BUILDING A DATABASE FOR PEKING UNION MEDICAL COLLEGE HOSPITAL OUTPATIENTS WITH RHEUMATIC DISEASES

BY

Xinping Tian M.D., Ph.D

A MASTER'S THESIS

Presented to the Department of Medical Informatics and Outcome Research and

the Oregon Health & Sciences University School of Medicine

in partial fulfillment of the requirements for the degree of

Master of Science

May 2003

TABLE OF CONTENTS

ACKNOWLEGEDMENTS ..... **I**

ABSTRACT ..... **II**

INTRODUCTION **1**

DATABASES IN RHEUMATOLOGY **1**

USABILITY TESTS IN MEDICAL INFORMATICS **2**

THE NECESSITY OF A DATABASE FOR THE OUTPATIENT CLINIC OF RHEUMATOLOGY OF PEKING UNION MEDICAL COLLEGE HOSPITAL **4**

GOALS OF PROJECT **6**

METHODOLOGY **7**

DATABASE DESIGN AND DEVELOPMENT **7**

USABILITY TESTING **15**

RESULTS **18**

THE DATABASE **18**

FORMS *18*

QUERIES *22*

REPORTS *23*

USABILITY TEST AND DATABASE MODIFICATION **23**

SUBJECTS *23*

TEST RESULTS *23*

INTERFACE MODIFICATION *29*

DISCUSSION **35**

DATABASE AND USABILITY TEST **35**

FUTURE WORK **37**

CONCLUSION **38**

REFERENCES **39**

APPENDICIES **41**

APPENDIX A **42**

TABLES OF THE DATABASE *42*

APPENDIX B **50**

THE RELATIONSHIP OF DATABASE TABLES AND FIELDS *50*

APPENDIX C **51**

FIGURES IN THIS DATABASE *51*

APPENDIX D **69**

QUERIES DEVELOPED IN THIS DATABASE *69*

ACKNOWLEGEDMENTS

My sincere thanks to my thesis committee members:

Dr. William Hersh, Advisor

Dr. Judith Logan

Dr. Robert Bennett

Without their wisdom, guidance, generous help and strong support, I could not accomplish this thesis.

I would like to thank my mentors in China, Dr. Gerald Lazarus and Dr. Audrey Jakubowski, for their strong support. It is through their and Dr. Hersh's efforts that made my study possible. I would also express my special thanks to Dr. Roy Schwarz, the Chair of China Medical Board, for his generous help. My sincere thanks to the leadership and friends of Peking Union Medical College Hospital for their help and support. I would also like to express my gratitude to my friends at OHSU who supported and helped me. Finally, I would like to give very special dedications to my husband and son for their strong support and confidence in completing this thesis.

ABSTRACT

Although several databases have been developed in rheumatology aimed at profiling the morbidity pattern of rheumatic diseases of the country, or to capture the detailed clinical and outcome information of patients with a specific rheumatic disease, we know of no database, as yet, for capturing visit-related health information of all outpatients with rheumatic diseases. In this thesis, a database was developed for rheumatic outpatient clinic of Peking Union Medical College Hospital (PUMCH). The primary goal of this project was to build a database that can collect and organize visit-related information of outpatients with rheumatic diseases to meet information needs of clinical research, health care continuity and teaching of rheumatic diseases. After the database was completed, a usability test was carried out to evaluate its functionality as well as the intuitiveness of the database interfaces. The results were analyzed and used for further modification.

This project consisted of two parts. In the first part, a relational database was built using Microsoft Access 2000. Patients' demographic data as well as visit-related information could be collected. Information retrieval was supported by a series of queries of the database. The database also provided a report print service for patients and physicians. A search function for identifying patients was also enabled. An overview interface was developed to provide quick access for physicians to review the overall medical information of a patient.

In the second part of this project, a "simplified thinking aloud" method was adopted to conduct a usability test. Five medical informatics graduate students were selected to evaluate the database and its interfaces. The results of the "think aloud" showed that the

database could accomplish all the expected functions. The interfaces were generally intuitive and easy to use and the general layout was good and made the data entry flow easy to follow. Queries were useful, helpful and flexible. The "Print Report" function was very easy and the reports were well labeled. They all thought that the "searching for patient ID" function was simple and very easy. Usability problems found in the test could be grouped into format problems, consistency problems, layout problems and other problems. The database was then modified based on the problems and the suggestions derived from the test.

In summary, a database was built for outpatients with rheumatic diseases for PUMCH that meets information needs of clinical research, health care continuity and teaching. The database interfaces are generally intuitive and easy to use.

INTRODUCTION

Rheumatic diseases are among the most prevalent chronic medical conditions in the world and may be one of the leading causes of disability. The consequences of rheumatic diseases have a significant impact on individuals with the disease, their families and society (1), but these severe consequences traditionally have been underestimated by physicians, reimbursement agencies, research funding programs and the general public (2). Physicians who treat these patients are concerned about the outcomes of rheumatic diseases and treated by them. They are not only interested in function, work ability, mortality and the response to treatment, but also in those disease factors that can be measured by radiological examinations and laboratory studies. Since these outcomes are chronic, longitudinal clinical research is needed. Therefore, accurate and complete medical data collection is of critical importance. Data collection is not only important for clinical research, it is also critical for delivery of high quality health care, particularly for ensuring continuity of health care in chronic disorders. Accurate clinical data collection and organization can also benefit medical teaching. However, how to collect and organize the data occurring in the clinic where most patients are being evaluated and receiving care is a big challenge. It is generally accepted that databases can offer the basic schema to help organize data and make it useful in meeting information needs of health care and clinical research. Databases that can capture visit-related medical data may meet these information needs well.

DATABASES IN RHEUMATOLOGY

Several rheumatology databases had been developed and reported in the literature. They are all relational databases and some were developed for longitudinal clinical

research. Generally they can be divided into two categories based on the primary goal: those which serve as rheumatic disease patient registries and those used for specific studies of individual rheumatic disease. None of them is designed to collect visit-related health care information of all outpatients with rheumatic diseases, however.

Registry databases have been developed in the USA, United Kingdom, Canada, Germany and Netherlands (3-9). The primary purposes of registry databases are to determine the morbidity profile of rheumatic diseases for clinical epidemiology studies. They can also be used for prevention or treatment of rheumatic diseases. Finally, they can be used for monitoring the changing patterns of rheumatic diseases and medical care (9).

Databases for individual rheumatic diseases have been developed to capture the detailed clinical and outcome information of patients with a particular rheumatic disease. Almost all of them were developed for longitudinal clinical research. For example, there are databases for Rheumatoid Arthritis, Systemic Lupus Erythematosus and Systemic Connective Tissue Disorders, as well as databases for Fibromyalgia and Back Pain (10-13). The purpose of these single databases is to enable physicians to collect patients' data for detailed clinical research.

USABILITY TESTS IN MEDICAL INFORMATICS

Usability, by definition, is "the effectiveness, efficiency and satisfaction with which specified users achieve specific goals in a particular environment" (14). Since the advent of usability engineering in the 1980s, usability testing has been widely acknowledged as a fundamental technique for evaluating user performance and acceptance of products and systems (15). Usability testing refers to the evaluation of information systems through

the analysis of typical end users interacting with the system (16). The "discount usability engineering" is used to describe the application of low cost methods for conducting usability tests. The "discount usability engineering" method is based on the use of scenarios, simplified thinking aloud and heuristic evaluation techniques (17). The basic strategy of the "thinking aloud" usability testing is to videotape all human-computer interaction (i.e., computer screens) and audiotape all subject verbalizations when they interact with the system prototypes. "Simplified thinking aloud" usability testing cuts the cost of usability testing down further by analyzing data based on the notes taken by the experimenter instead of by videotapes of the screens (17). Based on data collected from a representative sample of subjects, typically involving as few as three to five participants per study, the majority of usability problems and issues can be identified and summarized (17). The audio recording can then be analyzed using methods involving the coding and classification of user problems. The information which must be collected in the usability test includes: 1) suggestions by users for improvements to both the user interface and system functionality, 2) identification of usability problems such as lack of consistency in interfaces and operations, and 3) quantitative measurements including time to task completion and system response time (18). The resulting information from such evaluation is summarized and presented to the designers, allowing for iterative modification of the system.

In health care, increasing numbers of researchers are applying usability engineering methods for evaluation of information systems. A range of methodologies has been developed. Recent work has included the application of remote usability testing for distance evaluation of web-based health care information systems (16). It has been

reported that full usability testing can be efficiently inserted into the design cycle of patient electronic medical record systems, with a single initial design-evaluation-redesign cycle leading to as much as a ten-fold reduction in usability problems (19). As information technology application in health care has become more and more complex and the demands of particular user needs increases, usability engineering methods are likely to become more important.

THE NECESSITY OF A DATABASE FOR THE OUTPATIENT CLINIC OF RHEUMATOLOGY OF PEKING UNION MEDICAL COLLEGE HOSPITAL

Peking Union Medical College Hospital (PUMCH), which was founded by the Rockefeller Foundation in 1921, is one of the largest and most respected academic medical centers in China. It is the only hospital that reports directly to the Ministry of Health of China. It is also a national referral center and has more than 1 million patient-visits a year. About half of its patients are from areas outside of Beijing. The Department of Rheumatology and Clinical Immunology is the clinical, research and teaching center for rheumatic diseases in China. It has more than 40,000 outpatient visits every year. The majority of these patients are from areas outside of Beijing. However, due to the current medical record management system in the outpatient clinic, the majority of outpatients do not have a full medical record. Most outpatients have only a small notebook as the record of their visits. Patients themselves are supposed to keep their own "small notebook" medical records. So the rate of loss for this type of medical record is high. When medical records are lost, physicians' time is wasted by collecting all information again. In addition, because there is no Patient ID number that can be used to identify patients, it is almost impossible for physicians to review this kind of medical record. This threatens the ability of clinicians to carry out high quality clinical research.

So, this medical record management model not only compromises the continuity of health care, especially for patients with chronic conditions such as the rheumatic diseases, but also compromises the quality of health care delivery in the hospital. Therefore, it is necessary to find a tool for managing and organizing outpatients' medical records. Databases are a potential way of collecting and organizing visit-based health care information.

I am a rheumatologist, so it is quite natural for me to choose rheumatology as the target field for such a database. Another advantage for choosing rheumatology is that I understand the basic information needs of rheumatologists of PUMCH. Moreover, because of the chronic nature of rheumatic diseases, collecting and organizing patients' longitudinal medical information is very important for clinical research and high quality health care delivery. Therefore, developing a database that can capture and organize medical information in the rheumatology clinic was chosen to be my thesis project.

The specific objectives of this database are to enable rheumatologists of PUMCH to 1) get insight into the morbidity patterns of rheumatic diseases as seen at PUMCH; 2) have more comprehensive data for clinical research; 3) find clues for clinical epidemiology studies; 4) compare the differences between practices; 5) understand in depth the correlation between disease patterns and laboratory tests or radiological examinations; 6) retrieve better evidence about patients' responses to medications; 7) review patients' overall medical information; 8) access better cases for teaching.

GOALS OF PROJECT

There were two goals of this project:

1. Build a database that can collect and organize visit-related medical information of patients with rheumatic diseases from PUMCH. The information in the database should be able to meet the needs of clinical research, improve health care delivery and support teaching of rheumatic diseases,
2. Test the functionality and performance of the database and the intuitiveness of its interfaces by a usability test. The "simplified thinking aloud" method is chosen to carry out this test. The results of the usability test will be analyzed and used for database modification in order to make it functionally better and more user-friendly.

METHODOLOGY

DATABASE DESIGN AND DEVELOPMENT

1. Basic schema of the database

The database is named "PUMCH Rheumatic Clinic Database". Since the relational databases are still the most widely used database in medicine, this database is also designed as a relational database, similar to other databases in rheumatology reported in the literature.

In a relational database, tables form the fundamental structure. 14 tables were developed to collect and organize patients' medical information. The organization and composition of tables in the database are shown in Table 1. Among these tables, five are "lookup" tables: Diagnosis, LabTest, Drug, RadiologyTest and Procedure. These "lookup" tables contain no visit-related data, but have primary keys that form important data relationship between them and other data tables. All other tables are data tables of the database. They are used to store visit-related medical information of a specific visit and the general information of the patient.

The relationship between these tables and data fields is shown in Figure 1. It is basically a star schema. The right five tables are the lookup tables. The "Visit" table forms a bridge, which links the "Patient" table, containing general information of a patient with the visit-related data of that patient.

Table 1. Tables and fields of the database

Patient	**Visit**	**VisitDiagnosis**	**Diagnosis**
PatientID(KEY)	VisitID(KEY,Autonumber)	VisitDiagnosisID(KEY,Autonumber)	DiagnosisID(KEY,Autonumber)
Name	Date of Visit	VisitID	DiagnosisName
Gender	PatientID	DiagnosisID	SNOMEDCode
Age			
Date of Birth			
Occupation		**VisitLabTest**	**LabTest**
Marital Status		LabTestDetailID(KEY,Autonumber)	LabTestID(KEY,Autonumber)
Educational Level		LabTestID	LabTestName
Race		VisitID	
Address		Result	
Postcode			
Home Phone Number			
Contact Name		**VisitDrug**	**Drug**
Contact Phone Number		DrugDetailID(KEY,Autonumber)	DrugID(KEY,Autonumber)
Date of First Visit		VisitID	DrugName
Date of First Symptom		DrugID	
		Dosage	
		StartDate	
		StopDate	
		VisitRadio	**RadiologyTest**
		RadioDetailID(KEY,Autonumber)	RadiologyTestID(KEY,Autonumber)
		VisitID	ExamName
		RadioligyTestID	
		Result	
		VisitProcedure	**Procedure**
		PatientProcedureID(KEY,Autonumber)	ProcedureID(KEY,Autonumber)
		VisitID	ProcedureName
		ProcedureID	
		Result	
		VisitSerum	
		SerumID(KEY)	
		VisitID	
		VisitDNA	
		DNAID(KEY)	
		VisitID	
		Result	

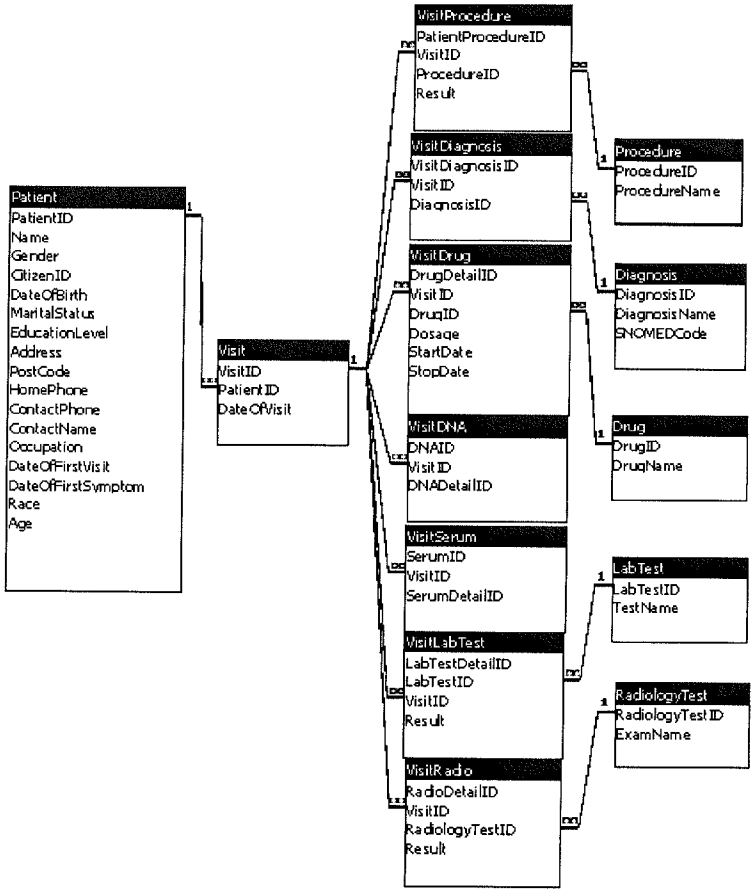


Figure 1. The relationship of tables and fields of the database

2. End user of the database

Since more than 40,000 outpatients visit the rheumatic disease clinic of PUMCH, each rheumatologist needs to see at least 25 patients in each half-day clinic. Therefore, it is difficult for them to enter patients' medical data into the database while seeing patients. For this reason, data entry clerks will initially be used for data entry.

Since queries are built for clinical research purposes, physicians will be the end-users of query functions of the database. Although the reports are designed to meet the requirements of physicians and patients who want to be informed about their own medical information, again, data entry clerks will be the intermediaries in this need. So

there are in fact two kinds of end-users of the database: data-entry clerks and the rheumatologists of PUMCH.

3. Developmental tool

Since this is the first clinical database in the PUMCH outpatient clinic, it should be inexpensive, easy to use and modify according to user feedback. Therefore, Microsoft Access 2000, which is widely available and inexpensive, is chosen to design the database.

4. Functions the database can accomplish

In order to achieve the objectives of the database, five functions can be accomplished by this database: data entry, information retrieval, report printing, searching for a patient record and reviewing the overall medical information of a patient.

5. Data collection flow (Figure 2)

When a new patient comes to the outpatient clinic, a unique Patient ID number is assigned and a Patient ID card labeled with the patient's name and Patient ID is given. Patient ID is the "key" for identifying individual patient. In the meantime, general information is collected and a medical record folder labeled with his or her Patient ID, name and gender on is given to the patient. When the patient sees the physician, he or she shows this folder to the physician. The physician writes detailed information about the patient's visit in the folder. When the patient leaves the clinic after the visit, the medical record folder is collected and sent to data entry clerks for data entry.

Return patients come to the clinic with their Patient ID card, and their medical record folders are sent to the physicians they are going to see. They then follow the same process as the new patient. It is unnecessary to enter the general information for return patients unless there is a change. However, information related to that particular visit does need to be entered.

The implementation of this database is integrated into the regular outpatient flow without creating extra work for both patients and physicians. The data collection flow is the same as a regular outpatient visit except patient ID number assignment and general data collection for new patients. For return patients, they follow the same process as there is no such a database.

After the data are entered by data entry clerks, another individual (the administrator of the database) checks the accuracy of the data and performs back up.

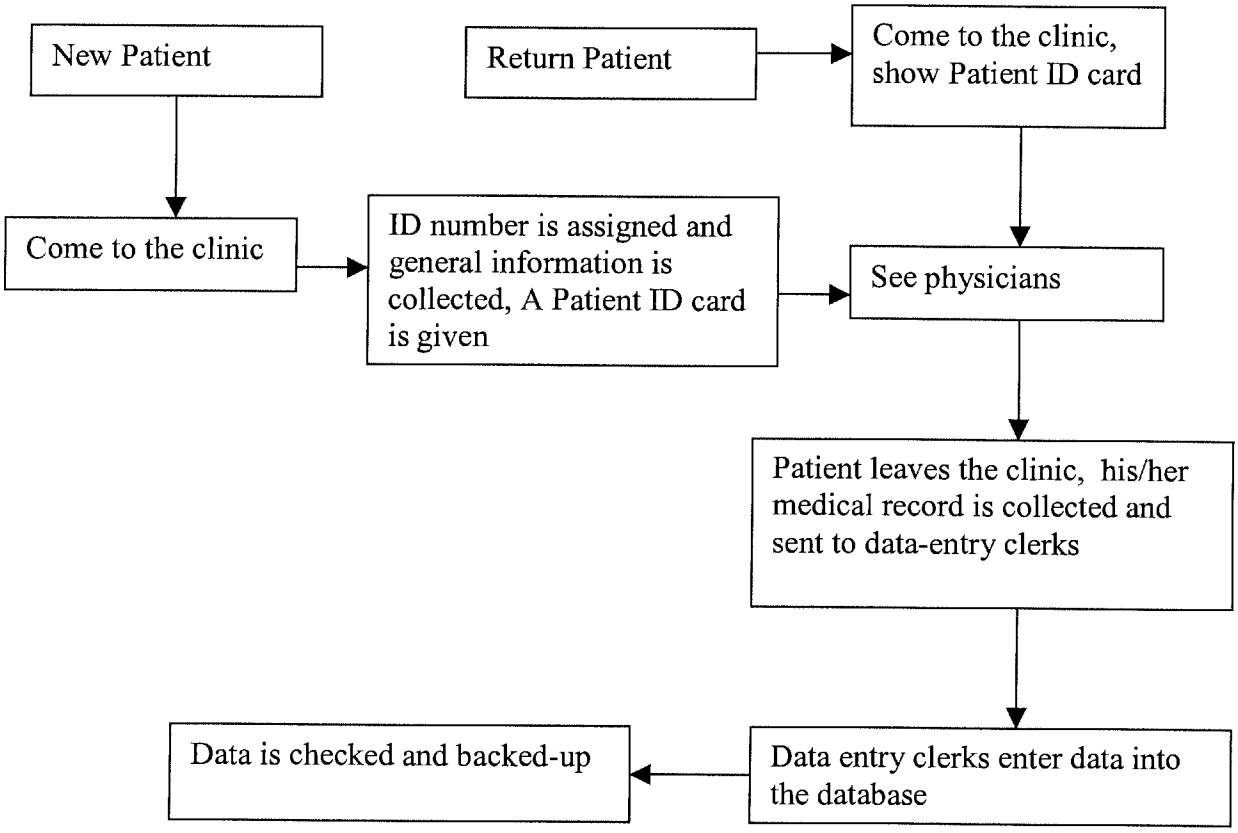


Figure 2. Data collection flow of patient visit (for all outpatients with rheumatic diseases)

6. Data collected in the database

To meet the objectives of this database, the following two parts of data are collected:

Part 1: General information about a patient. This includes:

a. Demographic data: name, gender, age, date of birth, race (ethnicity), occupation, marital status, education level, citizen ID (which was issued when an individual was born and is unique for each individual who is a citizen of the People's Republic of China. Since this number will never change in one's life, it can also be used to identify individual patients).

b. Contact information including: home phone number (if applicable), a contact name, contact phone number (if applicable).

c. Other information related to patient's illness and visit including: Patient ID (assigned at the first visit), date of first visit and date of first symptom.

Part 2: Visit and visit-related medical information. This is the core information of the database. It includes:

a. Date of visit.

b. Diagnoses. The diagnoses made by physicians at each visit. Diagnoses are limited to the 172 common rheumatic diagnoses as shown in Appendix A. The SNOMED Code for each concept ID is included for future indexing.

c. Laboratory tests and results. The 39 most commonly ordered biochemical laboratory tests and 33 antibodies (most are auto-antibodies) are included. The items are listed in Appendix A.

d. Radiological examinations and results. The 17 most commonly ordered radiological examinations in rheumatology are included (Appendix A).

e. Procedures. Ten procedures that are usually done in the outpatient clinic are included in the database (Appendix A).

f. Medications. This records the medications prescribed by physicians at the date of visit with the dosage, start date (date of prescription) and stop date (date of withdrawal) information also collected. The 38 most commonly prescribed immunosuppressants, Non-steroid anti-inflammatory drugs (NSAIDS) and other medications are included (Appendix A).

g. To meet the information needs of clinical research, samples of patients' serum or DNA are also collected and stored in the Serum and DNA Banks of the rheumatology department. The patient's Serum Bank ID, DNA Bank ID and date of sample drawn are included. Some patients can have more than one serum or DNA samples banked, but there is only one Serum or DNA Bank ID number. This ID is then used to identify samples.

7. Queries of the database:

One of the major goals of this database is to provide information for clinical research which involve specific queries. A series of queries were built into the database to make the querying task easy. The rheumatologists can select and run a query to get the desired information. All of these queries built are based on frequently asked questions by rheumatologists. Queries built in this database can be used to search on one medical domain (such as a single laboratory test, diagnosis or medication) and/or more than one medical domains. In order to verify the correctness of the queries, medical records of 20 patients with more than 80 visits were entered into the database. The output of every query was checked for its accuracy by reviewing these records.

8. Getting reports against the database

Since some patients may want summary information about their illness, and local referral physicians also want to know the information of the patient's visit in PUMCH, a series of reports that summarizes a patient's diagnoses, medications, laboratory tests, radiological examinations and procedures were built. A report that summarizes information of a specific visit was also built. This facilitates physician communication and also improves the continuity of health care. Following reports are developed in this database:

- Summary of a visit--Summarizes all information of a specific visit
- Summary of diagnoses--Summarizes all diagnoses of the patient has
- Summary of medications--Summarizes all medications of the patient
- Summary of laboratory tests--Summarizes all laboratory tests of the patient
- Summary of radiological examinations--Summarizes all the radiological examinations of the patient
- Summary of procedures--Summarizes all the procedures of the patient
- Summary of DNA Bank information--Summarizes all patients who had DNA sample(s) in the DNA Bank
- Summary of Serum Bank Information--Summarizes all patients who had serum sample(s) in the Serum Bank

9. Look up patients based on their citizen ID

A frequently encountered situation in the outpatient clinic is that a patient will lose his or her Patient ID card (Patients are asked to show this card when they come to the hospital after the first visit). For this case, there must be a way to recover the Patient ID. This database offers another function--searching for a patient ID using the patient's citizen ID.

USABILITY TESTING

After the database was completed, a usability test was carried out to evaluate the efficiency and functionality of the database, and the intuitiveness of its interfaces. A "simplified thinking aloud" usability test method was adopted by this study. Five subjects were recruited by e-mail from all on-campus medical informatics master program students. The first five respondents were chosen. Informed consent was obtained following procedures approved by the OHSU Institutional Review Board.

Before the beginning of the test, a scenario was read that describes what the subject was supposed to do and what they should pay attention to in entering data into the database. If the subjects became stuck on any task, they were instructed to move on to next task. Each subject was asked to do four jobs during the test:

1. Enter patient data. Each subject was asked to enter the complete medical information (general information and visit-related information) of two new patients and visit-related information only of two return patients. All subjects used the same medical record printouts. All of these medical records were from PUMCH and were translated into English.
2. Query the database. Each subject was asked to run a single query, chosen to present some challenge to the subjects. Specifically, each subject was asked to find the information of patients who had a diagnosis of "Systemic Lupus Erythematosus" and also had an abnormal ALT laboratory test value with the cutoff value of 30 IU/dl.

3. Print a report. Each subject was asked to print a summary of all the laboratory tests of a patient. A list of Patient IDs was provided so that the subject could choose any patient from it.
4. Searching for patients in the database. A list of patients, citizen IDs was offered to each subject. The subject was then asked to choose one from the list and find the corresponding patient ID. The output of this part is the patient's ID number, name, age and gender. This information can then be used to assure that the right patient is identified.

The experimenter did not demonstrate any of these tasks before the test but let the subjects perform all the jobs themselves. The experimenter just opened the database and showed the main switchboard of the database. While each subject was performing their tasks, the experimenter sat besides them to take notes about their performance and the problems in the test. The experimenter did not answer any questions during the test. The time spent on each job was recorded. At the end of each task, subjects were asked to verbalize his/her comments about that task. They could say anything they wanted, but their comments centered on problems, difficulties and suggestions. Their verbalizations were recorded on an audiotape recorder.

At the end of the session, each subject was asked answer the following four questions:

1. Write down your general comments about the usability of this database.
2. Please give your suggestions for modification.
3. What do you think are the major problems of this database (if applicable)?
4. Do you think the interfaces are easy to use?

The audiotapes and written comments as well as the notes were kept for later data analysis. The audiotape was transcribed and the notes were reviewed. The problems found in the test and the suggestions given by subjects were listed and classified. Then the database was modified.

RESULTS

THE DATABASE

FORMS

Sixty-two forms made of the interfaces of the database. These include fourteen switchboard forms that guide the user to navigate the database, 12 forms and subforms for data entry and 36 dialogue boxes that help to customize the querying process by allowing users to select items from a drop-down list.

1. Main Switchboard (Figure 3): This lists the functions the database has and is the main guide of the database navigation. When the database is opened, the main switchboard is shown on the screen. From it, the user can go to the "data entry" form to accomplish data entry, search the database to get the needed information by selecting "Queries" or "Print Reports" which summarize medical information of a patient, and recover "Patient ID" by "Searching For Patient ID" of a patient.
2. "Data Entry" Form (See Appendix C). This is the main entrance for data entry. It is linked to three forms: "Add New Patient", "Edit Patient Information" and "Enter Visit Information". "Add New Patient" (Figure 4) enables the user to enter the general information of a new patient. "Edit Patient Information" enables the user to modify a patient's general information in case there is a change. In order to ensure data integrity, some of the fields that should not be changed were locked, such as "gender", "citizen ID" and "date of birth" etc. "Enter Visit Information" is the core of the database because it is the entrance point where all visit-related information of the database is entered. It has a series of forms and subforms linked to it. When the user clicks the "Enter Visit Information" on the "Data Entry" form, a "Visit Information" form is shown.

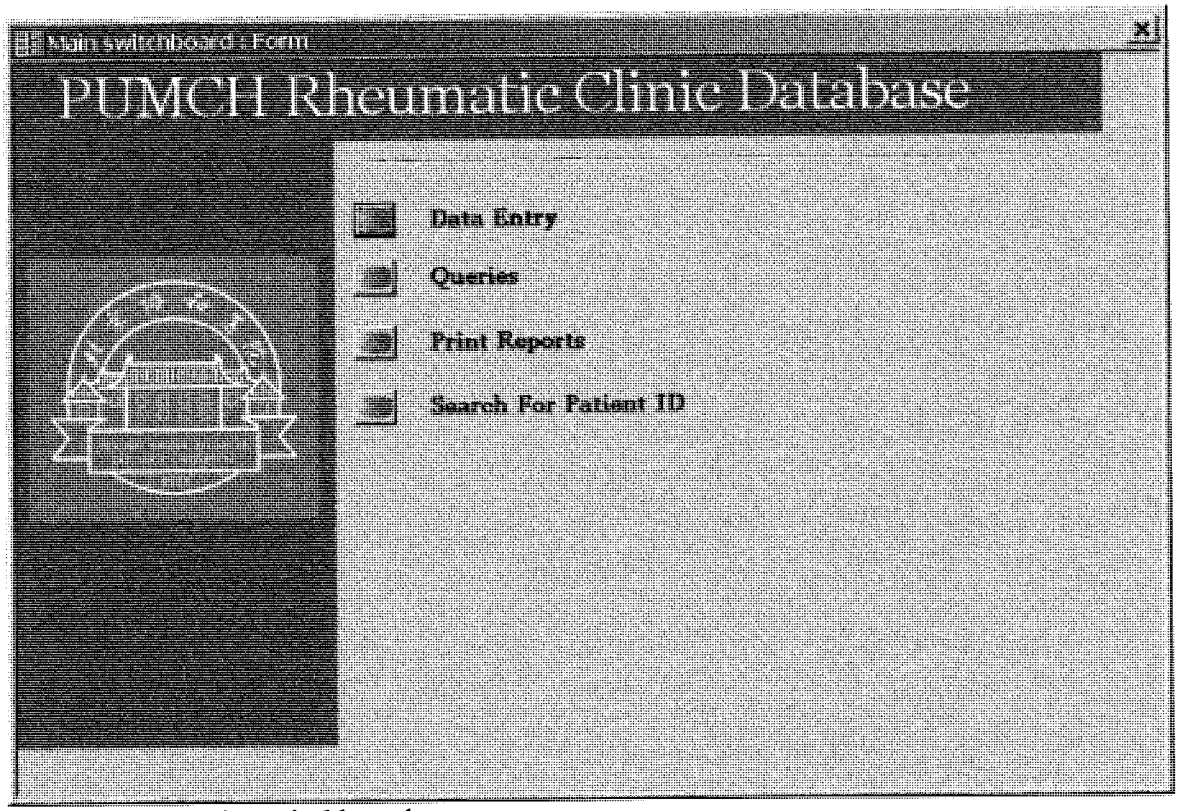


Figure 3. The main switchboard

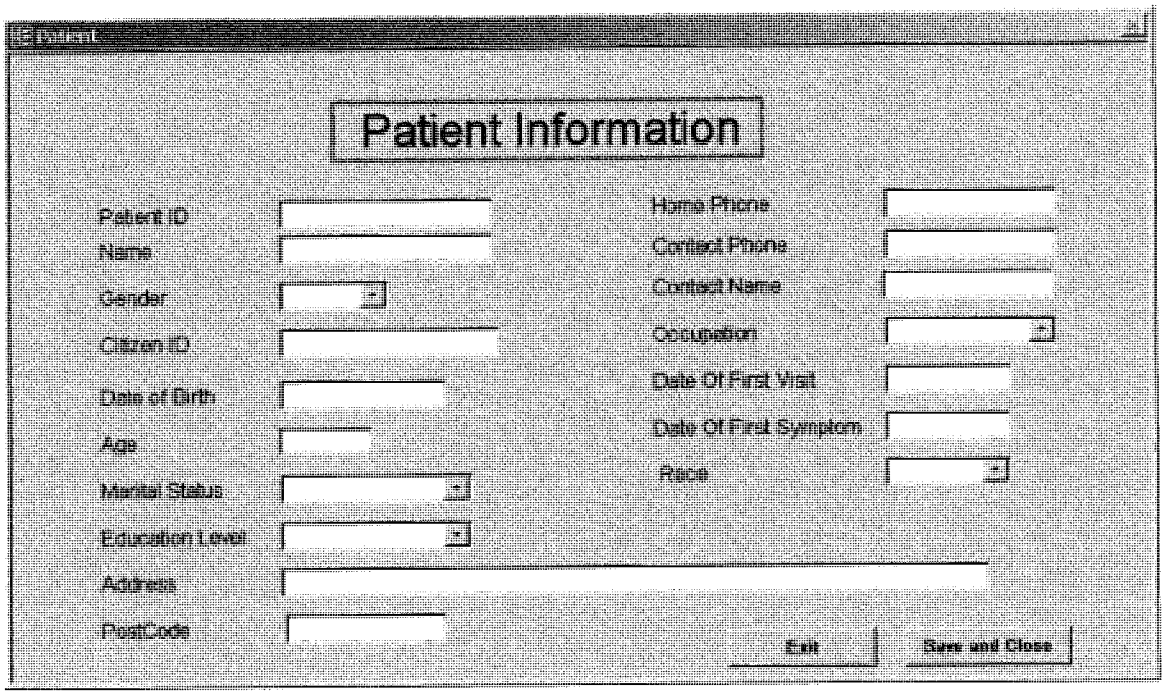


Figure 4. The "Patient Information" form

3. The "Visit Information" Form (Figure 5). The "Select Patient ID" field allows the user to select the Patient ID of the patient whose information is intended to enter by either selecting from a drop down list or typing in. When Patient ID is entered, the patient's name is shown automatically in the "Name" field. This provides a check to assure that the right patient is selected. The user needs to enter the "Date of Visit" and confirm it. Next, the user can enter a patient's laboratory tests, diagnoses, medications, radiological examinations, procedures, serum and DNA bank information for that visit date by selecting the corresponding button on the form. Each button is linked to a data entry subform that enables the user to enter related information into the database (See Appendix C). The "Visit Information" form also has a "Drugs currently used by the patient" subform that summarizes all the medications the patient is currently taking. When Patient ID is selected, this form is shown automatically. This helps the user to update patient's medication information such as change of dosage and withdrawal of a medication.
4. Queries list forms. All queries of the database can be accessed through a series of forms. They not only provide easy access to queries, but provide a guide for the user to find the right query (See Appendix C). 11 "Queries" list forms were built to guide users in finding the queries that can provide the appropriate information they want to search for. The "Main Queries Menu" and all other query list forms are shown in Appendix C. When the user selects a query and clicks the button beside it to run the query, a pop-up box appears. This box (Figure 6) customizes the searching by allowing the user to select the items he/she wants to search. Therefore, the user does not need to know exactly what the items

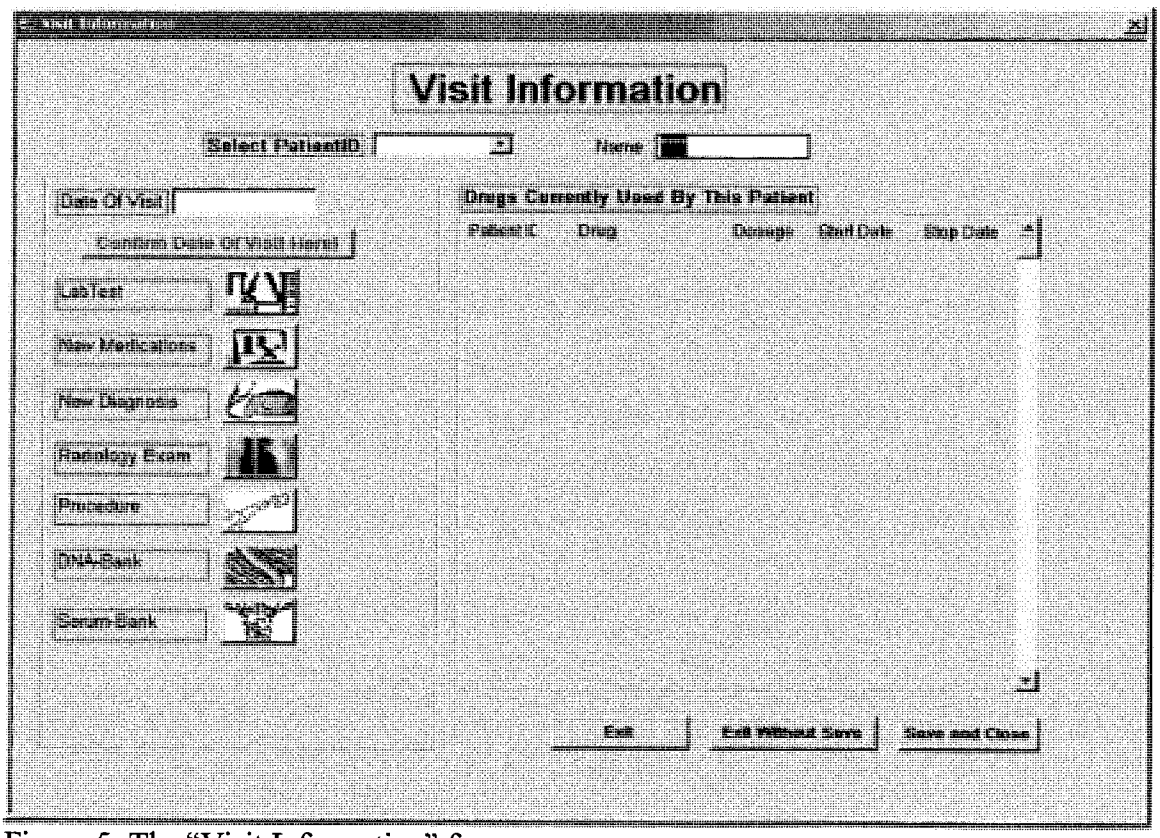


Figure 5. The "Visit Information" form

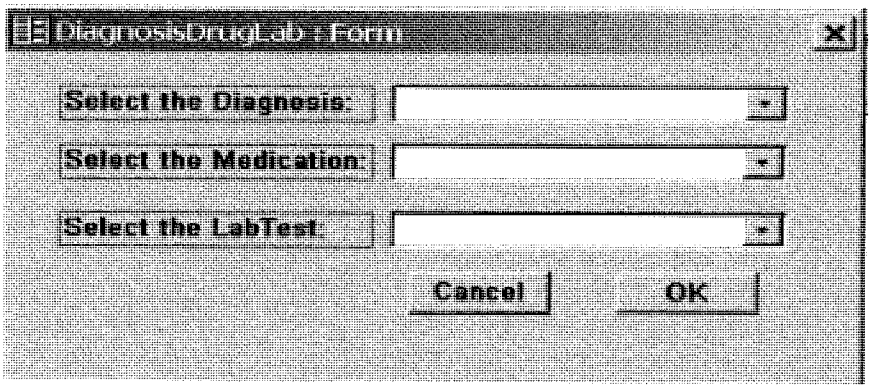


Figure 6. The pop-up dialogue form used for query items selection

are in the database such as, the names of diagnosis and medications. This can also avoid the tediousness and possible errors from typing.

QUERIES

The "Queries" button on the main switchboard leads the user to 33 queries built into the database. These queries were developed trying to answer the most frequently asked questions from clinical researchers.

Queries developed in the database can be generally classified into two categories: 21 queries about one field (such as queries about a single diagnosis, laboratory test or medication), and 12 queries about more than one field.

1. Queries about one field include:

- Queries about Diagnosis
- Queries about Medications
- Queries about Laboratory Tests
- Queries about Radiological Examinations
- Queries about Procedures
- Queries about Bank Information (Serum or DAN Bank)

2. Queries about more than one fields are more complicated than the first category, but more useful for clinical research. 12 queries were developed in this database including:

- Queries about Diagnosis and Medications
- Queries about Diagnosis and Laboratory Tests
- Queries about Medication and Laboratory Tests
- Queries about Diagnosis and Radiological Examination
- Queries about Diagnosis, Medication and Laboratory Tests
- Queries about Diagnosis, Medication and Radiological Examination

The details of these queries are shown in Appendix D.

REPORTS

Reports are another important component of the database. This database not only contains reports that summarize some of the medical information such as diagnoses, medications or laboratory tests of a patient, but also contains reports for rheumatologists to obtain information about Serum and DNA bank.

USABILITY TEST AND DATABASE MODIFICATION

The purpose of the usability test was to find usability problems and then modify the database to make it more efficient and more user-friendly.

SUBJECTS

Subjects were recruited as described in the "METHODOLOGY" part. Three of them were female and two were male. The average age was 27 years old (range 24-35). One had a Biology undergraduate background, one had business undergraduate education, one had nursing undergraduate background and two were physicians. Four were in the first year and one was the second year of the OHSU Master of Science in Medical Informatics program.

TEST RESULTS

For all subjects, no one failed to complete their tasks. The average total time to complete all tasks was 64 minutes (range 48-82 minutes including time for reading the scenario, verbalization recording and written comments). The average time for "Data Entry" was 38 minutes (ranges 25-48 minutes), average time for each case is 9.5 minutes. The subjects who spent less time did not enter the "Start Date" and "Stop Date" of medications. The average time to run a single query was 3.4 minutes (ranges 3-5

minutes). Only one subject spent five minutes in this part and all others finished this part in three or less minutes. The average time spent to print a report was less than one minute. The average time spent on identifying a patient was less than one minute.

All subjects thought the interface of the database was easy to use. The subjects indicated that the interface was generally "intuitive" although some modifications will make them more intuitive. They found the color of the interfaces was comfortable to the eye. The general layout of the database and the graphics were reported to be good and made the navigation of the database easy. Most subjects thought that it took a while to understand how to use the data entry screens, but not too long ("not more than a few minutes"). They thought the queries were useful, helpful and flexible. They allowed the user to select the items they want to search for and enter cutoff value for laboratory tests. They found printing reports to be very easy and well-labeled. They also thought that "Searching for Patient ID" was very simple and easy. It allowed the user to select citizen ID from the drop-down list, which avoids tedious typing of long numbers.

As expected, some problems were identified in the usability test. These problems can be classified into the following four classes:

- Format problems
- Consistency problems
- Layout problems
- Other problems

The following describes in more detail both the positive and negative findings for each task.

Data Entry

All subjects were able to complete this task. While it was relatively slow for them to enter the data from the first record, it became faster for the following records. Most subjects entered the data correctly, but one entered the wrong Patient ID number. Four of five subjects did not enter the start or stop date of medications.

Their comments included: "overall, it is pretty nice", " It is good that I can either choose the items such as laboratory test and procedure items from the drop-down list or I can also type in", "The interfaces for laboratory test, diagnosis and medication are good", "It is easy to add new patients".

Problems found in this task included the following:

- Format problems: All subjects complained about the difficulty in entering "dates". They all thought that the format of "date" was unclear. Because it was not clearly shown on the screen, it "takes a while to figure it out" and "it slows me down". Some subjects complained that "certain rules must be followed in order to enter the right date" for example, one must type '01' instead of '1'" and one must put the cursor in a certain place to "make sure that that you have enough space for it".
- Consistency Problems:
    a. When the data entry system was designed, all patient IDs began with "C-". However, they begin with "C/" on the "patient general information" screen and "C-" on the "Visit Information" screen. Three subjects complained about this because it made them confused about the correct format of the patient ID.

b. There is a "contact phone" field on the screen, but there is no such a field on the medical record print out. Instead, there is a "Day Phone" field. Subjects were confused, although with time all finally figured out that these two are the same.

c. The name of some items is not consistent. For example, the name of "Bone Marrow Aspiration is "BM Aspiration" in the data entry drop-down list of data entry subforms, but is "Bone Marrow Aspiration" in the medical record printout.

- Layout Problems:

a. Button layout. Two subjects mistakenly closed the "Visit Information" form when they actually tried to close one of the data entry subforms. Because there is a "Save and Close" button in all the data entry subforms and the "Visit Information" form, the layout of this button in the "Visit Information" form and its subforms was clearly troublesome for some subjects.

b. Layout of the "Visit Information" form. Some subjects complained that they needed to go back to check the medical record printout to see whether the patient had a new diagnosis made in the current visit or not. This slowed down the data entry. They felt that the display of previous diagnoses would be helpful.

c. The arrangement of data fields on the screen does not correspond to that in the medical record printouts. Some subjects complained that they needed to find the corresponding field from the printout.

- Other Problems

  Other complaints from subjects included:

  a. They were not used to typing so many long numbers such as citizen ID and phone number, so it was easy to make errors;

  b. There is a default "0" in laboratory test result fields. "It does not go away unless you manually delete it. Otherwise, you will have an extra '0' in the field".

Queries

Most subjects thought that the queries were very useful, particularly for physicians. "This database can make queries about more than one medical fields. This is very helpful to physicians", "The query part was good", "I was very happy that I could choose the diagnosis and laboratory tests that I wanted and also specify the cutoff values", " I like this query part" and "It is pretty clean".

Following problems were found in this task:

- Layout Problem

  a. Because the query the subjects were asked to run was about "diagnosis" and "laboratory test", many subjects tried to run the query using "Queries about diagnosis" button instead the correct "Queries about more than one fields" one. After finding that this button was not the correct one, they then had to search the list for the correct button. Therefore, one subject said "it is a little bit difficult to do a more than one field query".

b. Two subjects complained that the output of the query was separated into 2 pages, so they could not view the result of query if they did not scroll to the second page.

- Other Problems

One subject concerned that "queries about more than one fields" is not intuitive as the word "field" "does not mean too much for people have no knowledge of database".

Print Report

All subjects accomplished this job very quickly. They all thought that it was very easy and the reports were well labeled. "It is very clean", "I like the lines between the results so it is easier for you to read the information", "It is helpful to have the patient information on the top of the page such as Patient ID, gender and age", "It is nice."

There is no problem found in this part.

4. Results of "Searching for Patient ID" part of the database

All subjects found the right patient in less than 1 minute. They all thought it was simple, useful, convenient and easy. One subject said "I like the drop-down list for citizen ID so that I need not to type all the 15 characters by myself."

No problem was found in this part of the test.

Suggestions

15 items of suggestions were given by subjects:

- The field layout of the "Patient information" on the screen should follow that of the medical record printout.

- The "Back to Main Menu" button should be bigger than others, so that it is easier to see and does not look like it is to be one of the other navigation buttons.
- Allow user to add new queries.
- A dash separating sections of long numbers such as citizen ID and phone number.
- Separate home address into several fields with a street address and other fields in order to make "home address" entry easier.
- Have a drop down list with all the laboratory test results in order to avoid typing errors.
- Separate the "name" into "family name" and "given name" field to ease future data searching.
- Have a pop-up box to describe each query button;
- Integrate all data entry subforms of "Visit Information" into one interface;
- Provide copy and paste options;
- Show the patient's previous diagnoses on the data entry screen.
- Show the diagnoses on the report of "Summary of Medications" so it is easier for one to understand why some medications were prescribed to this patient;
- Create shortcut so that one can enter data (such as open and close a form) just using the keyboard without shifting between mouse and keyboard.
- Enter some "date" fields such as "Stop Date" of medications automatically by building a button beside them;
- Implement speech recognition to help entering long numbers such as citizen ID.

INTERFACE MODIFICATION

After results of the test were analyzed with problems classified, and suggestions listed, then the database was modified. Several changes were made.

Modifications in "Data entry"

- The format of "date" is clearly shown beside the "date" fields on the screen, so the user can know the format easily. Also the inflexible restrictions were removed so user needs not follow a certain rules (for example, user can just type "1" instead of "01") or worry about whether there is enough space for it or not.
- The field layout of the "Patient Information" screen was modified to match the layout of the medical record printouts. The modified "Patient Information" form is shown in Figure 7.
- Modify the layout of "Visit Information" form and its subforms
- Each individual data entry subform is fixed to a specific position so "Save and Close" button of the "Visit Information" form cannot be seen when any of the data entry subform is opened. This eliminates the chance of pressing the wrong button and closing the wrong form. In addition, this change can also enable the

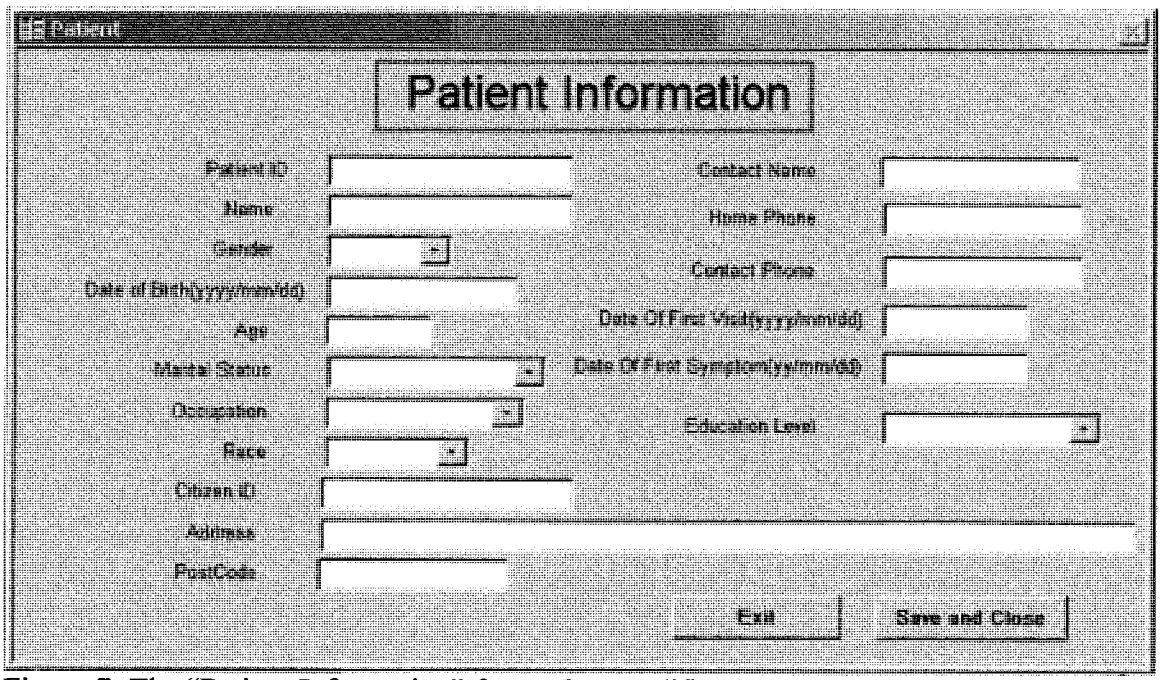


Figure 7. The "Patient Information" form after modification

user to see the Patient ID and name to make sure that information of the right patient is entered.

- Show Patient's previous diagnoses on "Visit Information" form
- In order to ease the updating of a patient's diagnoses information, a subform "Diagnoses This Patient Had:" was built into the "Visit Information" Form. When a Patient ID is selected, the current diagnoses that the patient has are shown automatically on the screen. This way, the user can see whether new diagnoses were made at the current visit or not. The modified "Visit Information" is shown in Figure 8.

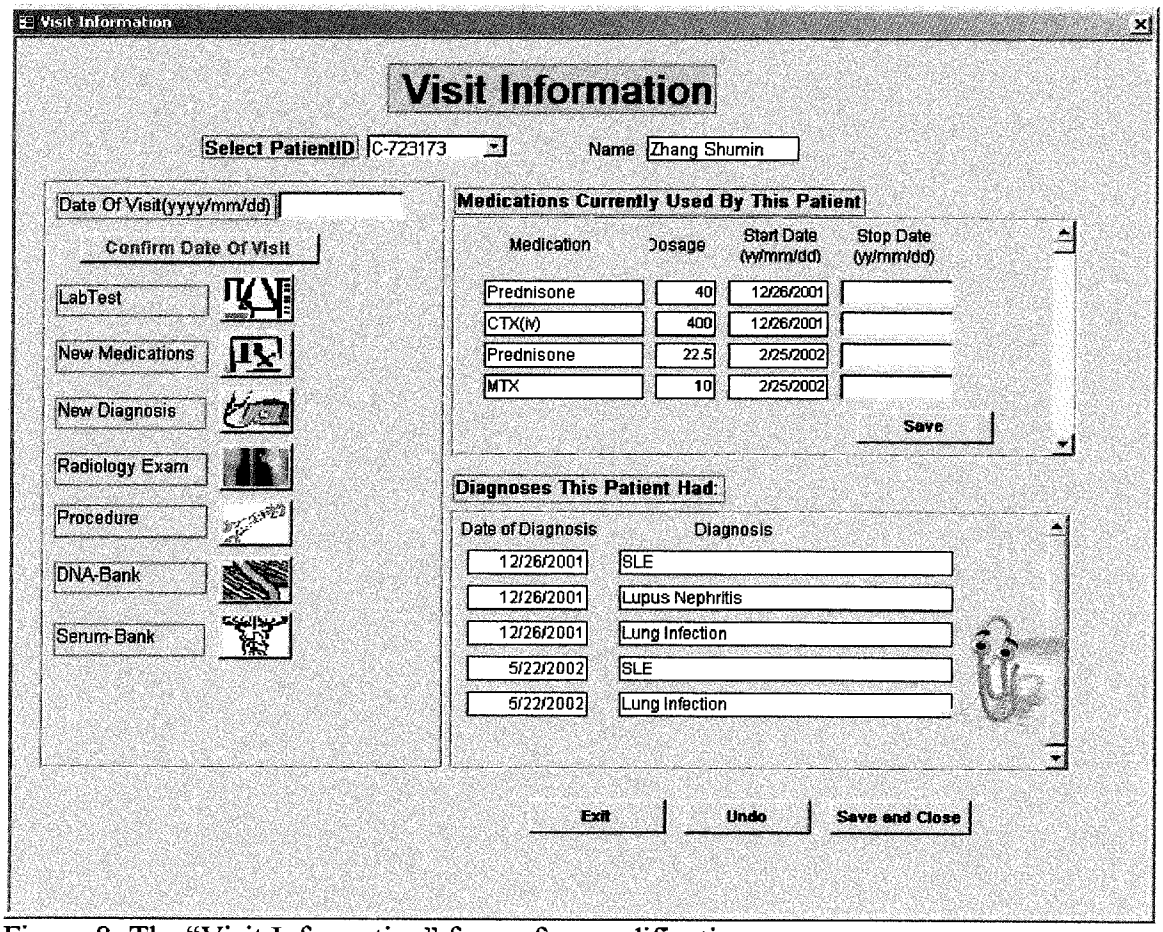


Figure 8. The "Visit Information" form after modification

- The default "0" in the laboratory test result fields was removed.

Modification in "Queries"

Since one subject concerned about the word "field" in the query might make no sense for users without database knowledge, the word "field" was changed to "medical domain".

- Because the "Queries About More Than One Fields" was mixed with other queries, some subjects thought that it was not easy to be found without reading through the whole list. Therefore, another form was inserted that lists the two general categories of queries in order to make it a "guide" for queries (See Appendix C). This not only makes "Queries About More Than One Domain" easier to be found, but also shows what kind of queries this database can provide.
- A mouse-over text tip that describes the query in detail was built for each query.
- The "Back To Main Query Menu" Button was enlarged and moved to the right lower corner of the form. This way, it does not look like one of other query navigation buttons. The modified form is shown in Appendix C.
- The layout of query output was modified so that the output of queries can be shown in a single page. Therefore, the user can view results of query without scrolling to another page.

Modifications in "Print Report"

A small box with a list of all the diagnoses the patient currently has was built at the right upper corner of the report (See Appendix C). In order to make the reports more informative, this list box was added to reports that summarize patients' medications, laboratory tests, radiological examinations and procedures.

Other Modifications

Adding more function to the database:

Although this is not based on the usability test, from a practical point of view, it would be helpful to rheumatologists if they could access the overall health information of patients they are interested in. Therefore, an overview function was added for this purpose. After selecting the target Patient ID, some basic information of the patient such as age, gender, date of first symptom and date of first visit would show automatically. Forms that show all diagnoses, medications, laboratory tests, procedures, DNA and Serum Bank information of the patient could be viewed by clicking on the corresponding buttons on the form. The main switchboard after medication and the overview form are shown Appendix C.

Suggestions not adopted for modification

Although most of the suggestions from subjects were adopted for modification, not all of them were used. The following suggestions were not adopted:

- A dash separating long numbers

  As the area code of phone numbers varies in different part of the country, it is difficult to put a "-" in a fixed position. Entering citizen ID is a challenge, because of the unique structure of it, it is difficult to put a "- " in an appropriate position.

- Separate home address into several fields

  While separating home address into "street" field and other fields may speed up entering home address of urban patients, it is not applicable to patients from rural

area because the format of home addresses of rural areas is quite different from urban areas.

- Have a drop down list with all the laboratory results to avoid typing errors

  It is almost impossible to include all laboratory test results in a single drop-down list because all of the biochemical laboratory test results are consecutive numbers in a wide range.

- Separate the "name" into "family name" and "given name" field

  As there are only two to three characters in Chinese names, it is still easy to be retrieved even when they are in the same field.

- Integrate all the subforms of "Visit Information" into one interface

  This is good for simple visits, i.e., visits with few data. However, this will cause confusion for complicated visits with a lot of information.

- Create shortcuts so that user can enter data using the keyboard only

  Since this database is not supposed to be used by sophisticated users and there are many shortcuts already been built into Microsoft Access itself, no more were created.

- Create a button that enables automatic "date" entry for some "date" fields

  Although this is a good suggestion to speed up data entry, a method for implementing it does not appear to exist.

- Implement voice recognition to help entering long numbers such as citizen ID.

  This might be a potential solution for long numbers. But the availability of software for Chinese is the concern.

DISCUSSION

DATABASE AND USABILITY TEST

Medicine is an information intensive arena. The importance of patient information on clinical research and quality improvement of health care delivery cannot be overemphasized. However, the collection and organization of health information is a big challenge. Relational database technology has offered an option. In this study, a relational database that aimed at capturing visit-related medical information of all rheumatic patients was developed. This database can be used not only for data collection and organization, but also to retrieve information for clinical research and teaching.

The importance of usability testing in software and system development is well recognized. It helps developers discover usability problems before the software or system is implemented in a real life setting. Although usability testing in real-life would be ideal, in this instance, testing was not carried out with actual end-users. However, this study still provided much valuable information. Usability problems were found as a result of the test and the database was then modified accordingly. Although most of the suggestions from subjects resulted in database modification, not all of them were adopted. Some of the suggestions were not applicable to situations in China and some were infeasible in a real life setting.

Although the modified interfaces were not retested, it is reasonable to assume that the modified database is more user-friendly than the original one. However, since usability testing is an iterative process, just one test will not make the database problem-free. More tests will be done in the future to make sure the database can handle the role that it has been designed for.

Besides usability problems discussed above, there are some other problems that need to be addressed. It was observed in the test that most subjects did not enter the start and stop date of medications. Since this database was designed for the outpatient clinic and all medications were prescribed on the day of visit, a default date was automatically entered into the "start date" field, which is the same as the date of the visit. Therefore, it is unnecessary to enter "start date" for most medications. However, rheumatology is a unique medical specialty and so some medications are used in sequence, so physicians usually prescribe them at one visit. Although the sequence of medications was clearly indicated in the medical record printouts, only one subject paid attention to this and changed the default date to the actual "start date" for some medications. The reason for this is that most of the subjects were not physicians (the subject who entered the right date was a physician), so they did not pay attention to the actual start date when they saw a default in it. Another reason is that the experimenter did not emphasize that the actual start date and the default for some medications might be different. So this issue should be emphasized in the training stage for data entry clerks and they should be alerted to pay attention to some special medications.

Although there was no default date in the stop date field, three subjects did not enter them. One probable reason is that it was rather slow to enter dates because of the previously discussed format problem, so they just did not want to enter it. Therefore, training is absolutely necessary to data entry clerks before they really start to use the system.

The time each subjects spent on data entry was relatively long. One reason was that they were very unfamiliar with the database. Since the majority of subjects were not

physicians, it was hard for them to understand the clinical data collected in the database. They did not understand the data entry flow because it was not shown to them before the testing. They needed to figure this out by trying. Another reason is that we wanted to know how much time would be spent on entering the most complicated cases--the worst cases of data entry in real life setting. So the new cases and return cases represent the most complicated cases that might be expected in real clinic setting--both have a lot of information. Most cases in real life settings are not as complicated as the testing cases and data entry clerks will be trained to be familiar and proficient in using the database before they begin. In addition, data entry clerks will be expected to have a medical background, so they can understand clinical data easily.

FUTURE WORK

As this database was developed for the rheumatic disease outpatient clinic of PUMCH, the current screens will first be translated into Chinese before implementation at PUMCH. Another usability test will be conducted in the real life setting before its implementation. The database will be modified thereafter.

This database is just a start. Although it was developed particularly for rheumatic disease clinic, it can be adapted for other internal medicine outpatient clinics since they share many common characteristics. It can also be modified to meet the information needs of a surgical department or other procedure-focused medical fields. Therefore, this database can be used as the prototype of an outpatient clinical system of PUMCH and hopefully can be built upon it in the future.

CONCLUSION

A relational database can be designed to collect, store, organize and retrieve visit-based medical information of patients from an outpatient clinic. In this thesis project, such a database was designed and built to meet information needs of clinical research, improving health care continuity and teaching of rheumatic diseases of PUMCH. The "simplified thinking aloud" testing method was a useful tool for carrying out usability testing because of its simple procedures and low cost. It was very efficient in discovering usability problems. The results from the testing can be used to guide database modification. Although the usability testing of this project was not conducted in real life setting, it is reasonable to assume that its interfaces are more user-friendly than the original one.

REFERENCES

1. Callahan L, Yelin E. The social and economic consequences of rheumatic disease. In: Klippel JH(ed). Primer on the rheumatic diseases, 12th ed. Atlanta: Arthritis Foundation, 2001; p1-4.
2. Theodore P. Why should rheumatologists collect patient self-report questionnaires in routine rheumatologic care? Rheum Dis Clin North Am 1995;21(2):271-319.
3. Singh G. Arthritis, rheumatism and aging: medical information system post-marketing surveillance program. J Rheumatol 2001;28(5):1174-1179.
4. Zink A, Listing J, Zeidler H. The national database of the German collaborative arthritis centers: I. Structure, aims, and patients. Ann Rheum Dis 2001;60:199-206.
5. Zink A, Niewerth M, Zeidler H. The national database of the German Collaborative Arthritis Centers: II. Treatment of patients with rheumatoid arthritis. Ann Rheum Dis 2001;60:207-213.
6. Midedema H,Van Der Linden S, Rasker J et al. National Database of Patients Visiting Rheumatologists in the Netherland: the standard diagnosis register of rheumatic diseases. A report and preliminary analysis. Br J Rheumatol 1998;37:555-5613.
7. Bowyer S, Roettcher P. Pediatric rheumatology clinic populations in the United States: results of a 3 year survey. J Rheumatol 1996;23(11):1968-1974.
8. Symmons D, Jones M, Osborne J et al. Pediatric rheumatology in the United Kingdom: data from the British pediatric rheumatology group national diagnostic register. J Rheumatol. 1996;23(11):1975-1980
9. Malleson P, Fung M, Rosenberg A. The incidence of pediatric rheumatic diseases: results from the Canadian pediatric rheumatology association disease registry. J Rheumatol 1996;23(11):1981-1987
10. Wolfe F. A database for rheumatoid arthritis. Rheum Dis Clin North Am 1995;21(2):481-500.
11. Edworth M. A database for systemic lupus erythematosus and systemic connective tissue disorders. Rheum Dis Clin North Am 1995;21(2):501-525)
12. Littlejohn G. A database for fibromyalgia. Rheum Dis Clin North Am 1995;21(2):527-557.
13. Raspe H. A database for back (axial skeletal) pain. Rheum Dis Clin North Am 1995;21(2):559-579
14. ISO. ISO CD9241-11: Guidelines for specifying and measuring usability; 1993
15. Wichansky A. Usability testing in 2000 and beyond. Ergonomics 2000;43(7):15-18
16. Kushniruk A. Evaluation in the design of health information system: application of approaches emerging from usability engineering. Comput Biol Med 2002; 32:141-149
17. Nielsen J. The discount usability engineering approach. In: http://www.useit.com/ papers /guerilla-hci.html ed: Nielsen J; 1994
18. Rubin J. Handbook of usability testing. Canada:John Wiley & Sons, 1994

19. Kushniruk A, Patel V, Cimino J. Usability testing in medical informatics: cognitive approaches to evaluation of information system and user interfaces. In: AMIA Annual Fall Symposium; 1997:JAMIA:218-222

APPENDICIES

APPENDIX A

TABLES OF THE DATABASE

I. TABLES AND FIELDS

Patient
PatientID(KEY)
Name
Gender
Age
Date of Birth
Occupation
Marital Status
Educational Level
Race
Address
Postcode
Home Phone Number
Contact Name
Contact Phone Number
Date of First Visit
Date of First Symptom

Visit
VisitID(KEY,Autonumber)
Date of Visit
PatientID

VisitDiagnosis
VisitDiagnosisID(KEY,Autonumber)
VisitID
DiagnosisID

VisitLabTest
LabTestDetailID(KEY,Autonumber)
LabTestID
VisitID
Result

VisitDrug
DrugDetailID(KEY,Autonumber)
VisitID
DrugID
Dosage
StartDate
StopDate

VisitRadio
RadioDetailID(KEY,Autonumber)
VisitID
RadioligyTestID
Result

VisitProcedure
PatientProcedureID(KEY,Autonumber)
VisitID
ProcedureID
Result

VisitSerum
SerumID(KEY)
VisitID

VisitDNA
DNAID(KEY)
VisitID
Result

Diagnosis
DiagnosisID(KEY,Autonumber)
DiagnosisName
SNOMEDCode

LabTest
LabTestID(KEY,Autonumber)
LabTestName

Drug
DrugID(KEY,Autonumber)
DrugName

RadiologyTest
RadiologyTestID(KEY,Autonumber)
ExamName

Procedure
ProcedureID(KEY,Autonumber)
ProcedureName

II. LOOKUP TABLES

Table 1. Diagnoses Included In the Database

Diagnosis ID	DiagnosisName	SNOMEDCode
1	SLE	55464009
2	Lupus-related Endocarditis	54072008
3	Lupus Nephritis	68815009
4	Lupus Hepatitis	19682006
5	Discoid Lupus Erythematosus	200938002
6	Systemic Sclerosis	89155008
7	Acrosclerosis	50803006
8	CREST Syndrome	31848007
9	Scleroderma	372929001
10	Systemic Sclerosis Lung Involvement	196133001
11	Systemic Sclerosis-related Myopathy	193252005
12	Primary Sjogren's Syndrome	239912009
13	Sicca Syndrome	267875002
14	Dermatomyositis	38826005
15	Polymyositis	31384009
16	Eosinophilia Myalgia Syndrome	95416007
17	Unspecified Diffuse Connective Tissue Disease	239918008
19	SLE glomerulonephritis syndrome	68815009
20	Lupus disease of the lung	233730002
21	Lupus Encephalopathy	95644001
22	Lupus Panniculitis	239888002
23	Bullous Lupus	239889005
24	Neonatal Lupus Erythematosus	95609003
25	Limited lLpus Erythematosus	239886003
26	SLE with pericarditis	309762007
27	Cutaneous Lupus Erythematosus	7119001
28	Reiter's Disease	67224007
29	Overlap Syndrome	276657008
30	Rheumatoid Arthritis	69896004
31	Felty's Syndrome	57160007
32	Ankylosing Spondylitis	9631008
33	Spondyloarthropathy	372109003
34	Undifferentiated Spondyloarthropathy	202649003
35	Psoriatic Arthritis	33339001
36	Enteropathis Arthritis	9350004
37	Juvenile-onset Spondyloarthropathy	239805001
38	Anti-phospholipid Antibody Syndrome	26843008
39	Giant-cell Arteritis with polymyalgia rheumatica	239938009
40	Temporal Arteritis	87511001
41	Takayasu's Arteritis	359789008
42	Polyarteritis Nodosa	66121003
43	Kawasaki's Disease	75053002
44	Wegener's Granulomatosis	195353004
45	Churg-Strauss Syndrome	82275008
46	Microscopic Polyarteritis	239928004
47	Henoch-Schonlein Purpura	191306005
48	Essential Cryoglobulinemia Vasculitis	190815001

Diagnosis ID	DiagnosisName	SNOMEDCode
49	Cutaneous Leukocytoclastic Angitis	238785001
50	Lupus Erythematosus Profundus	15084002
51	Mixed Connective Tissue Disease	33110008
52	Subacute Cutaneous Lupus Erythemmatosus	239891002
53	Polyarthritis	30701005
55	Juvenile Dermatomyositis	1212005
56	Myositis Associated with autoimmune Disease	239899000
57	Myositis Associated with Malignancy	239898008
58	Inclusion Body Myositis	72315009
59	Eosinophilic Polymyositis	240120003
60	Myositis Ossificans	44551007
61	Polymyalgia Rheumatica	65323003
62	Toxic Myopathy	66952001
63	Post-infection Arthritis	239783001
64	Myopathy Caused by Drugs	87858002
65	Juvenile Rheumatoid Arthritis	239801005
66	Secondary Sjogren's Syndrome	239915006
67	Juvenile Scleroderma	239904001
68	Rheumatic Fever	58718002
69	Gout	24595009
70	Hyperuricemia	35885006
71	Pseudo-gout	60782007
72	Jaccoud's Syndrome	84801008
73	Arthropathy	8316001
74	Osteoarthritis	267888004
75	Relapsing Polychondritis	72275000
76	Unspecified polyarthritis	270506005
77	Acute polyarthritis	42898009
78	Lyme Disease	23802006
79	Juvenile Psoriatic Arthritis	239802003
80	Acute Rheumatic Arthritis	81077008
81	Arthropathy secondary to inflammatory bowel disease	239814006
82	Amyloidosis	17602002
83	Sarcoidosis	31541009
84	Solitary Sacroiliitis	239815007
85	Unspecified Inflammatory Monoarthritis	220000
86	Hypertrophic Osteoarthropathy	203357004
87	Fibromyalgia	24693007
88	Undifferentiated inflammatory oligoarthritis	239819001
89	Behcet's Disease	310701003
90	Juvenile Chronic Arthritis	239796000
91	Adult-Onset Still's Disease	239920006
92	Reactive Arthropathy	239783001
93	Hypersensitivity Vasculitis	60555002
94	Isolated angitis of CNS	230733004
95	Erythema Nodosa	32861005
96	Erythema Multiform	36715001
97	Charcot's arthropathy	359554008
98	Juvenile arthritis of inflammatory bowel disease	239809007
99	Fungal arthritis	372247004

Diagnosis ID	DiagnosisName	SNOMEDCode
100	Post-immunization Arthritis	239789002
101	Generalized Osteoarteritis	201819000
102	Transient Arthropathy	66191007
103	Palindromic Rheumatism	50442003
104	Arthralgia	57676002
105	Enthesopathies	23680005
106	Pulmonary Tuberculosis	81483001
107	Herpes Simplex Infection of skin	240475000
108	Herpes Zoster Infection	4740000
109	Juvenile chronic polyarthritis	7441009
110	Candida infection of Mouth	79740000
111	Coccidioidal Meningitis	46303000
112	Coccidioidal Pneumonitis	88036000
113	Candidal Pneumonia	3487004
114	Harshimoto Thyroiditis	21983002
115	Thyroiditis	82119001
116	Gouty Nephropathy	190829000
117	Uric Acid Nephrolithiasis	267441009
118	Still's Disease(juvenile RA)	287984007
119	Hypogammaglobulinemia	119250001
120	Agammaglobulinemia	119249001
121	Selective IgA Immunodeficiency	29260007
122	Selective IgM Immunodeficiency	190980000
123	Selective IgG Deficiency	190981001
124	Common Variable Immunodeficiency	23238000
125	Autoimmune Hemolytic Anemia	25121006
126	Evans Syndrome	25331009
127	Juvenile Spondyloarthropathy	239806000
128	Juvenile AS	239805001
129	Juvenile Reactive Arthritis	239807009
130	Thrombotic Thrombocytopenia Purpura	78129009
131	Autoimmune Thrombocytopenia	128091003
132	Agranulocytosis	165508008
133	Lung Involvement of Sjogren's Syndrome	196137000
134	Lung Involvement of Polymyositis	196136009
135	Juvenile Reiter's syndrome	239808004
136	Pulmonary Amyloidosis	196135008
137	Rheumatoid Lung disease	111280008
138	Diffuse interstitial pulmonary fibrosis	196125002
139	Lung Infection	128601007
140	Monoarthritis Juvenile RA	83793004
141	Arthritis secondary to amyloidosis	237876008
142	Osteoarthritis of hand	267889007
143	OA of elbow	239866002
144	OA of wrist	239867006
145	OA of Knee	239873007
146	OA of ankle	239874001
147	OA of foot	309246000
148	Rheumatoid Carditis	28880005
149	Rheumatoid Arteritis	80172006

Diagnosis ID	DiagnosisName	SNOMEDCode
150	Necrotizing Vasculitis	11791001
151	Crystal Arthropathy	18834007
152	Gout secondary to renal impairment	239844009
153	Gout secondary to drug	239845005
154	Limited Systemic sclerosis	299276009
155	Linear Scleroderma	22784002
156	Acute Scleroderma renal crisis	236503001
157	Morphea	201049004
158	Steroid-induced myopathy	26715006
159	OA of hip	239872002
160	Vasculitis	31996006
161	Systemic Vasculitis	46956008
162	Tuberculosis Meningitis	58437007
163	Fungal infection of the Lung	63741006
164	Autoimmune Liver disease	235890007
165	Primary Biliary Cirrhosis	31712002
166	Sclerosing Cholangitis	235917005
167	Steroid-induced Diabetes	190447002
168	Aseptic Necrosis of head of femur	203476003
169	Cerebral Infarction	20059004
170	Cerebral Hemorrhage	274100004
171	Diffuse idiopathic skeletal hyperostosis(DISH)	31487001
172	SLE(acute flare)	
173	Baker's Cyst	82675004

Table 2. Laboratory Tests Included In the Database

LabTest			
LabTestID	TestName	LabTestID	TestName
1	Hb	43	Anti-Histone Ab
2	WBC	44	Acl Ab
3	Plt	45	Anti-CCP Ab
4	Urine Protein	46	LA(Lupus anti-coagulants)
5	Urine RBC/dl	47	APTT
6	Urine RBC cast/HP	48	Anti-Syphills Ab
7	Granular cast/HP	49	Anti-ENA Ab
8	24hr urine protein	50	Anti-auto-Ab
9	ESR	51	CMV-antigen
10	CRP	52	Anti-EBV Ab
11	ALT	53	Anti-TB Ab
12	AST	54	CH50
13	Cr	55	C3
14	AKP	56	C4
15	GGT	57	Anti-DNP Ab
16	CPK	58	PT
17	LDH	59	PAIgG
18	HBDA	60	Coomb's Test
19	Uric Acid	61	g-globulin
20	ANA	62	Urine Glucose
21	Anti-dsDNA Ab(TE-IF)	63	Serum Glucose
22	Anti-Sm Ab	64	ANCA-PR3
23	Anti-SSA Ab	65	ANCA-MPO
24	Anti-SSB Ab	66	ASO
25	Anti-RNP Ab	67	HLA-DR2
26	Anti-rRNP Ab	68	ACA
27	Anti-Jo-1 Ab	69	Tbil
28	Anti-Scl-70 Ab	70	Dbil
29	ANCA	71	Alb
30	AKA	72	Anti-dsDNA Ab(Farr)
31	APF		
32	RF		
33	HLA-DR4		
34	HLA-B27		
35	HLA-B51		
36	IgG		
37	IgM		
38	IgA		
39	Anti-SM Ab		
40	Anti-myocardiac Ab		
41	Anti-mitochondria Ab		
42	Anti-platelet Ab		

Table 3. Radiological Examinations Included In the Database

RadiologyTest	
RadiologyTestID	**ExamName**
1	Hands X-ray
2	Knee X-ray
3	Feet X-ray
4	Shoulder X-ray
5	Lumbar X-ray
6	Sarco-iliac Joint X-ray
7	Hip X-ray
8	Elbow-joint X-ray
9	Chest X-ray
10	Chest CT scan
11	Sarcoilliac Joint CT scan
12	Brain MRI
13	Brain CT scan
15	Abdominal CT
16	Esophageal Barium
17	High-resolution Chest CT scan

Table 4. Procedures Included In the Database

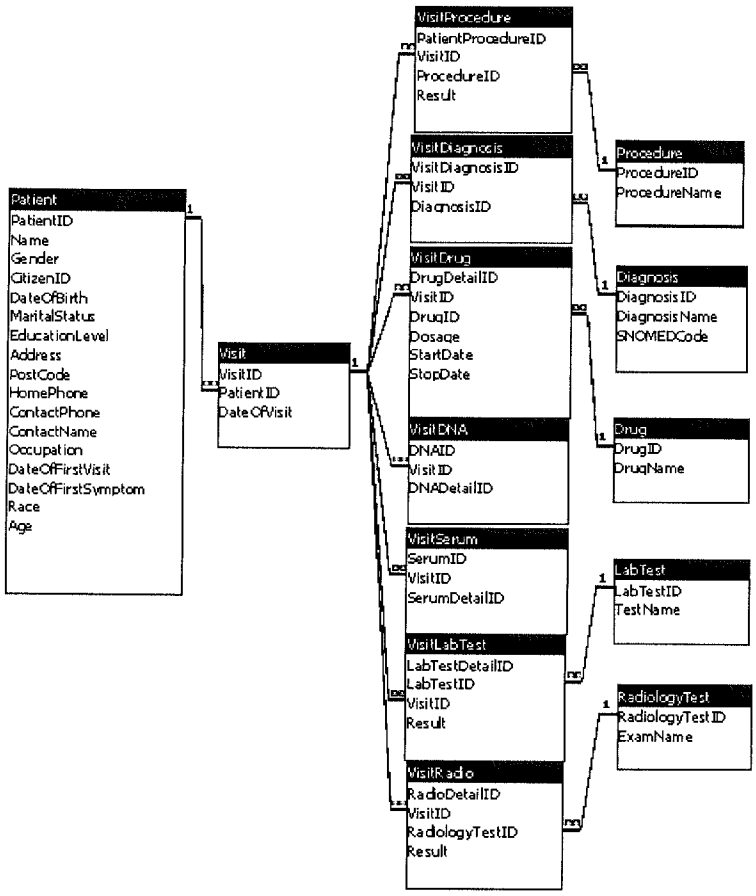
Procedure	
ProcedureID	**ProcedureName**
1	Bone Marrow Aspiration
2	Joint Aspiration
3	Bone Marrow Biopsy
4	Muscle Biopsy
5	Lymph Node Biopsy
6	Skin Biopsy
7	Subcutaneous Nodules Biopsy
8	Thoracic Parecentesis
9	Lumbar Puncture
10	EEG

Table 5. Medications Included In the Database

Medications	
MedicationID	**Medication Name**
1	Prednisone
2	CTX(po)
3	MTX
4	Prednisolone
5	Cyclosporine A
6	Azathioprine
7	Methylprednisolone
8	Colchicine
9	SASP
10	Allopurinol
11	Leflunomide(Arava)
12	Hydroxychloroquine
13	Chloroquine
14	D-pencillamine
15	Ridaura
16	CellCept
17	Warlfarin
18	Indomethacin
19	Ibuprofen
20	Naproxen
21	Sulindac
22	Diclofenic acid
23	Captopril
24	Acetaminophen
25	Tramadol
26	Celebrex
27	Vioxx
28	Aspirin
29	Nifedipine
30	IVIg
31	T2
32	Mobic
33	VCR
34	CTX(iv)
35	XiLuoMing
36	Gancyclovir
37	Imuran
38	Hydrocortison

APPENDIX B

THE RELATIONSHIP OF DATABASE TABLES AND FIELDS



APPENDIX C

FIGURES IN THIS DATABASE

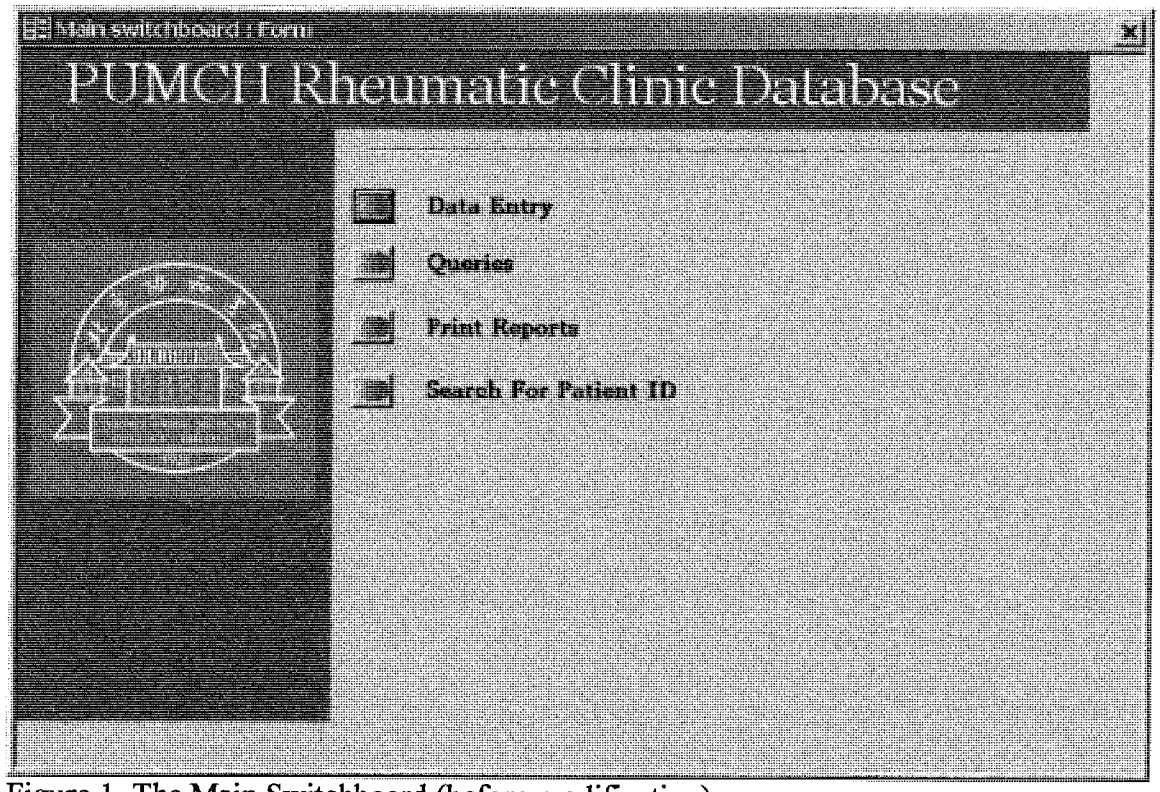


Figure 1. The Main Switchboard (before modification)

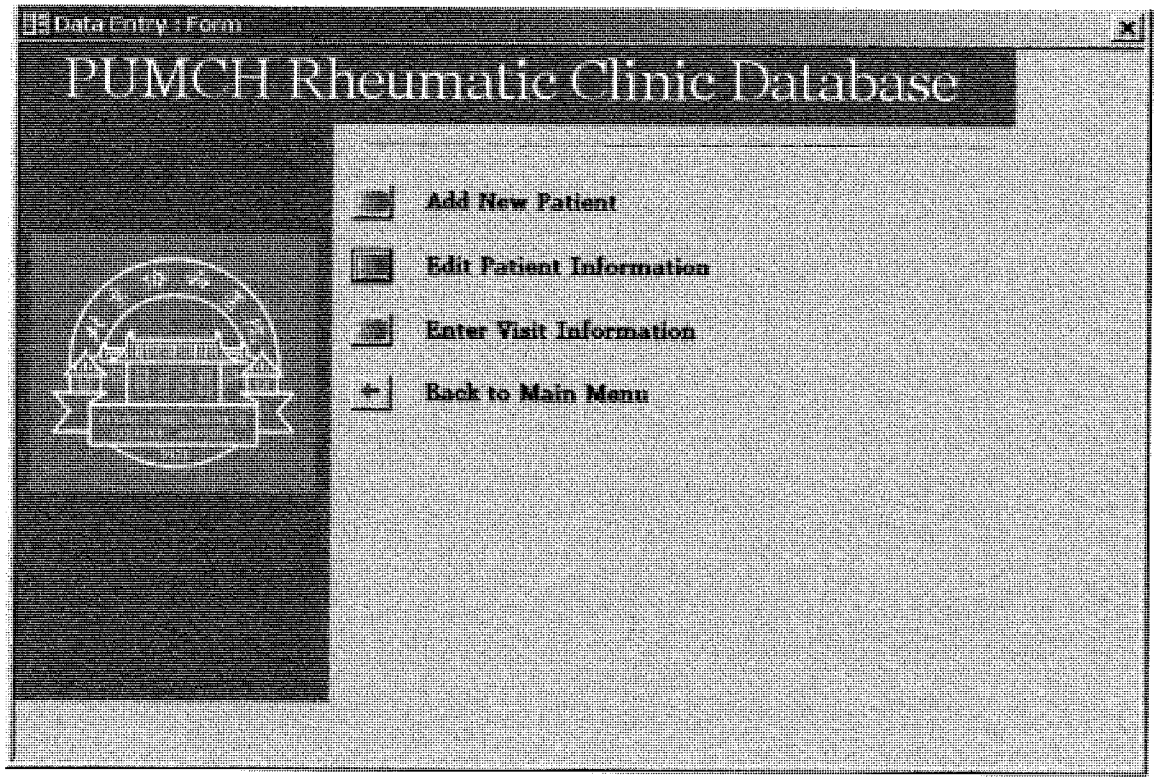


Figure 2. Data Entry Form (before modification)

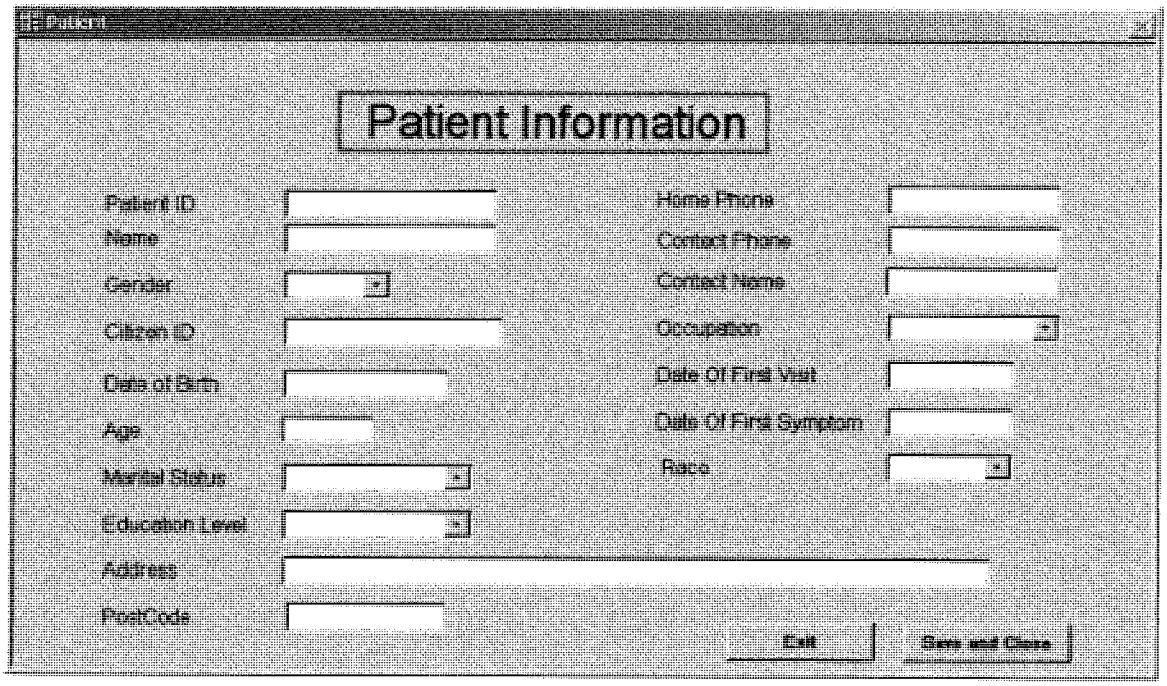


Figure 3. Patient Information Form(before modification)

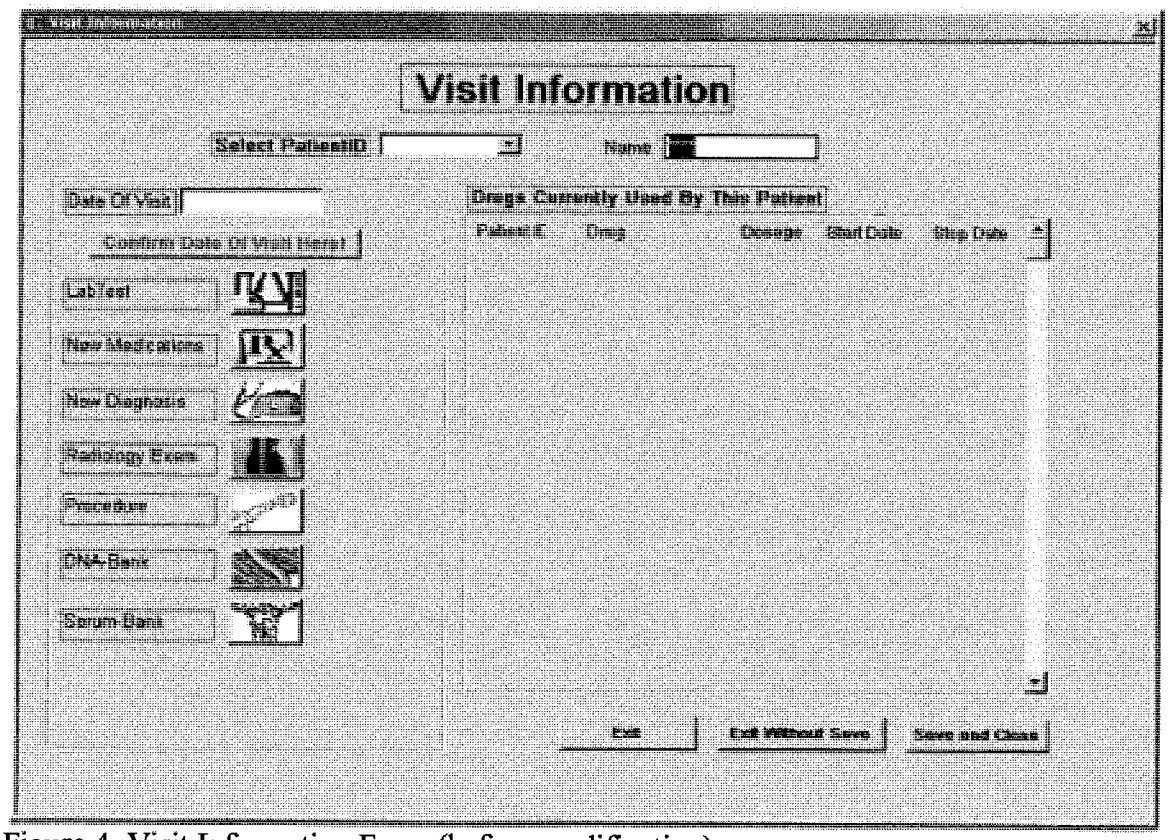


Figure 4. Visit Information Form (before modification)

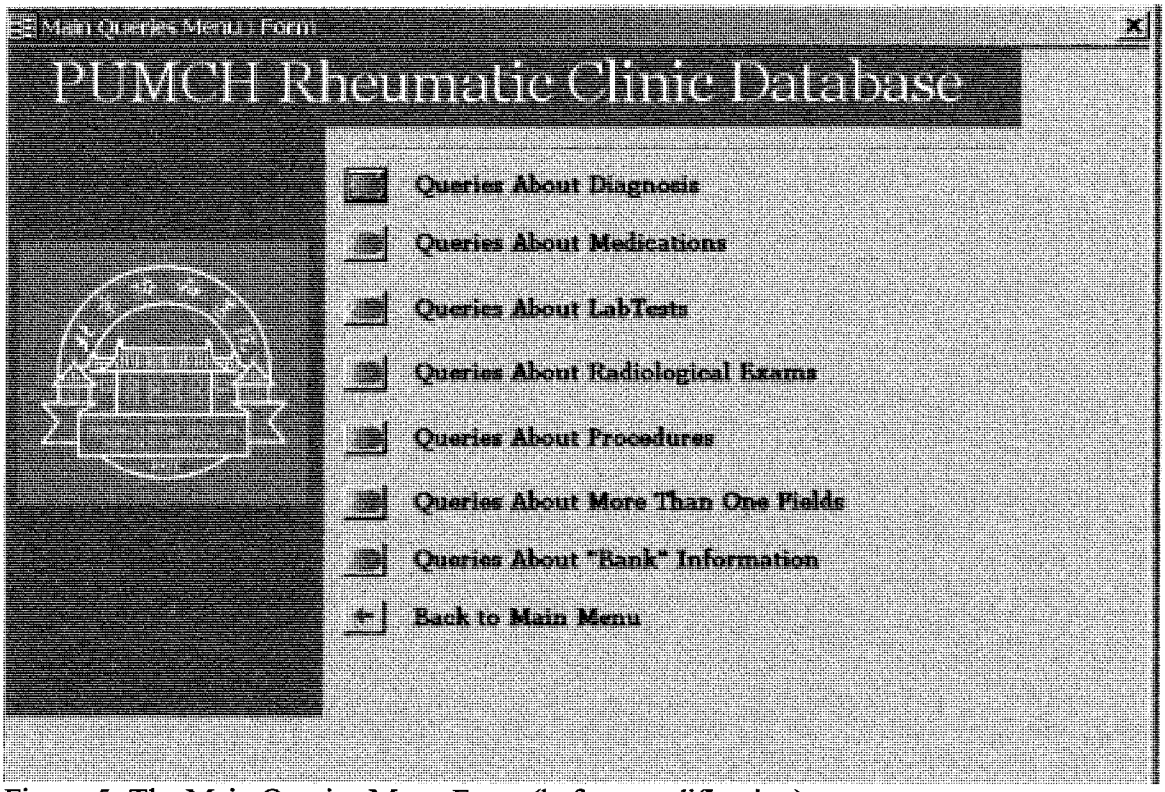


Figure 5. The Main Queries Menu Form (before modification)

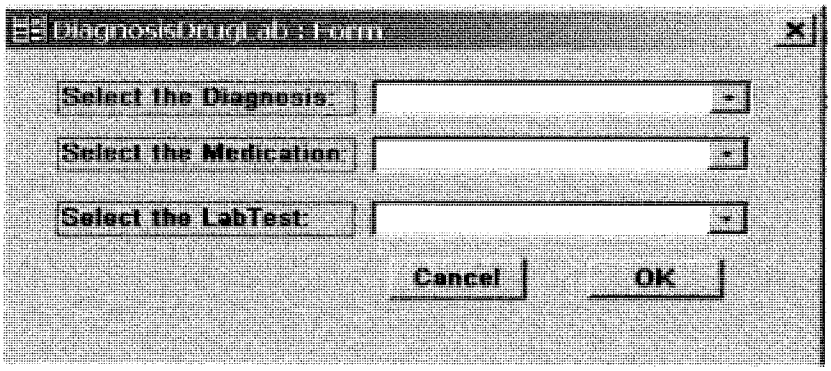


Figure 6. The Pop-up dialogue box form that enables to select items for query

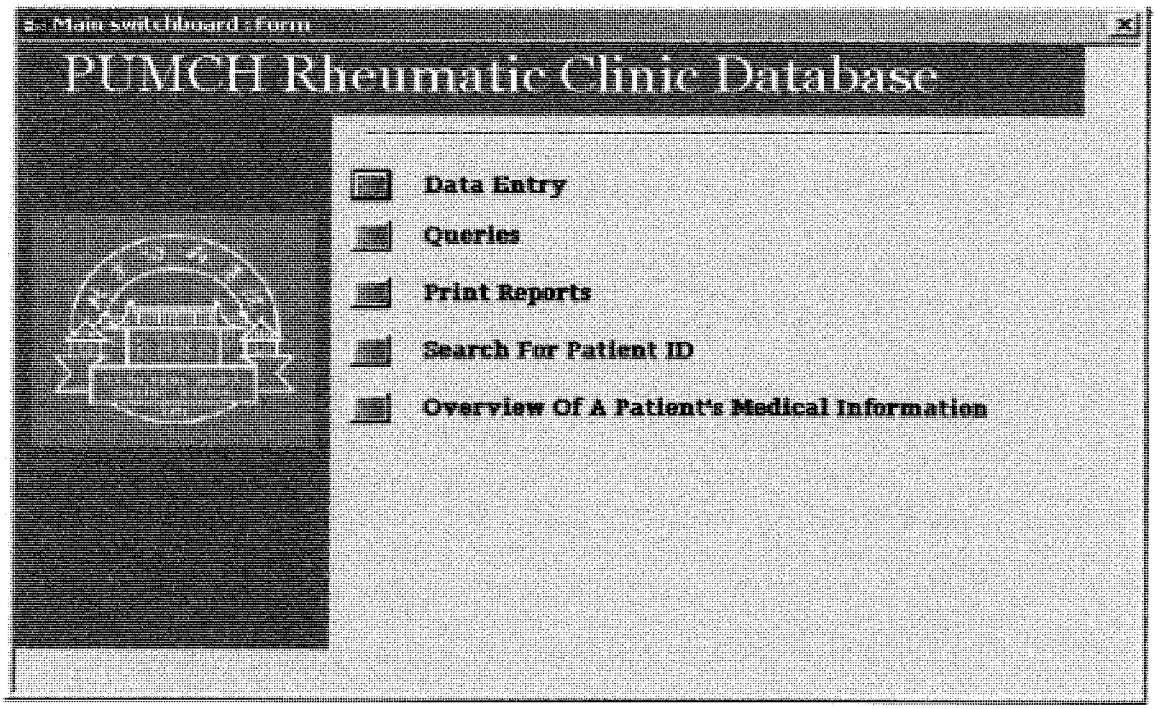


Figure 7. The Main Switchboard (after modification)

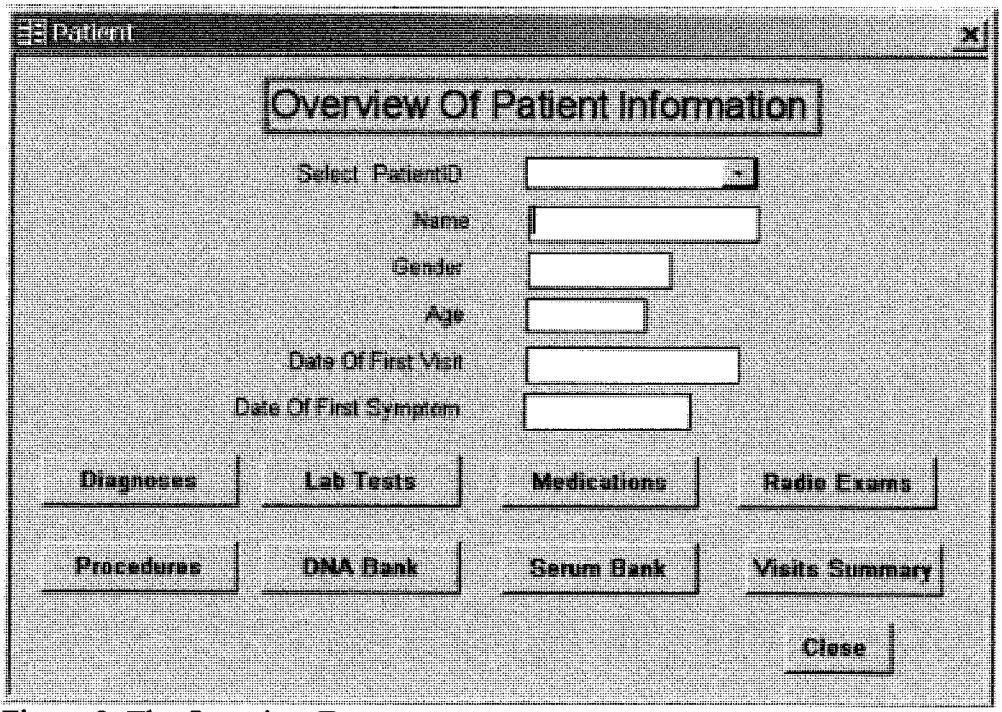


Figure 8. The Overview Form

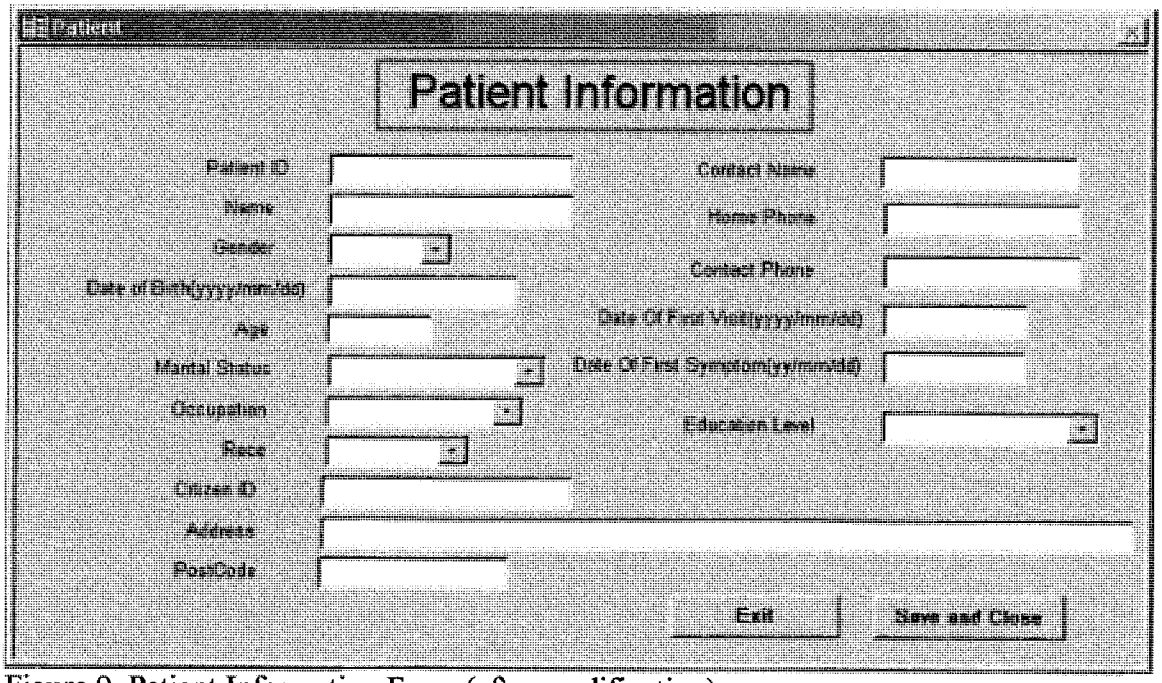


Figure 9. Patient Information Form (after modification)

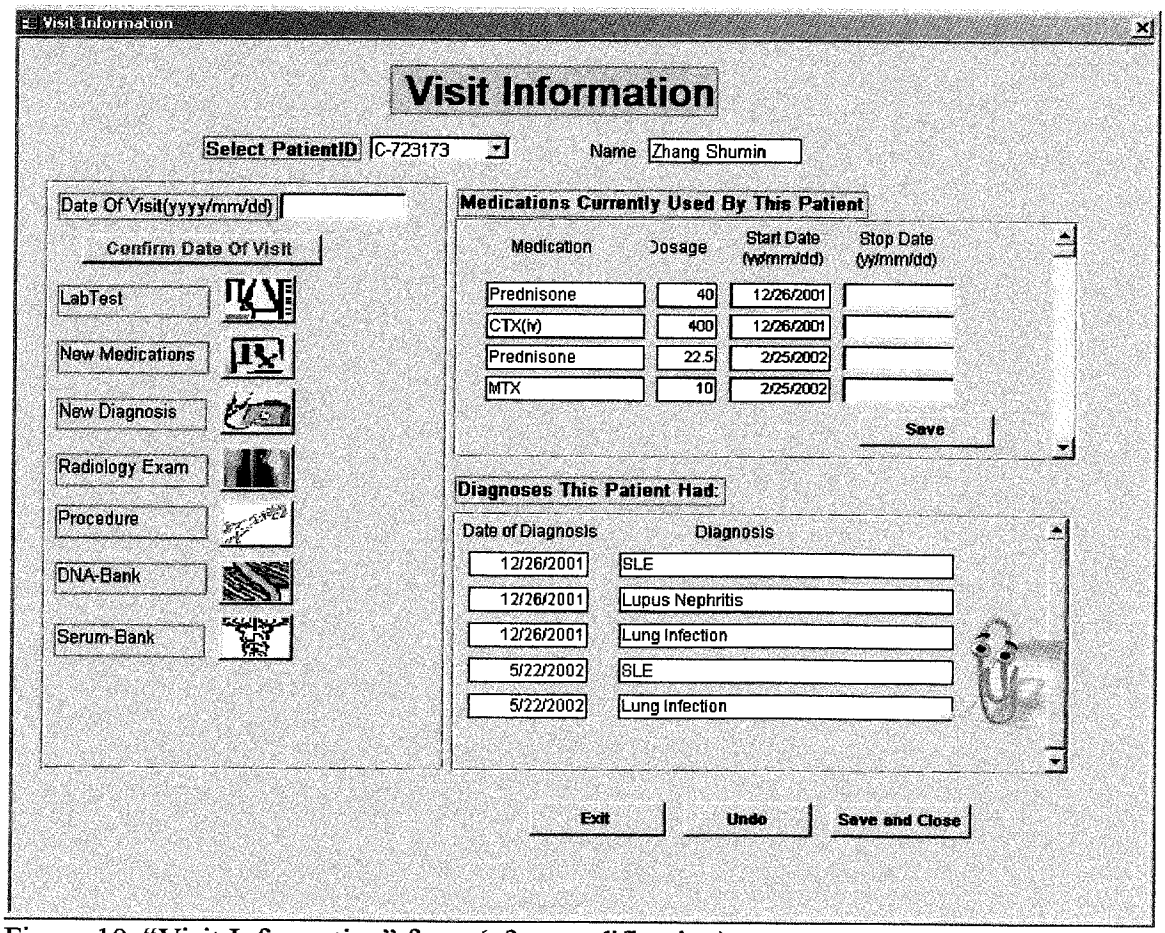


Figure 10. "Visit Information" form (after modification)

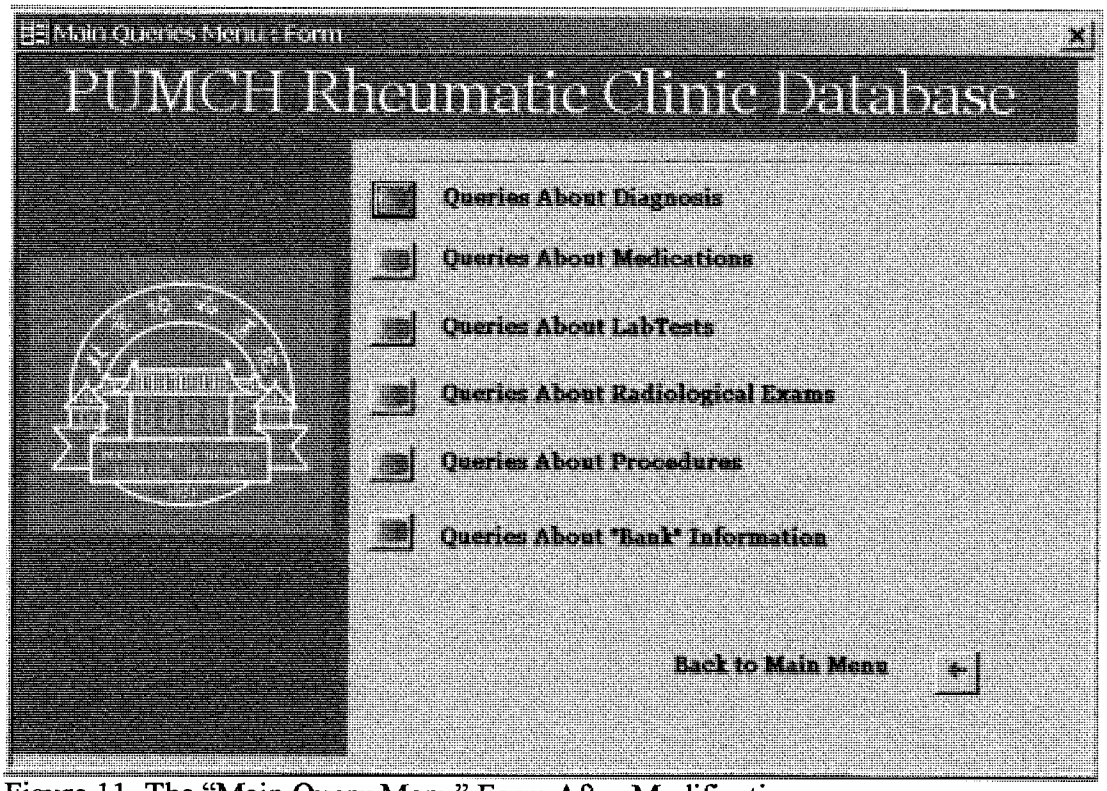


Figure 11. The "Main Query Menu" Form After Modification

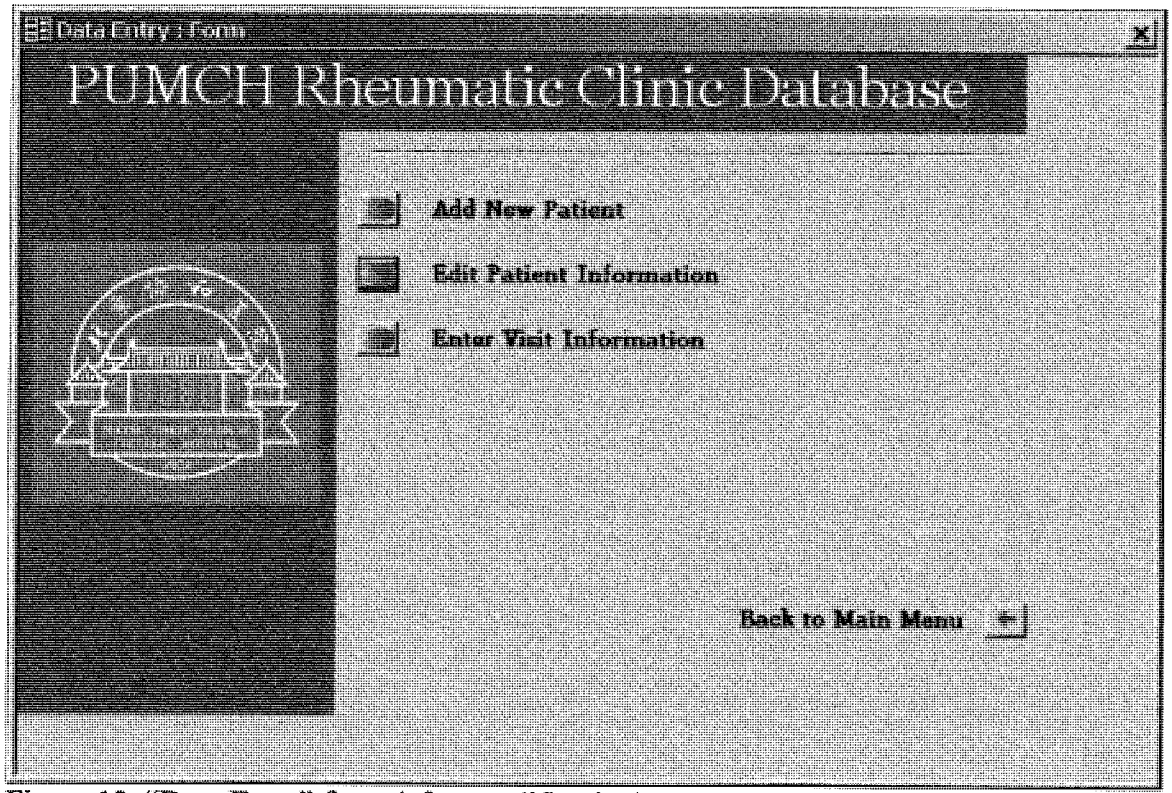


Figure 12. "Data Entry" form (after modification)

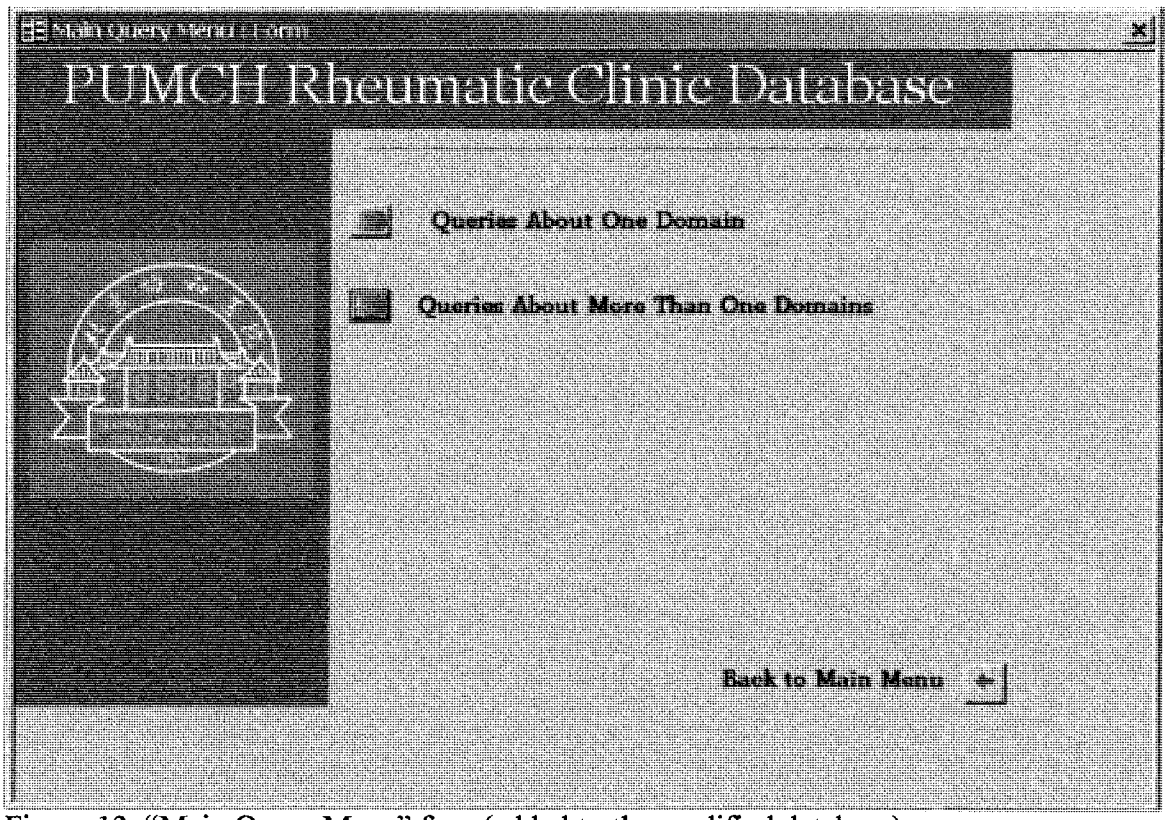


Figure 13. "Main Query Menu" form(added to the modified database)

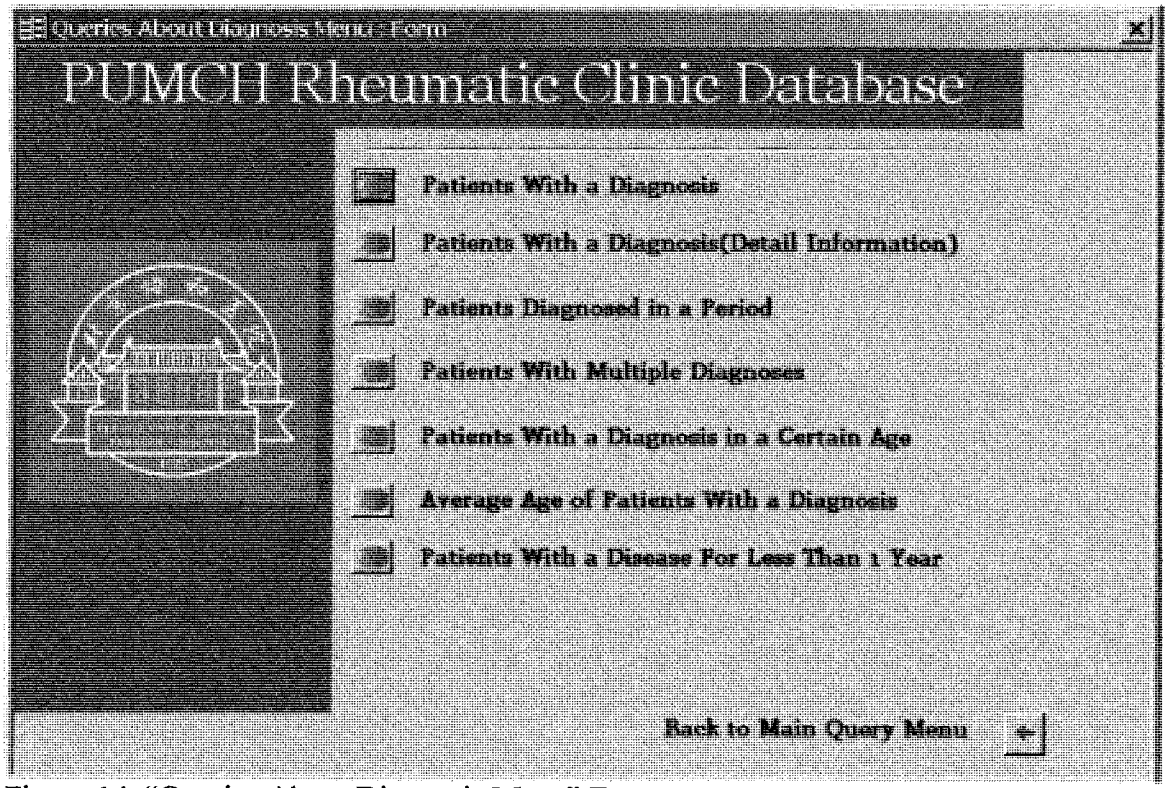


Figure 14. "Queries About Diagnosis Menu" Form

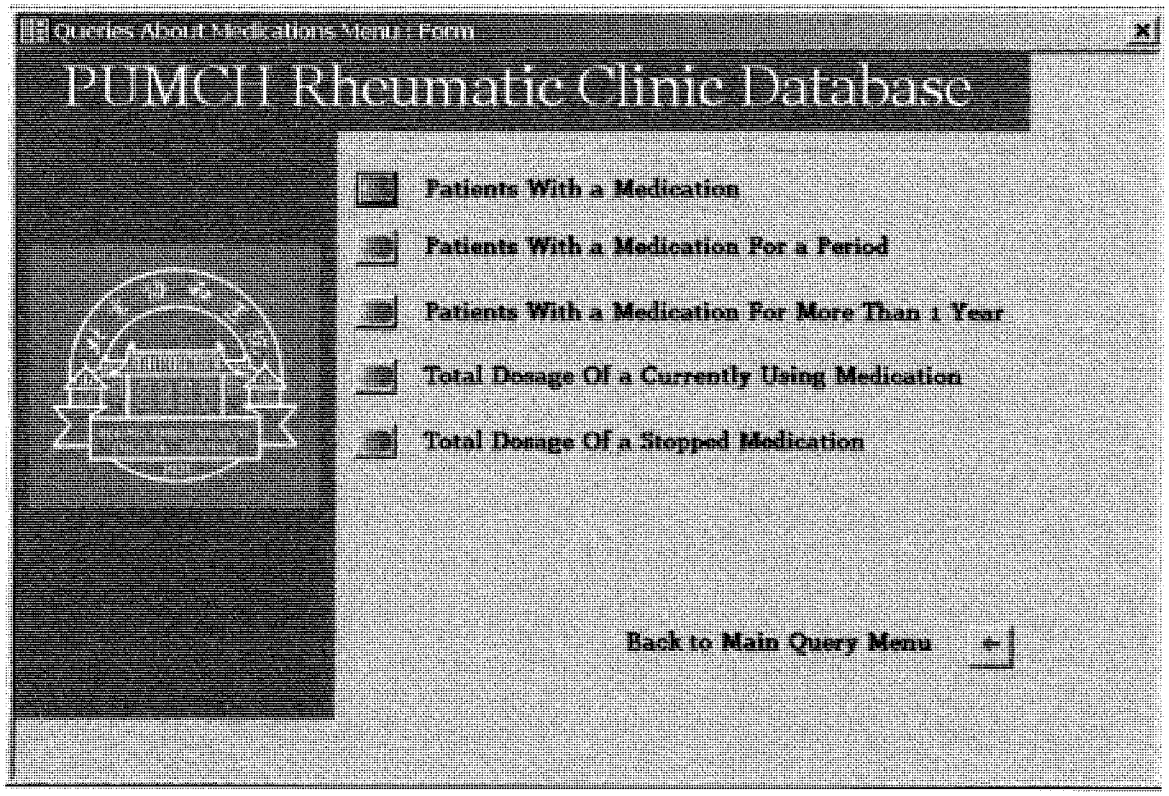


Figure 15. "Queries About Medication Menu" Form

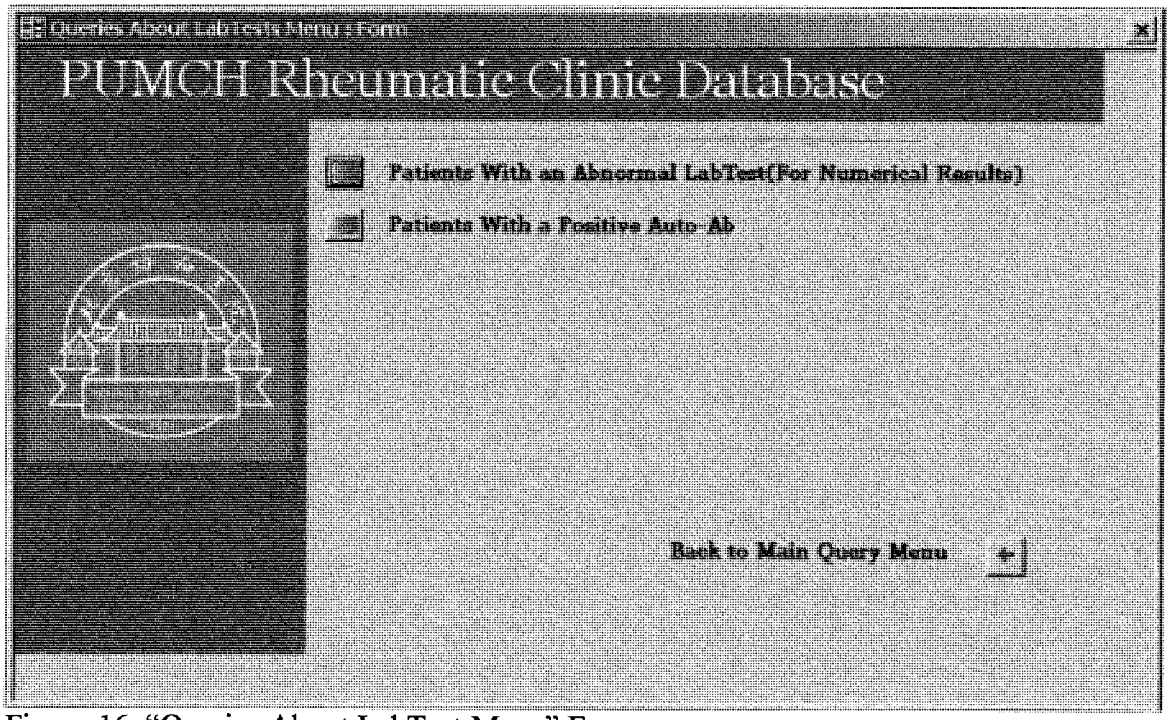


Figure 16. "Queries About LabTest Menu" Form

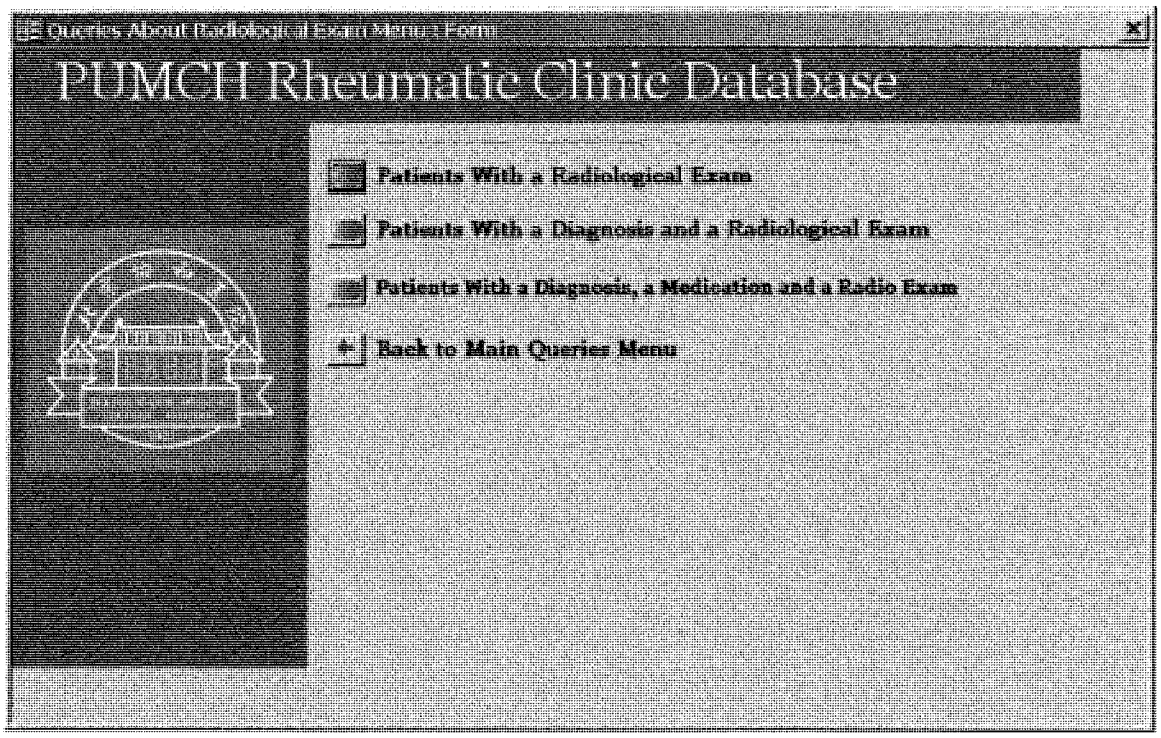


Figure 17. "Queries About Radiological Examination Menu" Form

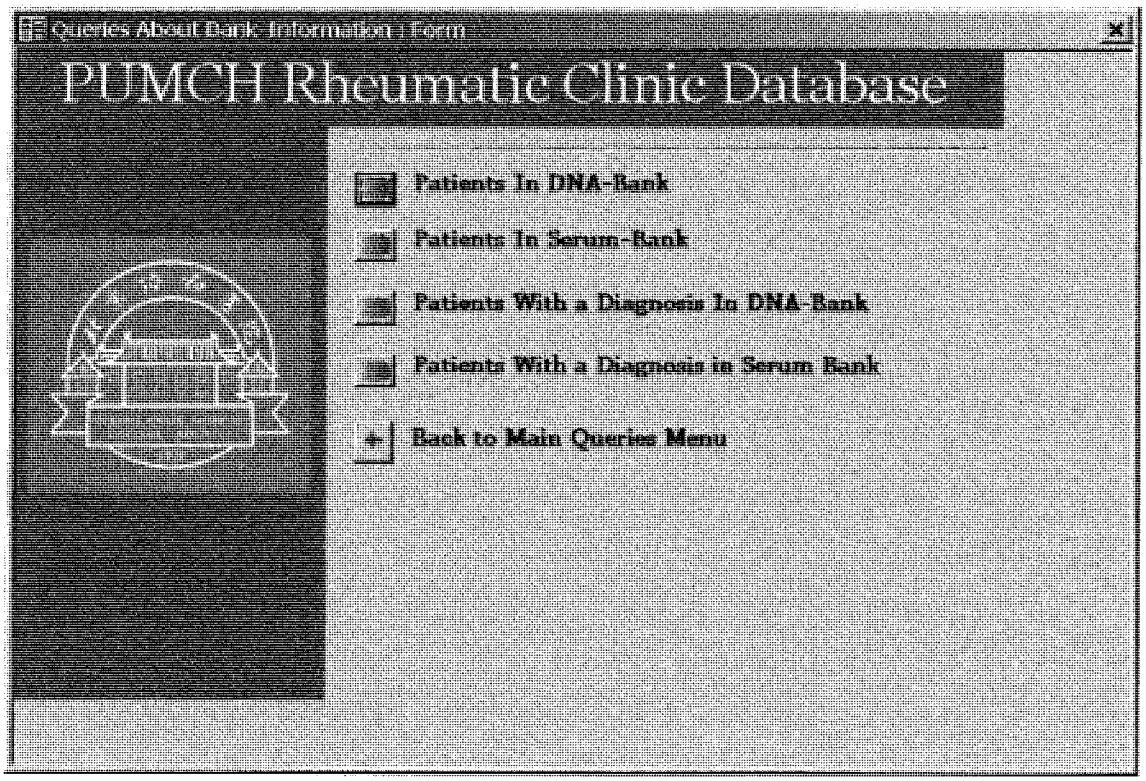


Figure 18. "Queries About Bank Information" Form

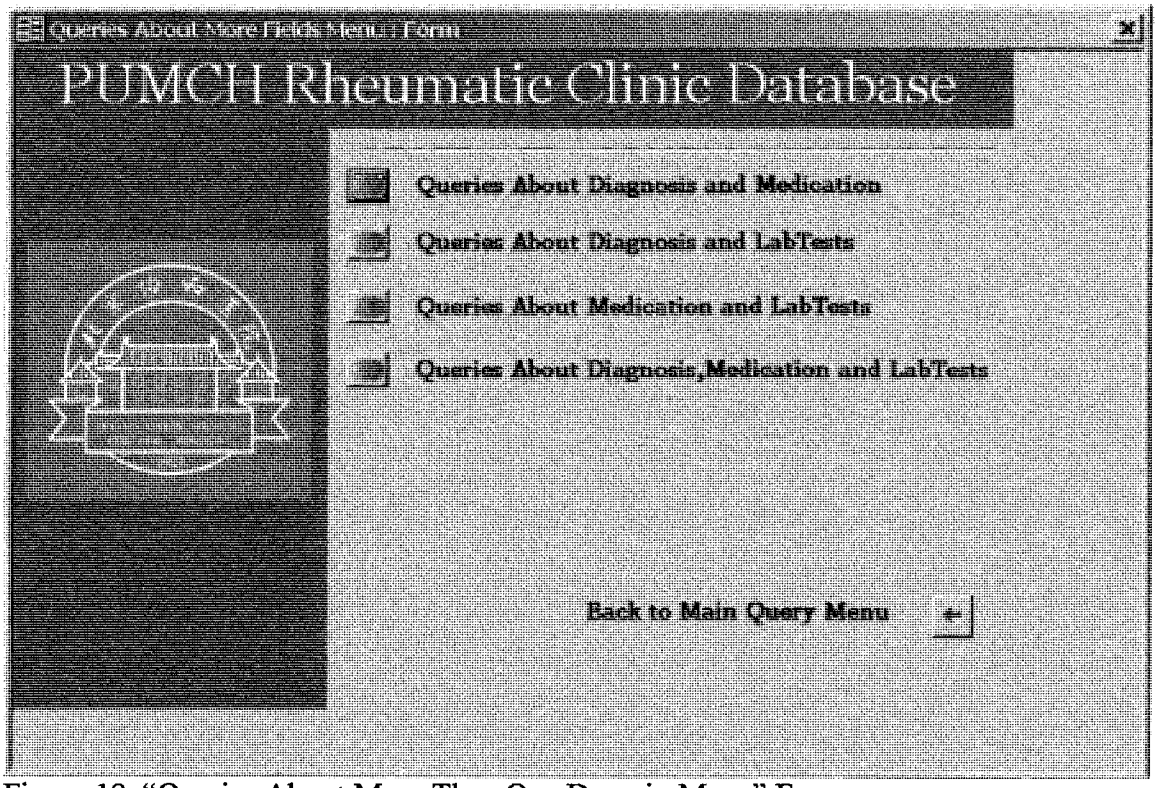


Figure 19. "Queries About More Than One Domain Menu" Form

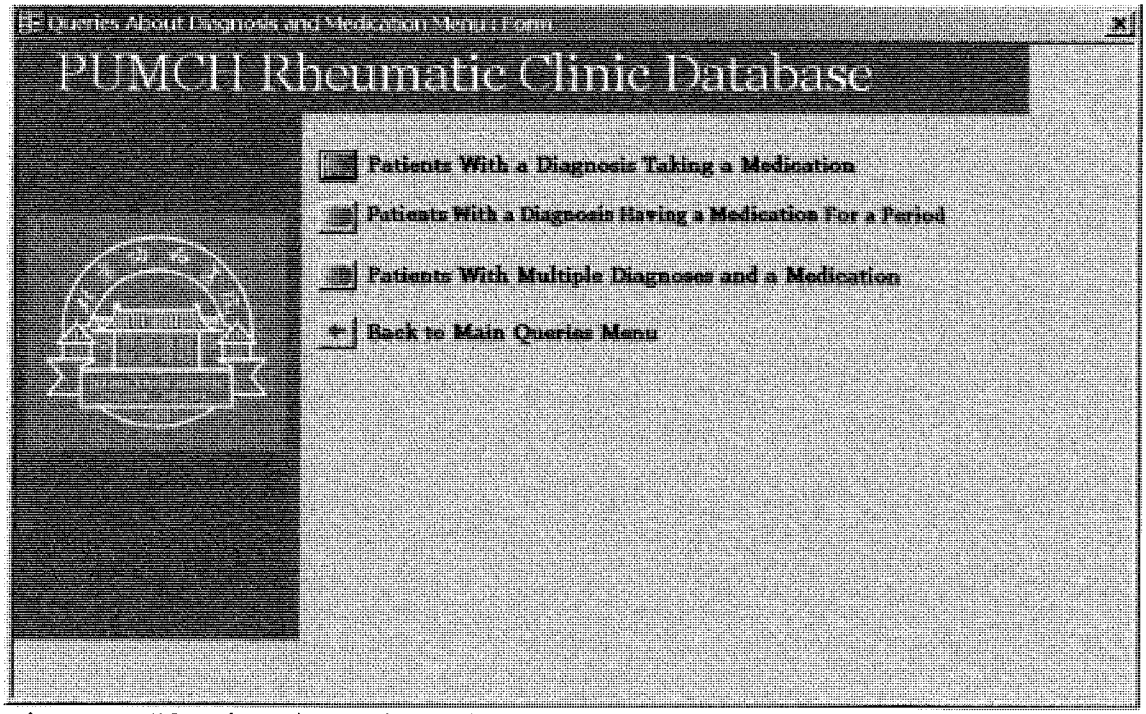


Figure 20. "Queries About Diagnosis and Medication Menu" Form

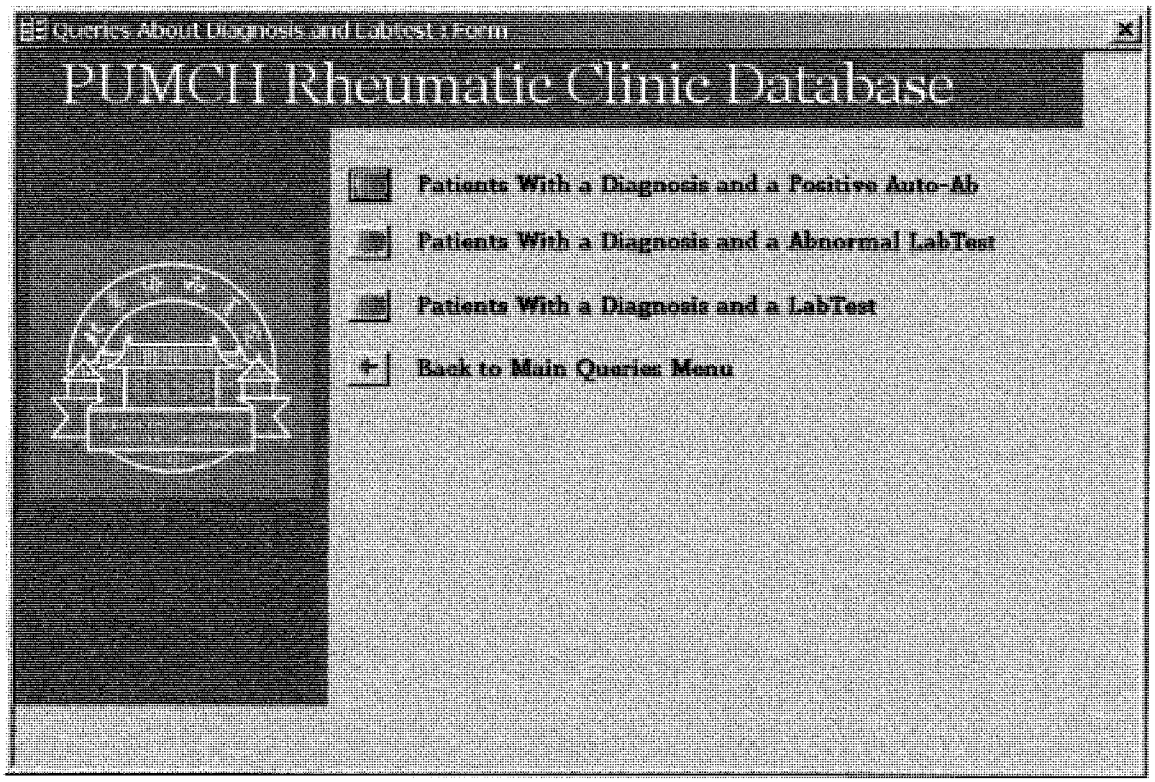


Figure 21. "Queries About Diagnosis and LabTest Menu" Form

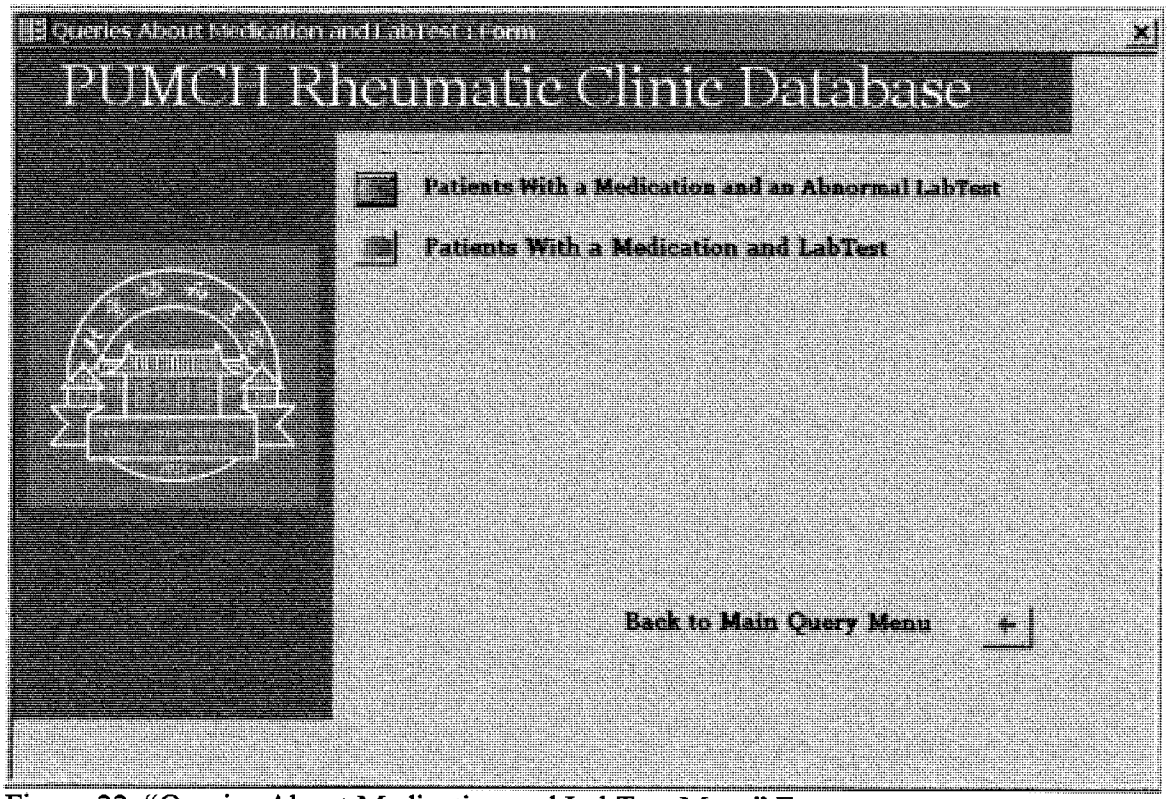


Figure 22. "Queries About Medication and LabTest Menu" Form

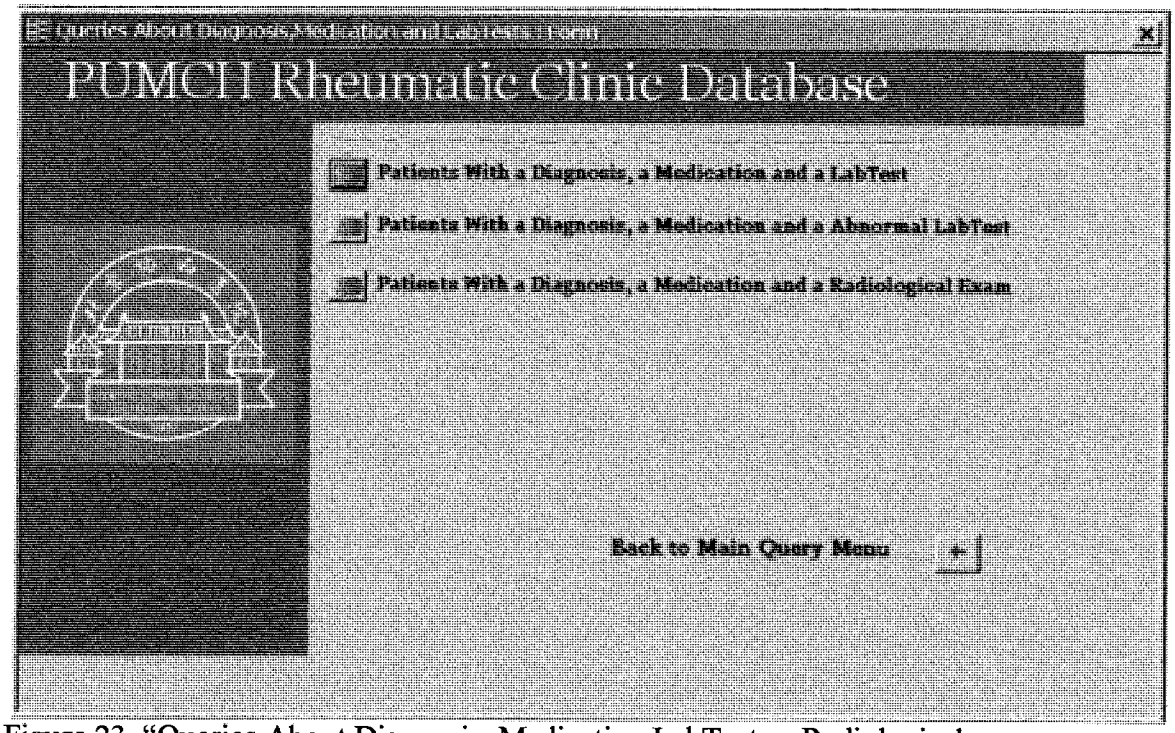


Figure 23. "Queries About Diagnosis, Medication,LabTest or Radiological Examination Menu" form

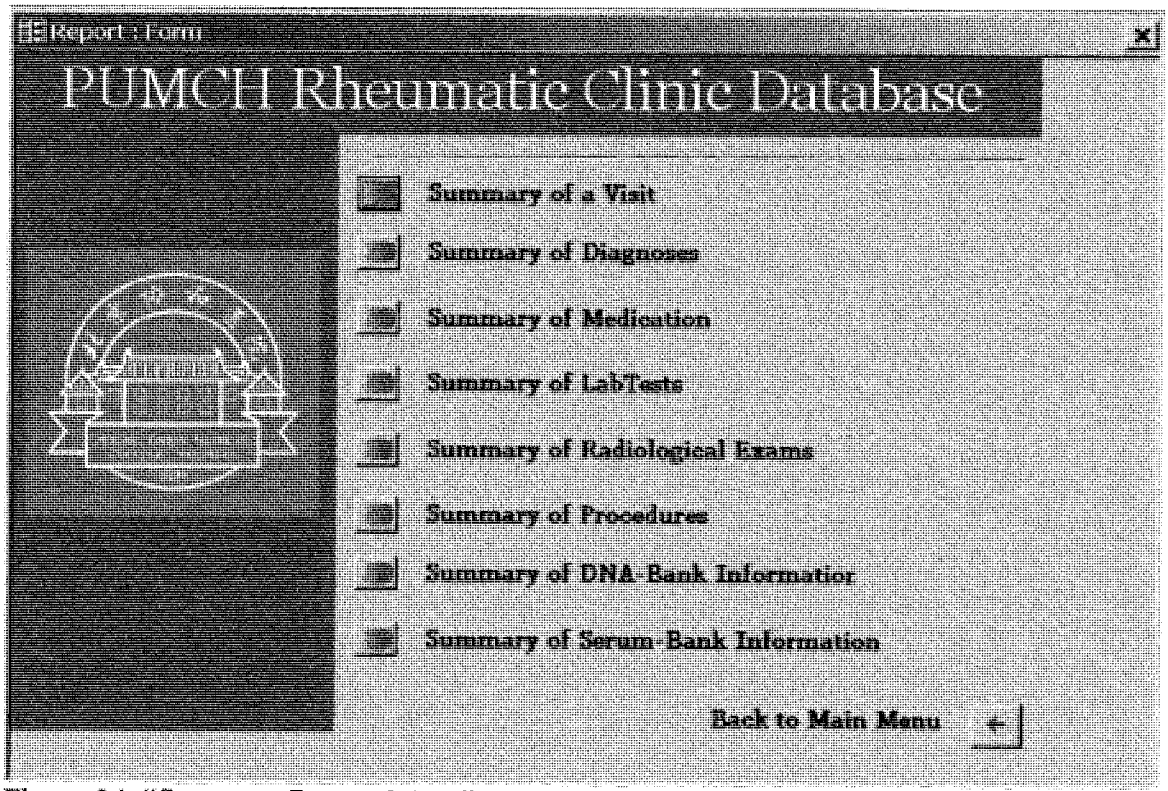


Figure 24. "Summary Report Menu" Form

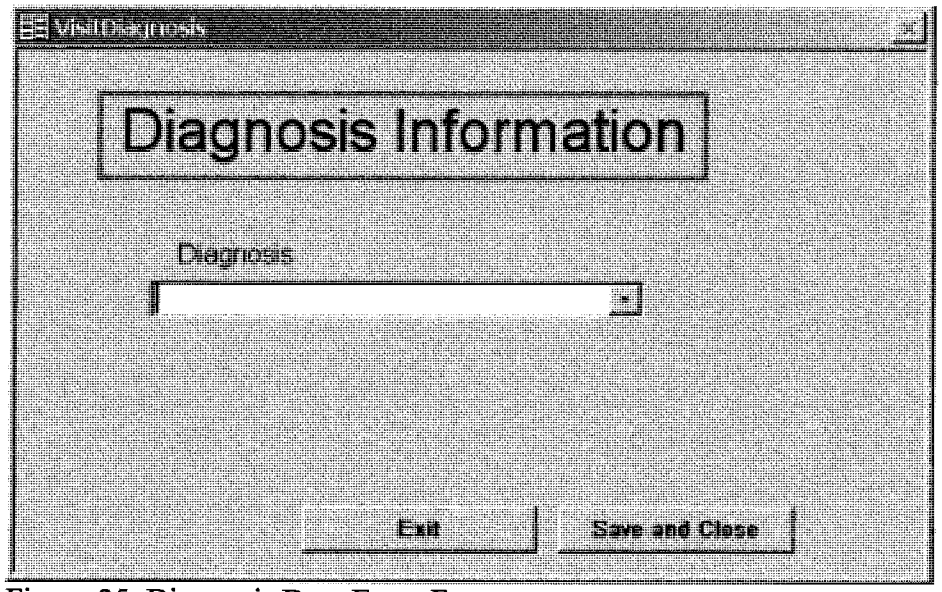


Figure 25. Diagnosis Data Entry Form

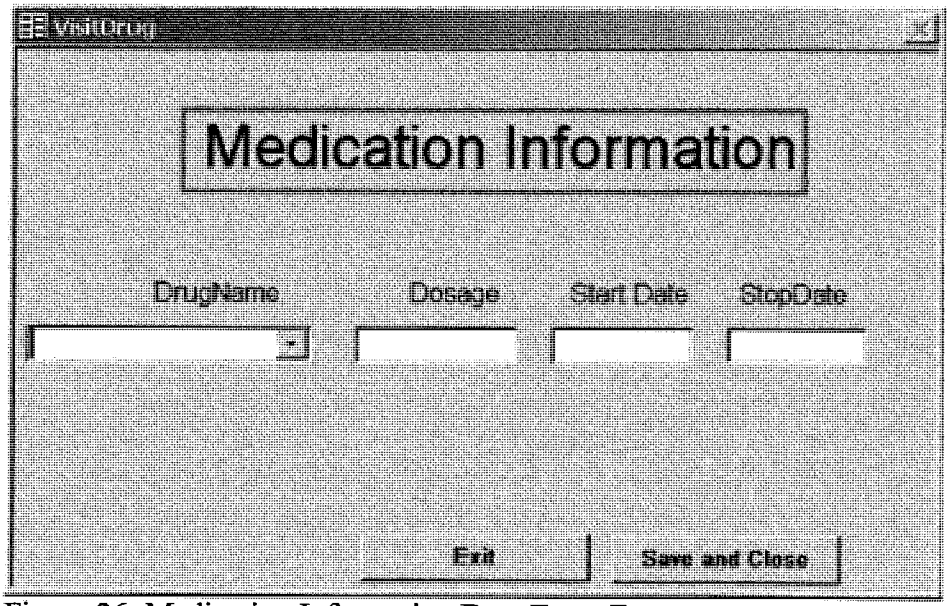


Figure 26. Medication Information Data Entry Form

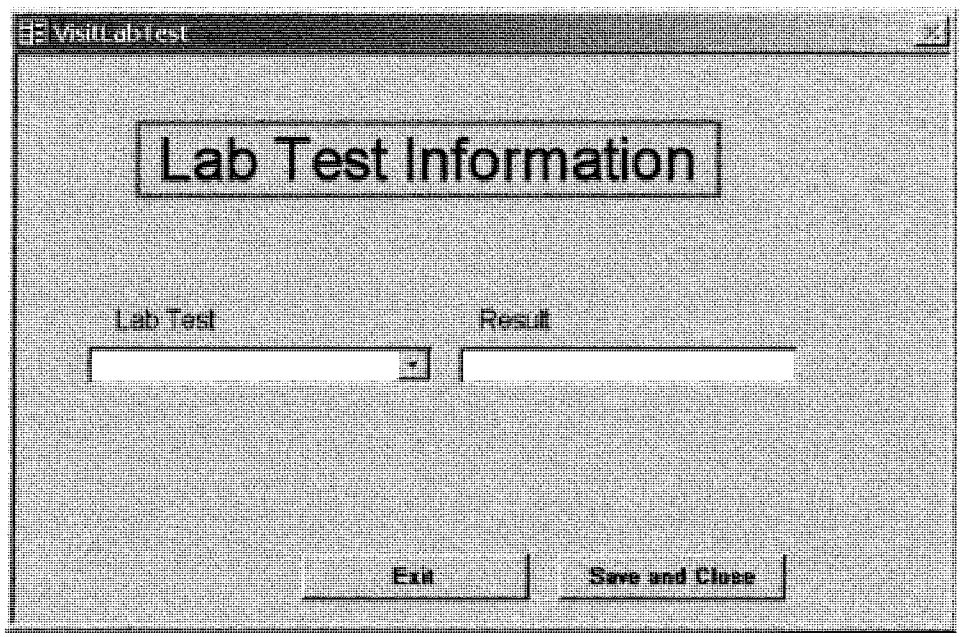


Figure 27. LabTest Data Entry Form

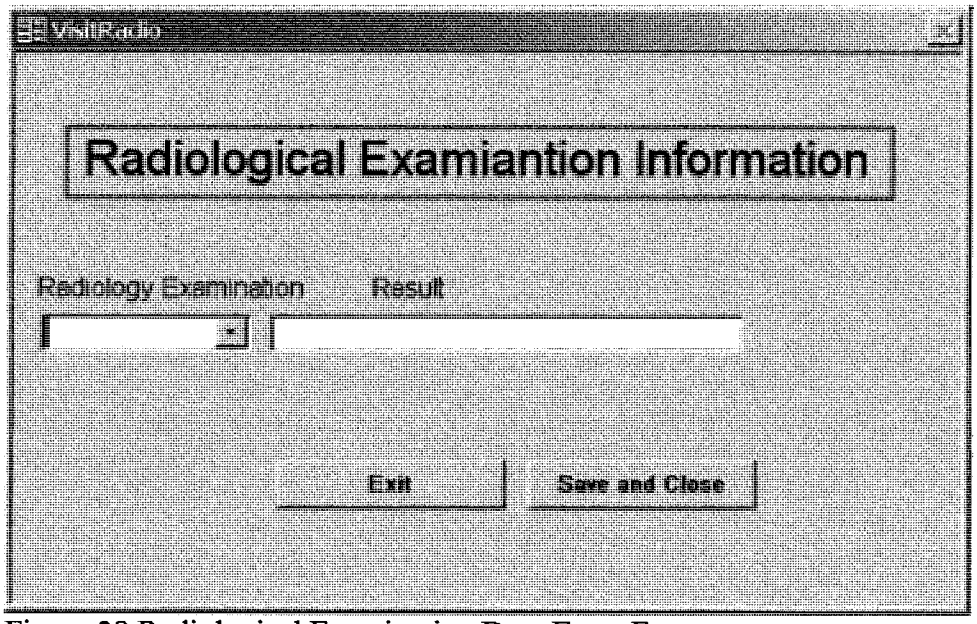


Figure 28.Radiological Examination Data Entry Form

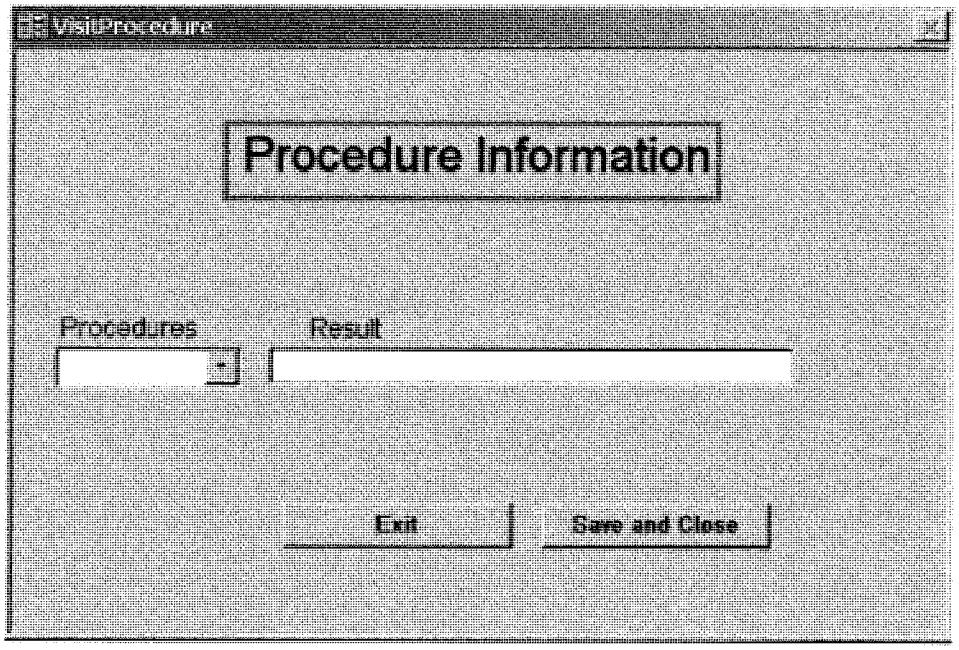


Figure 29. Procedure Information Data Entry Form

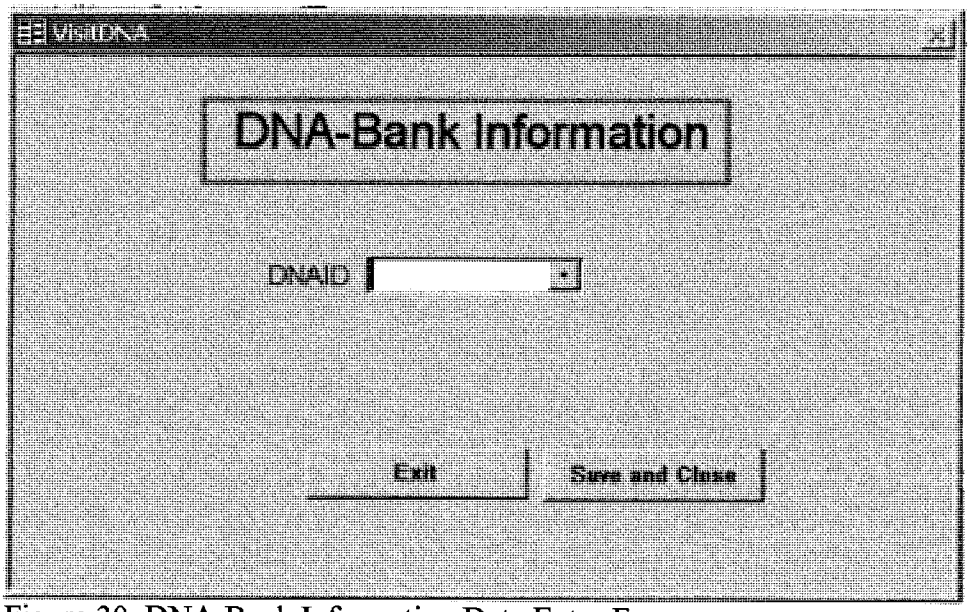


Figure 30. DNA-Bank Information Data Entry Form

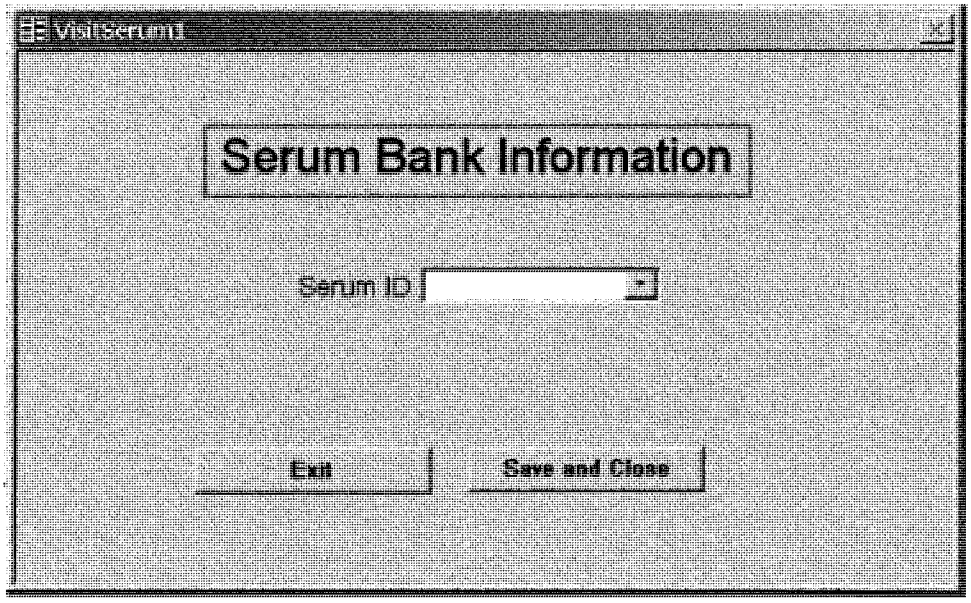
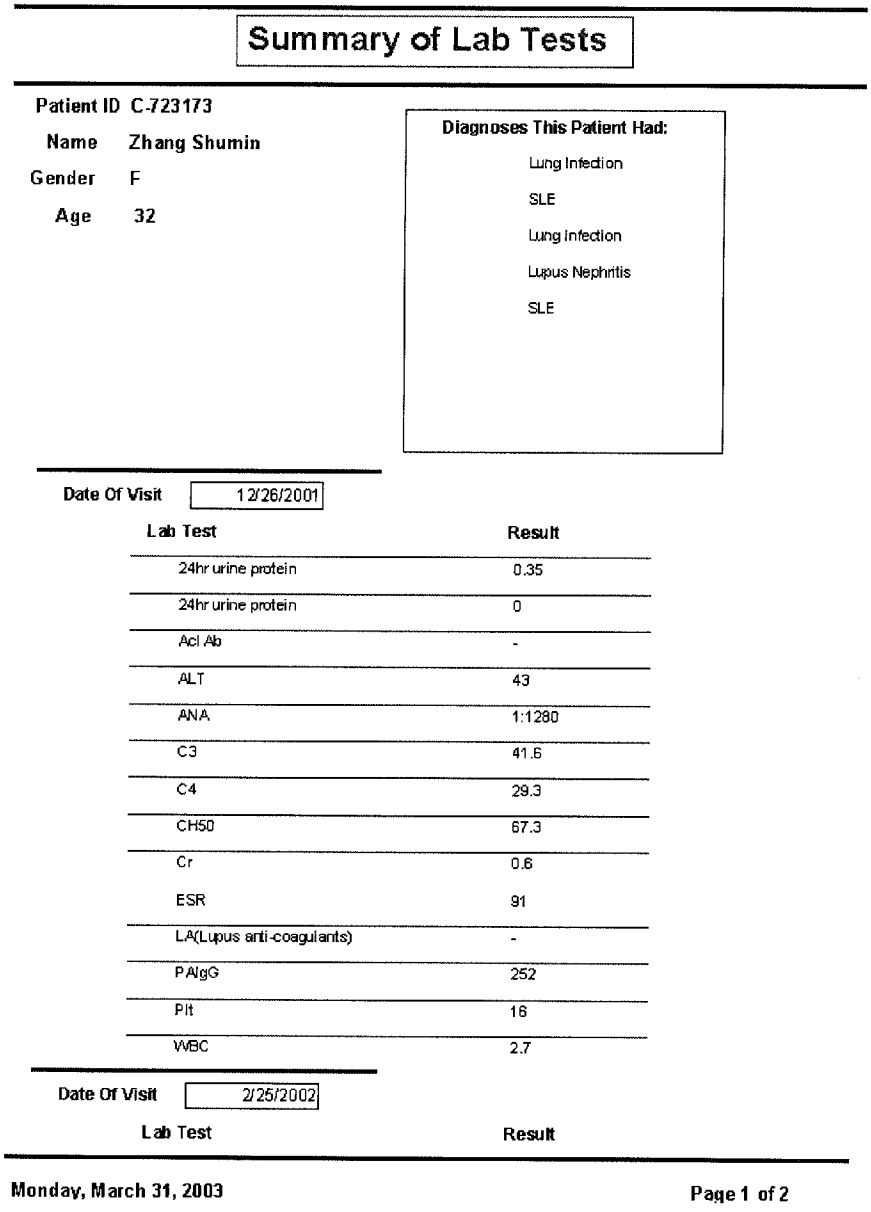


Figure 31. Serum-Bank Information Data Entry Form



Summary of Lab Tests

Patient ID C-723173
Name Zhang Shumin
Gender F
Age 32

Diagnoses This Patient Had:
Lung Infection
SLE
Lung Infection
Lupus Nephritis
SLE

Date Of Visit 12/26/2001

Lab Test	Result
24hr urine protein	0.35
24hr urine protein	0
Acl Ab	-
ALT	43
ANA	1:1280
C3	41.6
C4	29.3
CH50	67.3
Cr	0.6
ESR	91
LA(Lupus anti-coagulants)	-
PAIgG	252
Plt	16
WBC	2.7

Date Of Visit 2/25/2002

Lab Test	Result

Monday, March 31, 2003 **Page 1 of 2**

Figure 32. Report after modification (showing current diagnoses that patient has on it)

APPENDIX D

QUERIES DEVELOPED IN THIS DATABASE

A. Queries About One Medical Field:

1. Seven Queries about diagnosis were developed in this database(The italic part can be replaced to any diagnosis):

- Patients with a diagnosis
  For example: "How many patients were diagnosed with *Systemic Lupus Erythematosus*"?
- Patients with a diagnosis(detailed information that physicians can use to recruit patients into a clinical trail or to contact with patients if necessary)
  For example: "What are the general information of patients with the diagnosis of *Systemic Lupus Erythemtosus*"?
- Patients have a diagnosis in a certain period
  For example: "How many patients were diagnosed with *Systemic Lupus Erythematosus* during January 1, 2001 and December 31, 2001"?
- Patients with multiple diagnoses
  For example: "How many patients had been diagnosed with *Systemic Lupus Erythematosus* and *Systemic Sclerosis*"?
- Patients with a diagnosis in a certain age
  For example: "How many patients who were diagnosed with *Systemic Lupus Erythematosus* are older than 60"?
- Average age of patients with a diagnosis
  For example: "What is the average age of patients with a diagnosis of *Lupus Encephalopathy*?"
- Patients with a diagnosis for less than a year
  For example: "How many patients who were diagnosed with *Systemic Lupus Erythematosus* for less than 1 year"?

2. Five Queries about Medications were developed(The italic part can be replaced by any medication):

- Patients with a medication
  For example: "How many patients were treated with *Cyclosporine A*"?
- Patients with a medication for a time period
  For example: "How many patients have been treated with *Cyclosporine A* for more than 1 year"?

- Patients with a medication for less than *1 year*
  For example: "How many patients have been treated with *Cyclosporine A* for less than *1 year*"?
- Total dosage of a currently used medication
  For example: "What is the total dosage of *Cyclosporine A* that is currently used by patients"?
- Total dosage of a stopped medication
  For example: "What is the total dosage of *Cyclosporine A* that had been stopped"?

3. Two Queries about Laboratory Tests were developed in this database(the italic part can be replaced by any laboratory test):
   - Patients with an abnormal Laboratory test results
     For example: "How many patients had an abnormal *LA*(Lupus Anticoagulant) test"?
   - Patients with a positive antibody(generally auto-antibody)
     For example: "How many patients had a positive *Acl*"?

4. One Query about Radiological Examination (The italic part can be replaced by any radiological examination):
   - Patients with a radiological examination
     For example: "How many patients had a *Brain MRI*"?

5. Four Queries about Serum and DNA bank were developed. In fact, 2 queries are about diagnosis and Serum or DNA bank. As clinical researcher always asked patients diagnosis and their bank information at the same time, so I put these 4 queries under the list of queries about one medical field to facilitate searching. These 4 queries are(the italic part can be replaced by any diagnosis):
   - Patients in the DNA bank
     This query yields the overall information of patients who had their DNA samples in the DNA bank
   - Patients in the Serum Bank
     This query yields overall information of patients who had their serum samples in the Serum Bank
   - Patients with a diagnosis in the Serum Bank
     For example: "How many patients with *Systemic Lupus Erythematosus* had serum sample(s) in the Serum Bank"?
   - Patients with a diagnosis in DNA Bank

For example: "How many patients with *Systemic Lupus Erythematosus* had DNA sample(s) in DNA Bank"?

6. Two queries about "procedures" were developed. One of the queries is about diagnosis and procedure. I put these two queries together to facilitate searching about procedure information. These two queries are(the italic part can be replaced by any diagnosis and procedure):

- Patients with a procedure
  For example: "How many patients had a *Bone Marrow Aspiration*"?
- Patients with a diagnosis and a procedure
  For example: "How many patients with *Systemic Lupus Erythematosus* had *Bone Marrow Aspiration*"?

B. Queries About More Than One Medical Fields

1. Queries About Diagnosis and Medication:

Three queries in total were built. They are(the italic part can be replaced by any diagnosis and medication):

- Patients with a diagnosis taking a medication
  For example: "How many patients with *Lupus Nephritis* had been treated with *Cyclosporine A*"?
- Patients with a diagnosis taking a medication for a certain period
  For example: "How many *Lupus Nephritis* patients had been treated with *Cyclosporine A* for *1 year*"?
- Patients with multiple diagnoses taken a medication
  For example: "How many patients with both *Lupus Nephritis* and *Systemic Sclerosis* had been treated with *Cyclosporine A*"?

2. Queries About Diagnosis and Laboratory Tests

Three queries about diagnosis and laboratory tests were built. They are(the italic part can be replaced by any medication and laboratory test):

- Patients with a diagnosis and a positive auto-antibody
  For example: "How many patients with *Systemic Lupus Erythematosus* had a positive *Acl*"?
- Patients with a diagnosis and an abnormal laboratory test
  For example: "How many patients with *Systemic Lupus Erythematosus* had an abnormal *LA* test"?

- Patients with a diagnosis with a laboratory test
  For example: "How many patients with *Systemic Lupus Erythematosus* had a *LA* test"?

3. Queries About Medication and Laboratory Test

Two queries were built. They are(the italic part can be replaced by any medication and laboratory test):

- Patients with a medication and an abnormal laboratory test result
  For example: "How many patients treated with *Cyclosporine A* had an abnormal *ALT* result"?
- Patients with a medication and a laboratory test
  For example: "How many patients treated with *Cyclosporine A* had a *ALT* test"?

4. Queries About Diagnosis and Radiological Examination

One query about diagnosis and radiological examination was built. It is(the italic part can be replaced by any diagnosis and radiological examination):

- Patients with a diagnosis and a radiological examination
  For example: "How many patients with *Systemic Lupus Erythematosus* had a *hand X-ray*"?

5.Queries About Diagnosis, Medications and Laboratory Tests (or Radiological Examinations)

Three queries about diagnosis, medication and laboratory test or radiological examination were built. They are(the italic part can be replaced by any medication, diagnosis, laboratory test or radiological examination):

- Patients with a diagnosis, a medication and a laboratory test
  For example: "How many patients with *Systemic Lupus Erythematosus* who were treated with *Cyclosporine A* had an *ALT* test"?
- Patients with a diagnosis, a medication and an abnormal laboratory test result
  For example: "How many patients with *Systemic Lupus Erythematosus* who were treated with *Cyclosporine A* had an abnormal *ALT* test result"?
- Patients with a diagnosis, a medication and a radiological examination
  For example: "How many patients with *Polymositis* who were treated with *Methotrexate* had a *high resolution chest CT scan*"?