An Electronic Health Record Model for the Spatial Epidemiological Analysis of Clinical Data

Stelios Zimeras[1], Marianna Diomidous[2], Dimitris Zikos[2], John Mantas[2]
University of the Aegean, Department of Statistics and Actuarial-Financial Mathematics, Samos, Greece[1]
University of Athens, Department of Public Health, Athens, Greece[2]

SUMMARY

The main purpose of this work is to develop an Electronic Health Record (EHR), for the registration of clinical data, combined spatial Databases and GIS techniques. The main idea behind this work is the development of a real time efficient EHR environment in which clinical data could be analyzed and combined directly providing sections like statistical analysis, graphical presentation and GIS representation introducing disease monitoring. These monitoring results could be used as pilot directions for the spatial epidemiological analysis taking advantage geographical information for understanding the dynamics of health. Moreover the role of spatial information in the health sector in relation to the benefits of the GIS systems is illustrated combined with the spatial assumptions introducing a real-time health monitoring tools and geographic information systems. The EHR includes demographic data of the patients, followed by clinical data of the current and previous hospital admissions (like blood test data, medical actions, biological examinations, etc). The data can easily be extracted from the database for the implementation of statistical analysis with the use of the appropriate software. Demonstration of the system based on a short patient dataset was performed showing the benefits of the HER.

Keywords: Electronic Health Record Model

1. INTRODUCTION

Hospitals have a long road ahead to adoption of electronic health records. The EHR functions in which the greatest number of hospitals reported significant progress are: Order entry, Results management, Electronic health information/data capture, Administrative processes. By comparison, relatively few hospitals reported significant progress in clinical decision support health outcomes reporting (13%), and patient access (2%). Larger hospitals were further along in EHR adoption than were mid-sized or small hospitals and non rural hospitals were slightly further along than were rural hospitals. For these reason, it is crucial to develop an efficient, real-time EHR environment for monitoring and analyzing epidemiological changes, introducing healthcare policies.

Electronic health record systems hold the promise to address the two most crucial challenges to the healthcare system: controlling costs and improving quality. An electronic health record is a digital collection of a patient's medical history and could include items like diagnosed medical conditions, prescribed medications, vital signs, immunizations, lab results, and personnel characteristics like age and weight. The promises of EHRs are many: fewer adverse drug events, lower morbidity and mortality rates, seamless continuity of care, greater efficiencies, and lower costs. Electronic health records systems can also provide additional functionality, such as interactive alerts to clinicians, interactive flow sheets, and tailored order sets, all of which can not be done be done with paper-based systems.

Moreover geography plays a major role in understanding the dynamics of health, and the causes and spread of disease. The classic public health triad composed of man, agent/vehicle and environment emphasizes the importance of geographic location (environment or space where we live) in health and disease. Interactions within this triad can also change with time. Besides policy development, and provision and management of health services, public health practitioners have other important and related tasks including responding to health alerts and concerns, intersectional engagement, and community development initiatives. In all these tasks, they should try to incorporate searching and using best evidence in their everyday decision-making processes in order to minimize investment of efforts and funds in areas where there is solid evidence of no effect, or evidence of harm, or of poor cost-effectiveness. Evidence-based approaches can also highlight areas where the evidence may be less than reliable, requiring further assessment before expending large funds and efforts.

The main idea behind this work is the implementation of a real-time EHR environment, leading to epidemiological analysis combining geographical information for the evaluation of the healthcare sectors. Statistical analysis and graphical presentation tools have been included inside the system, for the analysis of the clinical data. Due to this geographical information, spatial databases could be used to support the proposed system. Main characteristic of these databases is the capability of managing large collections of relatively simple geometric objects (maps).

The main application driving research in spatial database systems are GIS. GIS are potentially powerful resources for community health for many reasons including their ability to integrate data from disparate sources to produce new information, and their inherent visualization (mapping) functions, which can promote creative problem solving and sound decisions with lasting, positive impacts on people's lives. For that reason, a GIS system has been develop to target resources for disease prevention by highlighting areas with significantly high rates, and to predict which areas might be at future risk and which may benefit most from future local population screening.

GIS provide the capability to perform two types of spatial analysis that could not be performed without GIS: finding areas of high disease incidence that can be labeled as statistically significant and worthy of further investigation, and examining the spatial relationship between disease incidence and information that is geo-referenced differently from the disease data.

Introduction to Electronic Health Record (EHR)

A carefully designed Health Information System which includes various fields related with the health of injured could be used as a tool for the extraction of statistical information regarding their health.

The use of information systems for this purpose is a practice that is being studied recently. The 2004 Utrecht study [1] combined the traditional epidemiological studies with the strength of the Electronic Health Record that is being kept in Primary Health Care Facilities. Another study was carried out in 2004 by [2], which concludes that the data files stored in electronic form are potentially available for process and analysis. [3] used Electronic Health Records, whose queries used statistical methods for the identification of cases of coronary disease. [4] studied the applicability of data extraction from an Electronic Health Record and concluded that the collection of data is feasible. In 1994, [5] used an Electronic Health Record to identify diabetic patients with or without coronary disease.

The utilization of a reformed Information System for Primary Health Care Units to improve organization of health data of patients is expected to assist Primary Health Care workers in their everyday practice. In addition, such a reformed Information System could be based not only on health data fields, but also on secondary information that may be correlated with specific health conditions.

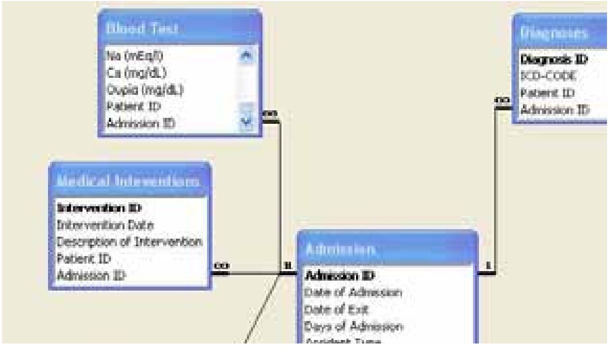

Figure 1: Schematic Presentation of the Database Relationships

Thus it can be used as a tool for the research and the fulfilment of epidemiological studies by extracting the data from the Information System. The development of a unique Electronic Health Record for the registration of patients injured by accidents aims to be used as a tool for the surveillance of various parameters concerning the specified population, as well as for the extraction of epidemiologic information from the database of the EHR.

2. METHODOLOGY

2.1. Database development

The database application that has been used for the development of the EHR build based on Microsoft Access 2007, which is considered as an all in one solution for Database and User Interface environment development. The fields included in the EHR are based on the handwritten patient data that is being collected in a Greek Island hospital.

The main table of the database contains the fields with demographic information, which are *Name, Surname, Age, Gender, Home Address,* and *Place of Birth*. The primary key of the table Demographics is the ***Patient_ID***. The table ***Demographics*** is connected with the table ***Admissions*** through a one-to-many relationship. This is based to the fact that one patient may have been admitted to the hospital more than once. The table ***Admissions*** includes relevant fields, which are the *Date of Admission, Date of Exit, Days of Stay, Accident Type, Place of Accident, Loss of Consciousness (Yes-No), Comments*. The table ***Admissions*** is connected with four additional tables, which are the *Medical Interventions, Blood Test, Diagnosis* and *Other Examinations* with one-to-many relationships. Each submission may require many blood tests, may include many diagnoses, more than one Medical Interventions and more than one *Other Examinations*. The table ***Medical Interventions*** includes the fields: *Date of Intervention, Description of Intervention*. The table ***Blood Test*** includes the fields *Date of Exam* as well as an additional 17 fields of relevant blood test results (i.e. *Ht, Blood Sugar, K, Na,* etc). The table ***Diagnoses*** include the field *Diagnoses*, which takes values from an additional table with the International Classification of Diseases ver.10 (ICD10). The normalization procedure of the database has been performed in a way that the schema of the database represents the actual procedure that takes place in the hospital and assures the integrity and validity of the

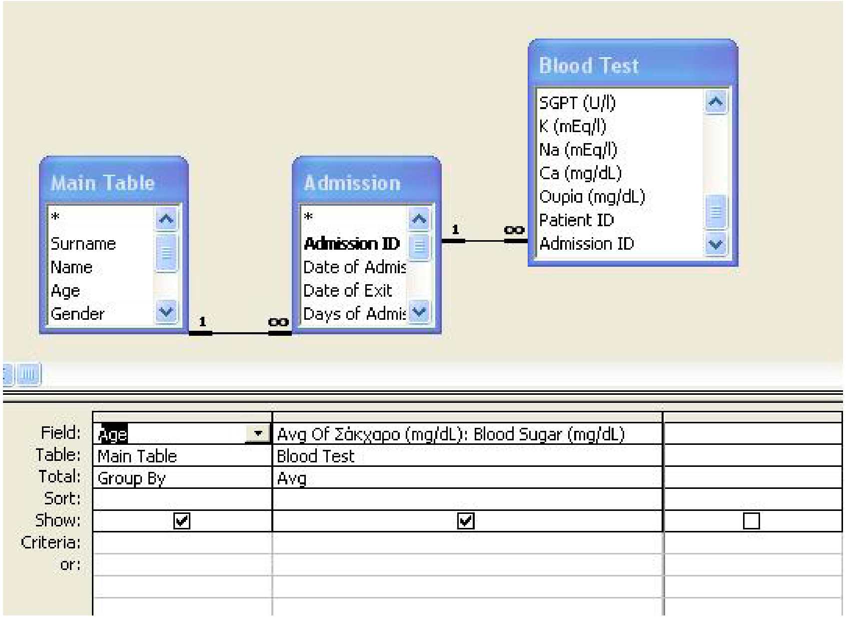

Figure 2: Query design: Blood Sugar level by age

collected data. Graphical illustration of the database relationships are presented in Figure 1.

The database collects health information as well as information of other characteristics that are not explicitly connected or might have correlation with health data fields. The collected data will be utilized for this purpose. Implementation of the system was applied based on 90 patients where for the most of them have been declared the types of accidents and for 34 cases the geographical position of the accident type.

The fields of interest that have been selected for further exploration consist of the following pairs:

Age-Days of Admission

- Diagnosis-Type of Accident,
- Blood test results-Age,
- Type of Accident-Days of Admission
- Type of Accident-Admission to Athens

The exploration of a possible correlation between the above parameters may give an answer to the following key questions:

- Does the age of an injured patient affect the days he/she stays in the hospital? Do specific types of accidents cause specific health problems?
- Does the patient age affect specific blood test values?
- Is there any correlation between specific types of accidents and admission days?
- Is the hospital unable to cure patients injured by specific types of accidents leading to the need for the admission to a larger hospital in city of Athens?

2.2. Query Development

A very important feature of the current HIS is the integration of queries for the extraction of specific information regarding the above parameters. The "correlation" queries are the basis for the creation of a graphical representation of selected parameters so that the user of the HIS can draw conclusions regarding the correlation between the selected factors that have already been described. The queries have been developed through the query design wizard of Microsoft Access. Each query refers to each of the correlations that are being explored. The appropriate query design techniques have been used to assure the validity of the query output. The query that is being used for the exploration of a possible correlation between Age and Blood Sugar Levels requires the participation of the field Age and the Average Value of the field *Blood Sugar Levels*. By executing the query, the output shows each value of the factor *Age* and the mean value of the *Blood Sugar Level* for each Age group (Figure 2).

The query "Days of Admission by Accident type" follows a similar procedure. The query contains two fields, *Days of Stay* and *Accident Type*, which are both included into the table *Admissions*. The Average Days of Admission are being calculated for each Accident Type (Figure 3).

The queries are being presented through an interactive user interface with the use of dynamic diagrams that show the results of each query. The diagrams are being updated automatically after each modification of the patient data. The diagrams appear in Visual Basic forms, which can be accessed through the unified Electronic

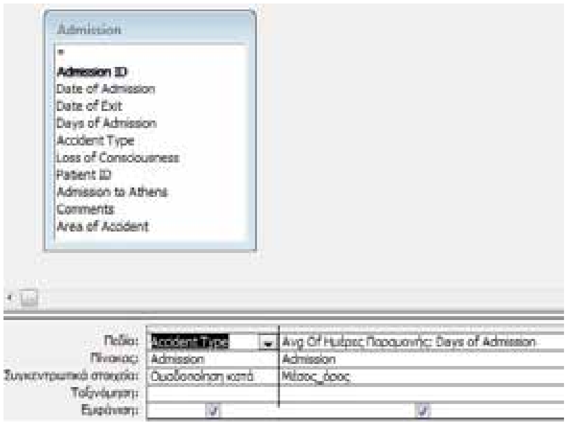

Figure 3: Query Design: Days of Stay by Accident Type

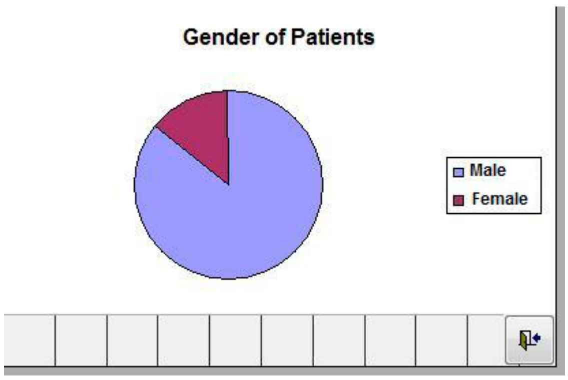

Figure 4: Graphical presentation of Gender Distribution

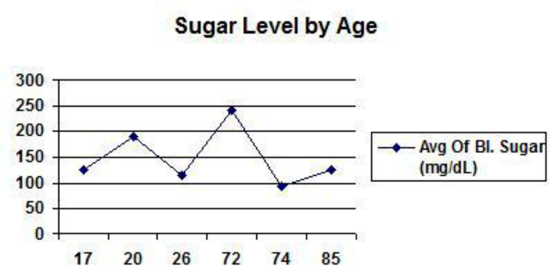

Figure 5: Blood Sugar Level by Age

Health Record environment. The graphical representation can be used by the healthcare professional as a tool for the surveillance of main aspects of the patients, as well as by the scientists to draw initial conclusions about a possible correlation and proceed to further data analysis by using more advanced statistical methods. Apart from the correlation diagrams, the EHR incorporates the graphical representation of demographic data of the selected population, such as age distribution and gender information (Figures 4, 5).

2.3. Development of the User Interface

The development of a friendly User Interface for the easy navigation though the information system and its options is vital. The information system introduces the user to its options through a startup screen, which lets the user choose the appropriate function (Figure 6). The functions provided are: the Electronic Health Record, the graphical presentation of the collected data through a series of forms which provide access to the diagrams that were described above, and the dynamic graphical presentation of certain health and demographic parameters in a map of the island of Samos, which depicts the values of these parameters in the various areas of the island.

Electronic health (e-Health) has become a very important area of focus and activity in multiple domains, such as health promotion, health care and maintenance, public health, medical science, health service, data management, image processing, telecommunication, wireless network, and operational research. The EHR, as defined by the Medical Records Institute, is electronically maintained information about an individual's lifetime health condition and health care. The EHR is expected to replace paper medical record(s) as the primary source of information for health care, and still comply with all clinical, legal and administrative requirements. Development of fully functional interoperable EHR system remains a major challenge. Recent research has proposed prototype service-oriented architecture (SOA) models for EHR in various contexts including clinical decision support, collaborative medical (mammogram) image analysis, and health clinic setting.

The Healthcare Information and Management Systems Society (HIMSS) define the EHR more comprehensively as a "longitudinal electronic record of patient health information generated by one or more encounters in any care delivery setting. Included in this information are patient demographics, progress notes, problems, medications, vital signs, past medical history, immunizations, laboratory data and radiology reports. The EHR automates and streamlines the clinician's workflow. The EHR has the ability to generate a complete record of a clinical patient encounter, as well as supporting other care-related activities directly or indirectly via interface–including evidence-based decision support, quality management, and outcomes reporting". A comprehensive EHR at the point of care could be created by aggregating and sharing data among all sites at which patient receives care, as well as with data supplied by the patient. To share and use data from multiple institutions, data must be built upon common words (data elements and terminology), structures, and organizations (interoperability) [13-16].

One of the greatest incentives to adopting EHRs will

Figure 6: Startup screen of the Information System

be through reaching a critical mass of information sharing investors in health care information technology are by and large dealing with internal information systems unable to interact with outside systems

The electronic record must include eight essential attributes [17]:

- Provide secure, reliable and real-time access to patient health information where and when it is needed to support care.
- Capture and manage episodic and longitudinal electronic health record information.
- Function as clinicians' primary information resource during the provision of patient care.
- Assist with planning and delivering evidence-based care to individuals and groups of patients.
- Support continuous quality improvement, utilization review, risk management and performance management.
- Capture information necessary for reimbursement.
- Provide longitudinal, appropriately masked information to support clinical research, public health reporting and population health initiatives.
- Support clinical trials.

The main benefits of EHR are:

- Improves patient care by allowing clinical decisions to be made using up-to-date documents
- Reduces errors
- Eliminates lost patient charts
- Strengthens both internal and external communication

To gain a more accurate appreciation of the EHR's

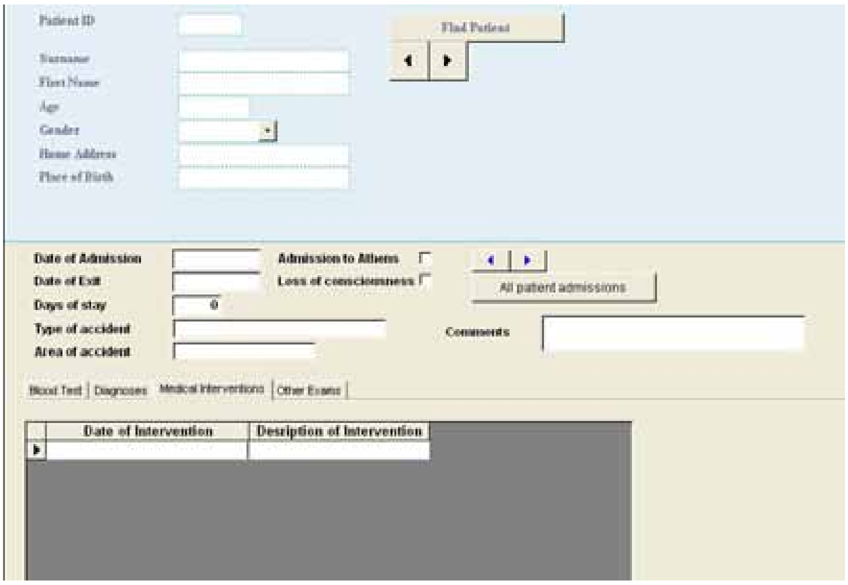

Figure 7: The Electronic Health Record Form

complexity and breadth of information, one must recognize the wide range of health information sources. Each time an individual visits a healthcare provider, data are generated.

The Electronic health record incorporates all given options through a single screen, which provides access to all patient data (Figure 7). The form is split into two parts. The above area contains the demographic data of the patient (Patient ID, Surname, First Name, Age, Gender, Home Address, and Place of Birth). This area includes navigation buttons as well as a search button for easy access to the required patient data by name or Patient ID. The bottom area includes the admission data of the patient that has been selected (Date of admission, Date of Exit, Days of Stay to the hospital, Type of Accident, Area of Accident, Admission to Athens, Loss of Consciousness). For each admission the user can gain access to a series of parameters which refer to the selected admission. The parameters are organized through four Tab Controls (cards), which are the *Blood test card,* the *Diagnoses card,* the *Medical Intervention card* and the *Other Examinations* card. The structure of the Electronic Health Record form clearly simulates the actual procedure that takes place in the hospital as far as the injured patients are concerned. A patient may have more than one admission and each of this may include more than one action (Diagnoses, Blood Tests etc). The selected color schema makes the discrimination of the Demographic Data and the Admission Data clear. Additional features such as the selection of the Gender through a list of reselected values (male-female) as well as the use of checkboxes when required, aim to create a more user-friendly environment.

The second section of the information system is the Graphical Presentation of demographic information and the correlation graphs of the parameters that have been selected for exploration. The user selects the fields that he wants to explore through a series of navigation forms. The user initially selects the first of the factors for investigation and a new form appears for the selection of the second factor. The selection of the second factor is followed by a final form where the user selects the final details of the requested information. i.e. the user who wants to see the diagram of the Examination Results by Age is requested to select the specific Blood Test he wishes (i.e. Blood Sugar). The step is followed by the appearance of the diagram (Figure 8). The manipulation of the collected information includes descriptive statistics, non-parametric statistical tests (Peason chi-square, spearman). These tests have been incorporated into the system through an integral user interface.

Figure 8: Blood Sugar Level by age: navigation and output

An additional feature of the section is the integration of 2x2 tables. A sample of three dynamically update 2x2 tables have been developed. The scope for further development of this feature with the addition of more correlation tables is very important (Figure 9). The correlation of the specified variables is performed in database-level by the transformation of the various correlation test function into SQL code.

The third section of the information system is the presentation of demographic and health information of the patient population through a pseudo-GIS interactive map of the Samos Island. **Geographical Information System (GIS)** is a system of hardware, software, data, and people to collect, store, analyze, manipulate, model, visualize, and disseminating information about specific areas [6-9]. Geographical Information Systems (GIS) can be used to undertake spatial analysis, quantitative representation and modeling of spatial data; making them fit for population analyses which uses attribute data about humans in order to get the size of population, its composition, characteristics, and how they are and will be spatially distributed. Data are normally collected at the point level (individuals) but it is always available spatial entities (e.g. administrative units) to allow tabulations according to various data characteristics and demographic analysis using statistical techniques. GIS, with its spatially referenced data and spatial analysis tools can provide solutions to display 2D or 3D spatial data [10-12].

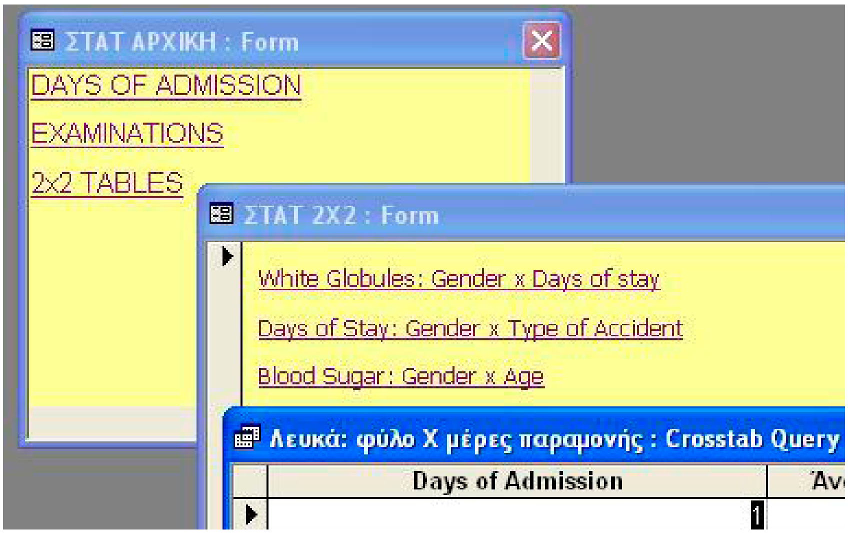

Figure 9: Statistical analysis based on 2x2 Tables

The user is given the opportunity to click on a specific geographical area of the map and acquire information regarding various parameters of the accidents according to the selected area. The unique feature will enable the healthcare professional to have an overview of the geographical distribution of health and health related parameters and study a possible correlation between these parameters and the selected area. The interactive geographical map that refers to the type of accident may show a possible correlation between the area of the accident and a specific type of accident that appears in high rates in the specific area, due to certain characteristics of the area (i.e. an increased rate of car accidents in a specific area may pinpoint the need for road work in the area). The interactive maps have been created by using links behind each map area. Each link enables the execution of a specific query that refers to the selected area (Figure 10).

3. EVALUATION

The statistical subsections of the system have been evaluated by testing the validity of the various tests. In order to test the SQL code, a comparative analysis of data has been performed, using a commercial statistical package (SPSS). The Information System (IS) datasets have been copied to SPSS, and engines were performed the same statistical tests. The output is expected to be the same. In case of inconsistency of the IS output, the database structure as well as the specified SQL commands was tested for logical, structural and syntax errors. This methodology is part of the overall IS evaluation, which includes post study usability questionnaires with usage scenarios.

4. CONCLUSIONS

The main purpose of this work is to develop an effective real-time Electronic Health Record (EHR), for the storing and analyzing of clinical data, combined spatial Databases and GIS techniques. Combination of these functions graphical presentation could be used for monitoring diseases mapping as well as healthcare management policies. These monitoring results could be used as pilot directions for the spatial epidemiological analysis taking advantage geographical information for understanding the dynamics of health.

Statistical analysis and graphical presentation tools (based on spatial Databases) have been included inside the system, for the analysis of the clinical data. Main characteristic of these databases is the capability of managing large collections of relatively simple geometric objects (maps). Finally, a pseudo-GIS system has been develop to target resources for disease to predict which areas might be at future risk

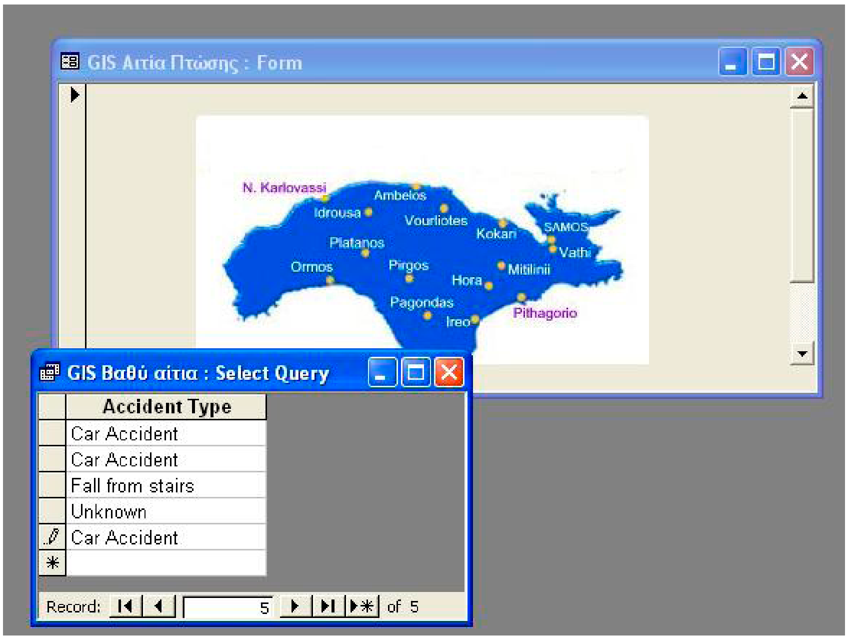

Figure 10: GIS: Accident type by Area of Accident

and which may benefit most from future local population screening.

REFERENCES

1. Grobee D, Hoes A, Verheij T, Schrijvers A, van Ameijden E, Numans M. The Utrecht Health Project: Optimization of routine healthcare data for research. European Journal of Epidemiology. 2005; 20: 285-7.
2. Majeed A. Sources, uses, strengths and limitations of data collected in primary care in England. Health Stat Q. 2004;21: 5-14.
3. Thiru K, Donnan P, Sullivan F. A validated logistic regression model to identify coronary heart disease patients within primary care databases in the United Kingdom. AMIA AnnuSymp Proc. 2003; 1030.
4. Mansson J, Nilsson G, Bjorkelund C, Strender LE. Collection and retrieval of structured clinical data from electronic patient records in general practice. A first-phase study to create a health care database for research and quality assessment. Scandinavian Journal of Primary Health Care. 2004; 22(1): 6-10.
5. Fuke D, Hunt J, Siemenczuk J, Estoup M, Carroll M, Payne N, Touchette D. Cholesterol management of patients with diabetes in a primary care practice - based research network. Am J Manag Care. 2004; 10 (2 Pt 2): 130-6.
6. Harris, B. Beyond Geographic Information Systems: Computers and the Planning Professional, Journal of the American Planning Association. 1989;55:85-90.
7. Harris B, Batty M. Location Models, Geographic Information and Planning Support Systems. Journal of Planning Education and Research 1993;12:184-98.
8. Klosterman R E. Planning Support Systems: A New Perspective on Computer-Aided Planning. Journal of Planning Education and Research. 1997;17(1):45-54.
9. Klosterman R E. The What if? collaborative planning support system. Environment and Planning. 1999; B26: 393-408.
10. Chou Y. Exploring Spatial Analysis in Geographic Information Systems. Santa Fe, OnWord Press, 1997.
11. DeMers N M. Fundamentals of Geographic information system, Wiley and sons, New York, USA, 2000.
12. Tomlin C D. Geographic Information Systems and Cartographic Modeling. New Jersey: Prentice Hill, 1990.
13. Daskalaki A, Lazakidou A, Philipp C, Jacob C, Berlien H P. Introducing electronic health record into laser medicine. Med. Inform. 2001; 5: 85-6.
14. Mantas J. Electronic health record. In J. Mantas & A. Hasman (Eds.), Textbook in health informatics: A nursing perspective (pp. 250-257). Amsterdam: IOS Press, 2002.
15. Shortliffe E. The evolution of electronic medical records. Academic Medicine, 1999;74(4), 414-9.
16. Wang S, Middleton B, Prosser L, Bardon C, Spurr C, Carchidi P, et al. Academic publication proves that EMRs are cost effective: A cost-benefit analysis of electronic medical records in primary care. American Journal of Medicine. 2003;114(5):397-403.
17. The Healthcare Information and Management Systems Society, http://www.himss.org.

Corresponding author: Stelios Zimeras. University of the Aegean, Department of Statistics and Actuarial-Financial Mathematics, 83200 Karlovassi, Samos, {zimste@aegean.gr}