COLLABORATIVE IT PLATFORM FOR RARE DISEASES

Dragan Jankovic, Tatjana Stankovic
University of Nis, Faculty of Electronic Engineering, Aleksandra Medvedeva 14, 18000 Nis, Serbia

Branimir Todorovic
University of Nis, Faculty of Science and Mathematics, Visegradska 33, 18000 Nis, Serbia

Keywords: Rare diseases, Collaborative IT platform, Web services, Data mining, OLAP, Medical data analyses, Electronic healthcare record.

Abstract: For a long time the rare diseases have not been in the "focus" of pharmaceutical companies and research because of potentially lower wages and the fact that very few institutions have a representative set of data necessary for quality research. Unfortunately, patients suffering from rare diseases are left in the margins of many societies, their drugs are usually not on the “positive lists” of insurance organisations and their price is extremely high. The number of rare diseases is between 6000 and 8000 and the estimated number of cases is about 5%, i.e. about 250 million. This paper presents Collaborative IT platform model for rare diseases by reviewing four important aspects: creating a national register of people suffering from rare diseases that can potentially grow into an international; establishment of a central repository for rare diseases with a collections of medical data characteristic for rare diseases, with modern data analysis tools in order to create better conditions for scientific research in the field of rare diseases, where some tools would be oriented to help doctors to more easily and with less cost came to proper diagnosis; improving living conditions and treatment of patients by forming a set of virtual patient’s associations to exchange experiences and find useful information; to create conditions for better education of medical workers and patients. The proposed platform is the subject of the project that we apply to the call for proposals of the Ministry of Science and Technological Development (MSc&TD) in the Republic of Serbia for a period 2011 - 2014.

1 INTRODUCTION

EHR (Electronic Healthcare Records) has had great public attention lately all over the world (Kukafka, 2007). As the amount of collected electronic medical data increases everywhere, the health-care services and supporting industry are making efforts to identify better ways to use this data for patients care (Ford, Menachemi, Phillips, 2006). Ideally, data is collected in a real time, can support point-of-care clinical decisions, and, by providing instantaneous quality metrics, can create the opportunities to improve clinical practice as the patient is being cared for (Michael, Holl, Badawi, Riker, Silfen, 2010).

However, all segments of the health-care industry are plagued by many challenges that have made it a latecomer to business intelligence and data mining technology (Wickramasinghe, Schaffer, 2006). For example, the adoption of electronic medical records is delayed in many countries, integration between different medical information systems (MIS) is poor, and there is a lack of uniform technical standards. There is a poor interoperability between complex medical devices and MIS. Until basic technical infrastructure and well designed clinical applications are implemented through the health-care system, data aggregation and interpretation cannot effectively progress (DesRoches, 2008).

Unfortunately, all these facts affect the area of rare diseases tracking even more than other healthcare related processes. Rare diseases occur with significantly less frequency than common diseases. But, their frequency should not reduce the professional attention that is given both to diseases and to patients suffering from them. Unfortunately, practice often shows just the opposite: patients are

Jankovic D., Stankovic T. and Todorovic B..
COLLABORATIVE IT PLATFORM FOR RARE DISEASES.
DOI: 10.5220/0003171003090314
In *Proceedings of the International Conference on Health Informatics* (HEALTHINF-2011), pages 309-314
ISBN: 978-989-8425-34-8

usually marginalized by the society. The number of medications used for rare diseases treatment is relatively small, many of them are not in the list of medicine that government pays for, and their price is often even ten times higher than the price of medications for common illnesses. This situation may be justified by the fact that the interest of researchers and pharmaceutical companies especially in these diseases and these drugs was minor, when compared to the interest in common diseases. Rare diseases are too specific to study, to find the appropriate causes of disease (most of rare diseases are genetic), treatment methods, and the appropriate preparations. Researching difficulties are certainly the result of a small amount of relevant data in this area, in research centres separately. The only goal of this paper is to emphasize the problem of patients suffering from rare diseases, and to propose one way to help those patients, their doctors, and the researchers interested in the area of rare diseases. Our suggested solution introduces one model of Collaborative IT (CIT) platform for rare diseases (Fig. 1).

The proposed platform is the subject of the project we apply to the call for proposals of the MSc&TD in Serbia. Another purpose of this paper is to find potential partners for an international project with the same subject and goals and a wider set of users from different countries.

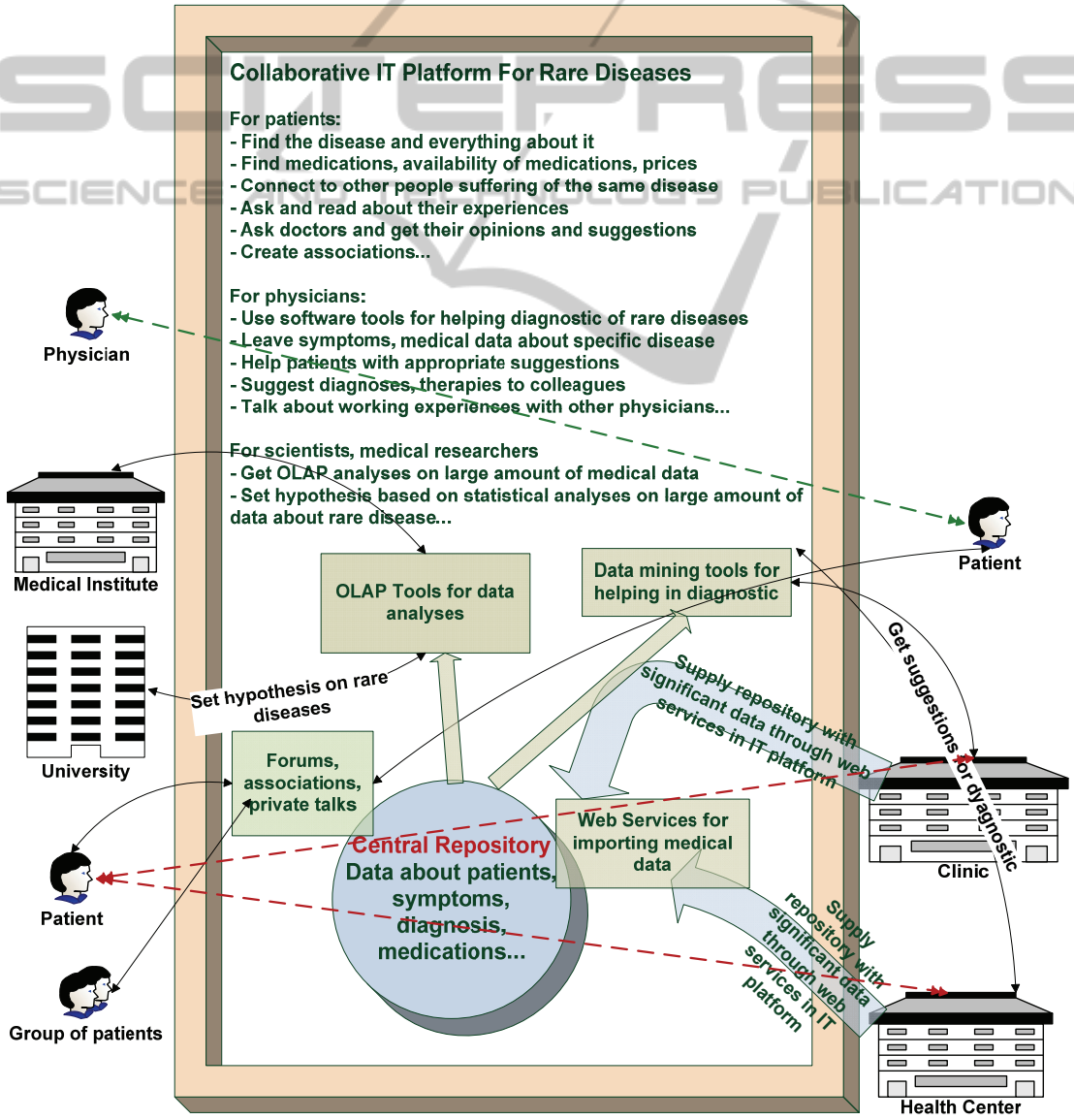

Figure 1: CIT platform for rare diseases illustration.

The existence of a project DILS (Delivery of Improved Local Services) is very important for the success of our project (http://www.dils.gov.rs). DILS main objectives are: to increase the capacity of institutional actors and beneficiaries in order to improve access and the efficiency, equity and quality of local delivery of health, education and social protection services (see http://web.worldbank.org/external/projects/main?pagePK=64312881&piPK=64302848&theSitePK=40941&Projectid=P096823). One of the DILS components is establishment of the information systems that connect local service providers (health centres, schools, centres for social work and non-governmental organizations) with the relevant ministry and allow efficient and transparent financing of services, service delivery and inter-agency information sharing. 157 health centres in Serbia will be provided by computer equipment and relevant MIS. Therefore, our platform for rare diseases will have sufficient IT base for successful work.

2 RESEARCH SIGNIFICANCE

This research will represent continuation of our previous work (http://medisc.elfak.edu.rs, Rajković, Janković, Tošić, 2009, Rajković, Janković, Stanković, 2009) on development of different MIS. Working on both ambulatory and clinical MIS, we have offered many data retrieving forms to medical professionals as well as connection to laboratory devices, different reporting and data analyzing tools. Our systems are primarily oriented to fit the needs of Serbian public healthcare, but designed configurable so they could be applied anywhere. Mentioned solutions are in process of adoption within medical facilities in Southern and Eastern Serbia. Population of this region is about 1.7 million people, which ensures representative population for data retrieval. Since we have realized that rare diseases had more complex medication process we have decided to start research in this area, having in mind wider sociological significance.

Research and activities that took place during the modelling of the platform are important from several aspects: scientific research, national, social, and educational.

2.1 Scientific Research Aspect

Our first goal is to establish central repository for rare diseases that would be, in the beginning, used in the Republic of Serbia. The establishment and operation of a central repository will provide, in a relatively short time, critical amount of data necessary for quality research to a number of researchers and research centres dealing in the area of rare diseases. Without the existence of a central repository, with variety of relevant medical information about patients suffering from the rare diseases, research is performed on the small amount of data from real patients, in separated centres, which can potentially reduce the quality and effectiveness of research. Initially, central repository will import data retrieved by previously developed medical information systems that are in active use in Nis region. Next step will be defining data interchange standards (under the authority of the Serbian Ministry of Health and National health insurance) to allow other medical information systems to share data with central repository. Also, legal point of view must be included to ensure that data exchange process will not intrude patients' privacy. As a side repository effect, the complete list of patients suffering from rare diseases (that does not exist at the moment in Serbia) will be formed.

A very important point of the research aspect is a novelty in the approach to data analyses. Specifically, power software tools will be developed for analysis of data entered by patients themselves. Similar approach is mentioned in (Michael, Holl, Badawi, Riker, Silfen, 2010), but not against the wide spectrum of medical data entered by patients.

2.2 National Aspects

By establishing a rare diseases central repository, we will create a suitable location that offers different kinds and pieces of information useful for patients. Data such is common symptoms of diseases, specific symptoms, medical and general advices, lists of medicaments, trends in medical practice, geographical distribution, and the success rates of different treatment under different circumstances will be maintained by applications based on the repository. This will bring national (even international) importance to a repository. Such a repository should provide access to all researchers in this field, as well as clinicians who participate in the project, to store the data they have on their clinics. One goal is to enable all doctors to store the relevant data in the central repository, whenever they detect rare diagnosis.

Another crucial national aspect is bringing systemised set of data directly to the patients, in their own language. There are many different

sources about rare diseases on the Internet, but there are just few South Slavic languages sites, having pure information about rare diseases.

The existence of repository will allow appropriate planning of the activity, primarily in the Ministry of Health and National health insurance organisation, and then in the medical institutions. That way rare diseases data repository should become a significant referent centre, which should be recognizable in the region, and we hope even in Europe.

2.3 Social Aspects

Serbian government has made significant efforts during the last decade to educate people in the health issues area, but those suffering of rare diseases are still treated, unfortunately, as marginalized group of citizens. Having CIT Platform for rare diseases, patients would be able to actively participate in their own treatment in different ways. They will have better status in society along with better ability to independently organize themselves, and the motivation to do additional efforts to decrease the consequences of their illness and healing where possible. They will be able to contact and share experiences with other people, find information about the disease, medications, therapies, medical institutions that deal with their disease, research centres, pharmaceutical companies, the latest scientific developments and other useful links.

In addition to the project significance (which will over time get more and more important) goes the fact that the estimation of the number of rare diseases patients (about 5%) is very significant and it grows. Only in Serbia the number of those patients reaches 400000, about 30 million in Europe, i.e. 250 million around the world! As science progresses, the number of rare diseases will probably increase, even though it has not been small yet.

Because of these figures, pharmaceutical companies nowadays increase attention directed to this segment. Their income from this area becomes more important. We can expect that their further engagement will grow, and then, medical data collected and stored in repositories can be essential for them, both for planning and for the research itself, in the area of new medications. The problem of rare diseases is known and marked as extremely important all over the world. The number of projects, international associations, organizations and sites dealing with this stuff grows every day (www.rarediseases.org, www.raredisease.org.uk, www.eurordis.org, http://www.crdnetwork.org, etc.).

In Serbia, some medications for rare diseases lately became a part of government's so-called "positive list", which means that patients can get them free of charge. Unfortunately, only few of them are on positive list currently.

3 CENTRAL REPOSITORY

The number of patients suffering from rare diseases is not significant in health-care facilities; there are diseases with one or zero patients per facility. The problem with rare diseases is that critical amount of data needed for research and education is spread all over the world. Also, due to the procedural, administrative and other problems those data cannot be gathered easily from different clinics located in different countries which are pretty interesting from researchers' point of view. For this reason we are commencing development of this platform, having leading idea to easy gather medical data about rare diseases as much as possible at one place.

The first step in implementing CIT platform for rare diseases would be the classification of diseases and gathering significant medical information about every entity (name, type, common and specific symptoms, possible treatments, and many more). That would be much easier part of the research. The harder one is gathering medical data about patients suffering from these diseases. The idea how to gather significant medical data about particular cases is described later. This process must be performed in order not to harm patients' privacy and make their lives even more complex.

3.1 Collecting Rare Diseases Data

We will try to define several use cases how to collect data in proper way, both from medical and legal point of view. Our intention is, again, to collect more relevant medical data without making patient's personal data directly connected and exposed in public. Possible cases are:

- Physicians directly involved in rare disease medical treatments will be able to enter medical and demographic patients data (with the respect of patients privacy) when they discover new case of rare disease (through the corresponding web application). They will leave symptoms, used medicaments, results of laboratory analyses, recommended therapies, results of therapeutic treatments, etc.
- Web services that could collaborate with differ-

rent medical information systems will be developed as a part of IT platform, to help clinics partners provide electronic data from theirs EHR. This step is very difficult to project, because of databases heterogeneity in different MIS. To avoid bad data, beside data recognition, human factor will have to take place in this way of collecting data.

- Patients themselves will be able to leave their data. This kind of data will be marked as 'patient left', so it could be included or excluded from research on researcher's request.

3.2 Data Evaluation

All gathered data should be evaluated in order to avoid bad or missing data for further researching. Curtain procedures will be developed to check data, and to filter or delete bad data. In some cases procedures will add references to provide missing referential data integrity. The set of procedures that will be performed on gathered relevant data will present "Data evaluation tool".

Relevant data, in this context, is any medical information relevant to the disease. It is possible to develop various software tools for analytics, faster diagnosis, conclusion based on data-mining algorithms, and so on.

4 SOFTWARE TOOLS

Three kinds of software tools will be developed inside of the platform: communication tools, business intelligence (BI) tools and data mining tools as described further.

4.1 Communication Tools

Communication tools are of importance both for physicians and for patients themselves. Generally, there is a possibility to use different tools for patient to patient communication, patient to doctor, patient to institution, and patient to patients association. Communication tools will provide technical ability to the patients to create "virtual associations for certain types of rare diseases" through CIT platform. If implemented in international level, the platform would be localized to different languages.

The idea is every patient to have account, every association to have account and responsible person, and every institution or facility to have account and responsible person. The relation between entities will be precisely defined.

4.2 Business Intelligence Tools

BI in medicine can be referred to as clinical intelligence (Wickramasinghe, Schaffer, 2006). In our plan, this kind of tools represents the set of software tools for analyzing data repositories that can trigger periodically or on demand, and are designed to find appropriate templates related to specific diseases. Beside analysis tools, there can be software tools for setting hypotheses about the relationship of appropriate data and factors that could potentially be of use to researchers dealing with rare diseases.

These tools would be available to any medical institution or university and its researchers who want to participate in the project. Researchers should also share their experiences, opinions, suggestions and results between them on platform's forums.

4.3 Data Mining Tools

Each patient record consists of medical history symptoms, medical conditions, and various tests and lab results. In order to categorize those symptoms, and to recognize the existence of rare disease, we will apply machine learning classifiers: decision tree, support vector machine (Vapnik, 1998) and Bayesian sparse trained logistic regression (Tipping, 2001).

Basically, there are two main tasks which will be pursued during the building of these classifiers. First task will be of course, to obtain as accurate as possible classifiers which will be able to recognize (based on symptoms and lab results) the presence of rare disease. The second one, also important will be to recognize which symptoms and lab (test) results are relevant features for obtaining good recognition accuracy. By selecting relevant test results which are sufficient for accurate categorization (recognition) of the disease we are able to cut the cost of the potentially expensive tests and save the patients from unnecessary physical exertion.

5 PLANNED ACTIVITIES

Brief overview of CIT platform planned activities can be presented in two parts, as follows. Activities related to data repository:

a. classification of diseases and data structure modelling;

b. definition of information relevant to certain types of diseases;

c. central repository database designing;

d. design and creation of web application for manual data entry by physicians;

e. design and creation of web services for automatic data extraction;

f. definition of tools for data analyzing and reporting (Data Mining and BI);

g. definition of tools for publishing the results of analysis;

Activities related to collaborative platform:

h. portal design and implementation;

i. setting up an initial rare diseases data;

j. Portal promotion in national and international level.

The most complicated phase in the CIT platform project development is expected to be the phase e., because it implies collecting and importing data from heterogeneous medical data sources. Differences in medical standards between countries will create special difficulties, which could be partly overcome by involving international medical centers in the very beginning of the project.

6 CONCLUDING REMARKS

Due to the project specificity and its national (international) importance, precise analysis of the evaluation plan for the return of investment is relatively difficult. The very nature of the platform is such that its' result is significant at the national level and in a broader perspective. For these reasons we made a project application and funding request that has been sent to Serbian government. Since we are supported in this work partly by Nis Clinical Centre and partly from National health insurance organization we expect that our efforts will be recognized as national interest.

At the same time there is a benefit immeasurable financially, reflected in the satisfaction of patients, and their restored sense that society cares about them, and kind of returning them from the social care margins. The other useful effect can be achieved by including the pharmaceutical companies in the platform through advertising their manufactures, through the use of data collected and payment for the service. A special aspect is the possibility of forming a Balkan or even European data centre to collect the data on rare diseases. Certainly, the perspective is that after the project achieves the planned results, it can be spread in some international projects (for example within some of FP7 calls).

There is also a factor directly immeasurable: the patients themselves can access useful information to reduce their cost of treatment and personal problems. There is a great probability that the costs of rare diseases diagnostic can be significantly reduced, if we successfully develop such a platform. If we involve medical doctors employed in the public health in education and use of such CIT platform, they can reduce the number of expensive medical analyses for diagnostic in rare diseases.

Software tools can greatly assist rapid diagnosis of rare diseases, of course after a period of data collection in the repository, in order to create a sufficient quantity of data to perform the conclusions based on the knowledge base.

And at the end, software tools that will be developed to analyze the data stored in the repository can be used for many similar and commercial databases. Also designed model of CIT platform can be used in other social and public needs (justice, sport, investment, etc.).

REFERENCES

Michael, M., Holl, R., Badawi, O., Riker, R.R., Silfen, E., 2010. *A Collaboration Between Industry, Health-Care Providers, and Academia*, IEEE Engineering in Medicine and Biology, Volume 29, Number 2, pp. 18-25.

Wickramasinghe, N., Schaffer, J. L., 2006, *Creating knowledge-driven healthcare processes with the Intelligence Continuum*, International Journal of Electronic Healthcare, Vol. 2, No. 2, pp. 164-174.

Rajkovic, P., Jankovic, D., Tosic, V., 2009, *A software solution for ambulatory healthcare facilities in the Republic of Serbia*, in Proc. of HealthCom2009, Sydney, Australia, IEEE, pp. 161-168

Rajkovic, P., Jankovic, D., Stankovic, T. N., 2009, *An e-health solution for ambulatory facilities*, in Proc. of ITAB 2009, Larnaca, Cyprus, IEEE, pp. 1-4

Kukafka, R., August 2007, *Redesigning electronic health record systems to support public health*, Journal of Biomedical Informatics Volume 40, Issue 4, Pages 398-409.

Ford, E.W., Menachemi, N., Phillips, M. T., 2006, *Predicting the adoption of electronic health records by physicians: when will health care be paperless?*, J Am Med Inform Assoc 13, pp. 106–112.

DesRoches, C. M., at all, *Electronic Health records in Ambulatory Care – Survey of Physicians*, The New England Journal of Medicine, 2008, 359: 50-60.

Tipping, E. M., 2001, *Sparse Bayesian Learning and the Relevance Vector Machine*, Journal of Machine Learning Research, pp. 211-244.

Vapnik, V. N., *Statistical Learning Theory,* Wiley, 1998.