount of research projects in the field of interventional radiology, storing and organizing the patients' data and information effectively is becoming more difficult. Existing hospital electronic medical record and archiving systems store patient information in the form of reports which put the physicians and researchers in a situation where they have to use time-consuming manual search through reports for suitable patients for different clinical researches.

Spreadsheet programs such as Microsoft Excel® (Microsoft, Washington, USA) are often used as a data storage place to save and organize patient data for research. Once the spreadsheet is completed filled up, it is manually filtered by preset study's parameters and then sent to statistical analysis software for further analysis.

For statistical analysis, columns containing text are translated into binary values (1 or 0) to be in a format acceptable by statistical analysis software. For example, each underlying liver disease is assigned

a new column. Patient clinical, laboratory and pathological data are read manually to assign a 1 or 0 to each underlying liver disease entity column, 1 for positive, 0 for negative. Under underlying liver disease column, researchers would write a 1 for all patients with positive data 0 for all patients with negative results.<sup>[1]</sup>

This method of data storage has limitations in both organization and the quality of the data. Higher risk of incorrect data entry and introducing duplicates are associated with data input and analysis without a database system. Furthermore, selection, as well as calculation of data, is time-consuming.

The purpose of this article is to provide an improved workflow efficient tool optimized for interventional oncology clinical research.

## Database and Query Interface Design

The ideal database management system should have two distinct parts, the database server and client interface, illustrated in Figure 1.<sup>[1]</sup>

**Amr Maged  
Elsaadany,  
Fida Hasan Alaeddin<sup>1</sup>,  
Mohamed Farouk  
Badran<sup>1</sup>,  
Hamad Saleh  
Alsuhaibani<sup>1</sup>**

*Department of Radiology, King Faisal Specialist Hospital and Research Center, Body Interventional Unit, <sup>1</sup>King Faisal Specialist Hospital and Research Center, Body Interventional Unit, Riyadh, Kingdom of Saudi Arabia*

### Address for correspondence:

*Dr. Amr Maged Elsaadany,  
Department of Radiology,  
King Faisal Specialist  
Hospital and Research Center,  
Body Interventional Unit,  
P.O. Box 3354, Riyadh 11211,  
Kingdom of Saudi Arabia.  
E-mail: aelsaadani@kfshrc.edu.sa*

### Access this article online

Website: [www.arabjir.com](http://www.arabjir.com)

### Quick Response Code:



**How to cite this article:** Elsaadany AM, Alaeddin FH, Badran MF, Alsuhaibani HS. How i do it: A practical database management system to assist clinical research teams with data collection, organization, and reporting. Arab J Intervent Radiol 2017;1:10-4.

This is an open access article distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 License, which allows others to remix, tweak, and build upon the work non-commercially, as long as the author is credited and the new creations are licensed under the identical terms.

For reprints contact: [reprints@medknow.com](mailto:reprints@medknow.com)

The illustrated database is run by software (MySQL, Oracle Corporation, California, USA) on a central computer server within the department.<sup>[2,3]</sup>

In the interventional oncology section in our institution, which is a large tertiary care center till now the database management system units are connected to each other via the hospital intranet network through a shared folder rather than a database server. Granted access to client graphical user interface (GUI) is provided to authorized users which enables them to add, edit, and query data concurrently utilizing Filemaker Pro Advanced (FileMaker Inc., a subsidiary of Apple Inc., USA). It is all about easy access to information, but that is a double-edged sword. If you do not add security, every person who uses your database has unrestricted access and can add data, tables, and scripts as freely as you can. Of course, you can take all the usual precautions (give your computer a password, install virus-protection software, lock your office door, and so on). However, the minute you let anyone else into the database, you have all kinds of security challenges. This database management security system has two primary levels of control: Who can get into your database in the first place; and what they can do once they are there. You determine who gets access to your database by setting up user accounts, and you control what each person can do by assigning privilege sets to each account. Any updates in the data are immediately available for others to see. The database performed automatic calculations using queries, user-defined search criteria. Queries are saved, rerun, and exported to spreadsheets. Queries aid in data analysis and increase study productivity.<sup>[4]</sup> These tools can be very useful in filtering and sorting the data sets.

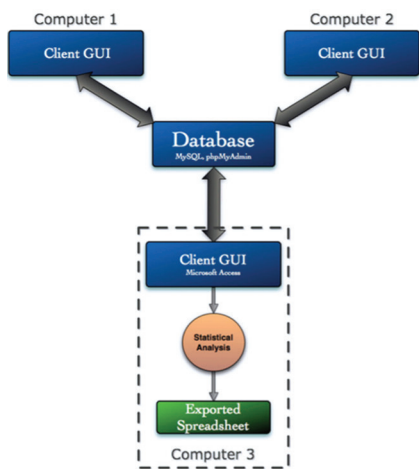


Figure 1: The data flow chart this diagram demonstrates a general design of the database server and its users. Clinical users can use the database client graphical user interface to import data easily

Figure 2 illustrates the query interface and an example of a request from the database.

### Graphical User Interface Design and Utility

In our interventional oncology research environment, the database GUI was populated to make patient data entry easier [Figure 3].

This was accomplished by creating and customizing user-friendly interface forms that contain textboxes and labels including biographic data, laboratory, clinical and imaging information (e.g., number, size of the lesions, underlying liver disease,... etc.), tumors types, treatment options (transarterial radioembolization, trans-arterial chemoembolization, radiofrequency ablation, or microwave ablation), etc. The GUI is accustomed to view patient data and allows users to add/edit data [Figure 4].

In addition, different photos, images, and even videos can be added to each patient’s data entry [Figure 5].

### Automatic Calculations

Automatic calculations can be used between different events, such as dates. For example, the database may calculate the time between the trans-arterial treatment and the transplant date or patient’s date of death after a particular treatment (e.g., radioembolization), which is essential for survival studies. The database does also automatically calculate clinical scores such as Child-Pugh score and Barcelona Clinic Liver Cancer (BCLC) stage as shown in Figure 6.<sup>[5]</sup>

### Statistical Output

The ability to provide the first tier of statistical information is another powerful feature of the database. The user defines the search criteria and runs queries to obtain immediate statistical information about a particular set of parameters. With this feature, the database can easily output an accurate summary of patient data such as, how many patients per each BCLC stage [Figure 7].

### Discussion

The aim of this article is to present a database management system that we believe is user’s friendly, very time efficient and accurate in feeding, storing, retrieving, and analyzing patients’ data for any clinical research projects. This is achieved through the availability of automatic calculations, interface forms, queries, etc. The way of customized interface, data access, entry, organization, automatic calculations, and export processes are clearly performed to assist the researchers with data and statistical analysis.

::Number_Of_Le...	::Largest	Aim of Therapy	::Treatment_Opti...	::STAUS	::Transplantation...	BCLC::Stage
Multiple		Down-staging	Segmental	2 Transplanted	1 Living Related	Stage B

Figure 2: The query interface. In this example query, a list of the number of patients with hepatocellular carcinoma who underwent transplantation after down-staging them with transarterial hepatic radioembolization

With the presented database management system put into use, the effort for clinical studies can truly focus on conducting a variety of statistical analysis and data interpretation rather than preparing data for analysis.<sup>[6]</sup> All retrospective and prospective data can be merged into this database, enabling a centrally maintained and shared resource.

A typical Excel spreadsheet for some research studies can have more than 100 columns, which make it very difficult to track medical data due to a large number of columns in the spreadsheet. Using a spreadsheet in data collection and analysis, converting data to binary/numeric format is very time-consuming and impractical. In contrast to the typical Excel spreadsheet with so many columns, the database management system presented in this article

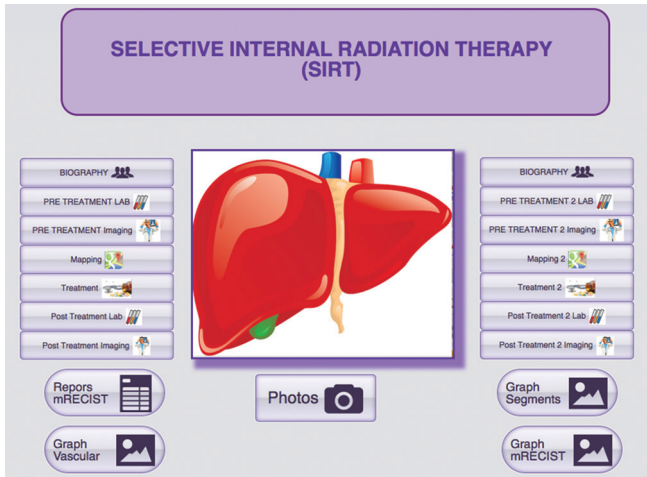


Figure 3: The front page of graphical user interface showing different buttons that facilitate accessing different layouts

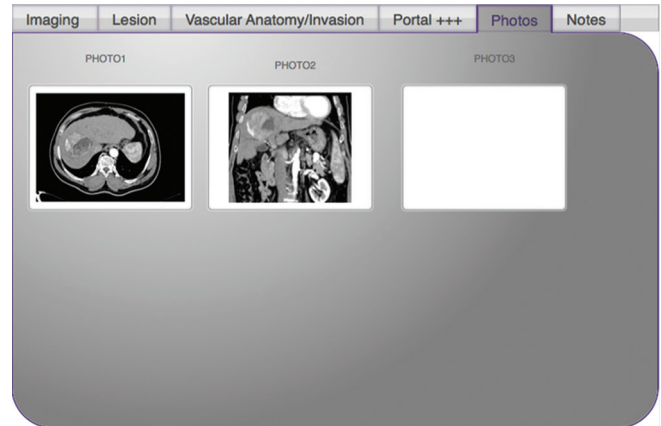


Figure 5: The layout for storing the imaging of different modalities for each patient

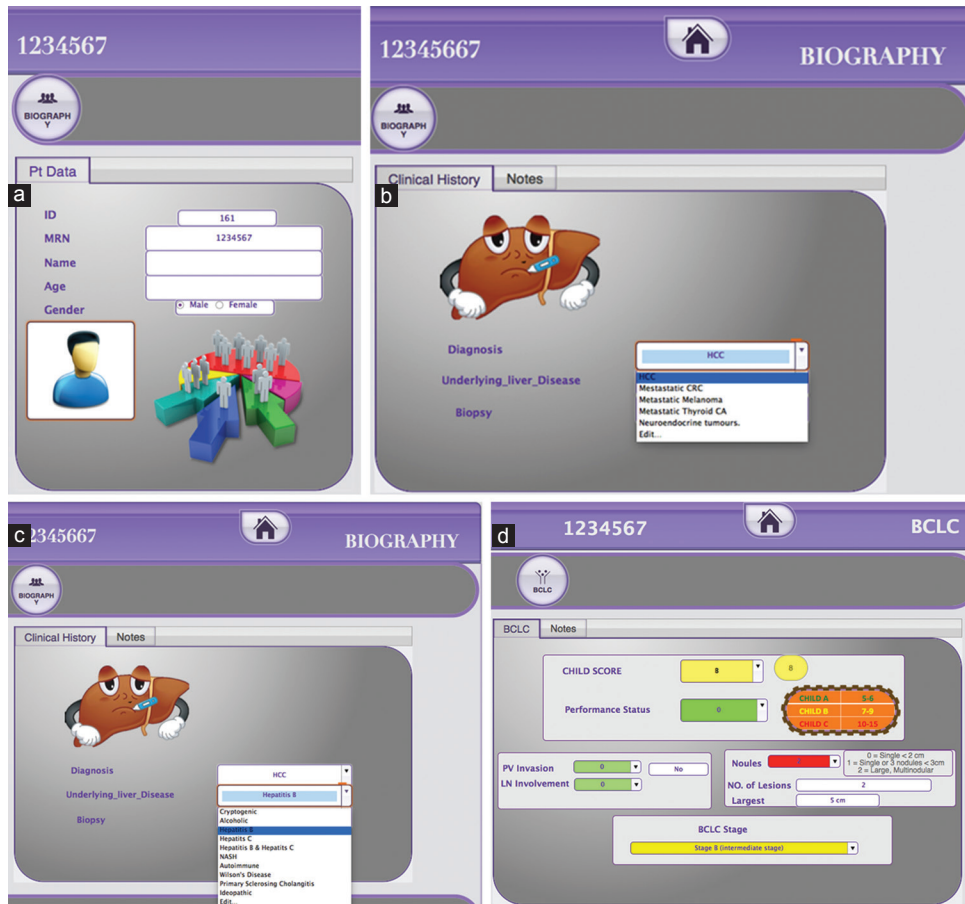


Figure 4: The users can store data into the different database store boxes. (a-c) Showed different store boxes regarding patient demographic data. To minimize human error factor, drop menu lists were used to enter the data. (d) Showed pretreatment laboratory data essential to calculate the Child-Paugh score

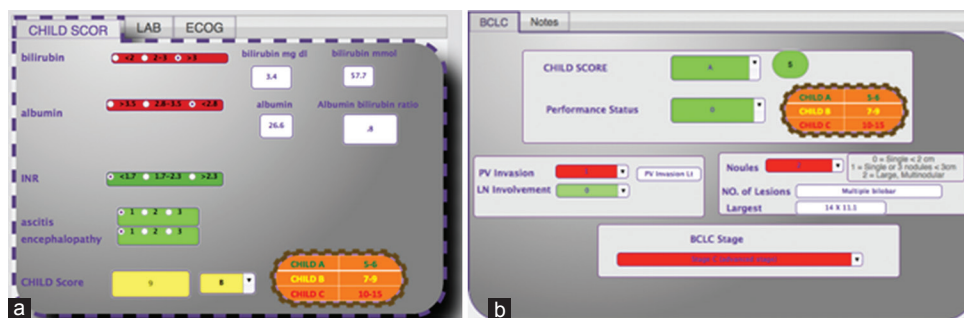


Figure 6: A patient's Child-Pugh score (a) and Barcelona Clinic Liver Cancer stage (b). They are automatically calculated when provided with pertinent patient data

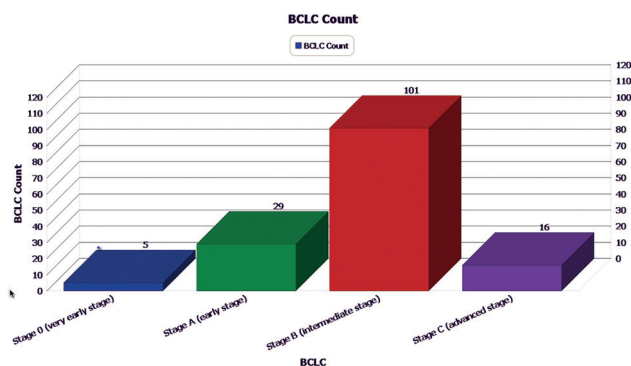


Figure 7: The automatically created bar graph regarding the number of patient of each Barcelona Clinic Liver Cancer stage

alleviates the inconvenience of manually searching, organizing, and calculating data, especially more complex ones. In addition, browsing and adding prospective and retrospective data through the database GUI presented here is more organized and practical. Combining the database's ability to calculate statistical analysis with automatic calculation queries, reports can be generated with virtually any parameter. This is not only helpful in radiology, but it is also beneficial for other studies and hospital information systems.

Such systems need to be tested thoroughly before being used. It is highly recommended to have a special server to host the database. Training is crucial for success. Users have to be trained fully by experts on the system and to have access to them after implementation to solve any unexpected problem or answer practical questions as well as develop and maintain the system as required.<sup>[7,8]</sup> The maintenance challenges include backups, modifying the database structure to serve different and developing data and to keep the database and software up to date. The server host can be on PC or online. Either way should be accessible for users in local or global network. It is important to have intercommunication between different database systems.

The learning curve is usually steep at the setup. The database allows an easy and quick way to feed the data in systematic way, searching results with statistics and calculations and keeping data while enabling multi-user

access. In day-to-day practice, each departmental section and research unit should have their own separate database. For example, while interventional oncology unit has its own research database other teams in an intervention like critical limb ischemia and hepatobiliary have their own. However, these separate databases should be interconnected for cross-departmental research to provide a wider range of data and research.<sup>[9]</sup>

## Conclusion

The implementation of the current database and interface allows a much faster and more detailed retrospective analysis of patient cohorts. In addition, it facilitates data management and a standardized information output for ongoing prospective clinical trials. The database management system with an interface is a very time-efficient and robust tool that provides a significant edge over the manual retrieval of patient records by filtering data and assisting statistical analysis in a study-relevant fashion.

## Financial support and sponsorship

Nil.

## Conflicts of interest

There are no conflicts of interest.

## References

1. Lee H, Chapiro J, Scherthaner R, Duran R, Wang Z, Gorodetski B, et al. How I do it: A practical database management system to assist clinical research teams with data collection, organization, and reporting. *Acad Radiol* 2015;22:527-33.
2. Stobart S, Vassileiou M. *MySQL Database and PHP My Admin Installation. PHP and MySQL Manual*. London: Springer; 2004. p. 461-73.
3. Kuenz D. *Manage Data for Free with MySQL*. Book Manage Data for Free with MySQL. (City: Element K Journals); 2001. p. 7-10.
4. Coronel CM, Rob P. *Database Systems: Design, Implementation, and Management*. 9<sup>th</sup> ed. Boston, Massachusetts: Cengage Learning; 2009.
5. Llovet JM, Di Bisceglie AM, Bruix J, Kramer BS, Lencioni R, Zhu AX, et al. Design and endpoints of clinical trials in hepatocellular carcinoma. *J Natl Cancer Inst* 2008;100:698-711.
6. Kanas G, Morimoto L, Mowat F, O'Malley C, Fryzek J,

- Nordyke R. Use of electronic medical records in oncology outcomes research. *Clinicoecon Outcomes Res* 2010;2:1-14.
7. Xie SX, Baek Y, Grossman M, Arnold SE, Karlawish J, Siderowf A, *et al.* Building an integrated neurodegenerative disease database at an academic health center. *Alzheimers Dement* 2011;7:e84-93.
  8. Parkes D, Lowman M, Andres C, *et al.* Pro Python System Administration. Automatic MySQL Database Performance Tuning. Apress; 2010. p. 329-48.
  9. Piriyaongsa J, Bootchai C, Ngamphiw C, Tongsimma S. MicroPIR: An integrated database of microRNA target sites within human promoter sequences. *PLoS One* 2012;7:e33888.