We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900
Open access books available

185,000
International authors and editors

200M
Downloads

154
Countries delivered to

Our authors are among the

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

BOOK CITATION INDEX
CLARIVATE ANALYTICS
INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

Monitoring Human Resources of a Public Health-Care System through Intelligent Data Analysis and Visualization

Aleksander Pur[1], Marko Bohanec[2], Bojan Cestnik[4,2], and Nada Lavrač[2,3]
[1] Ministry of Interior Affairs, Ljubljana,
[2] Jožef Stefan Institute, Ljubljana,
[3] University of Nova Gorica, Nova Gorica,
[4] Temida, d.o.o., Ljubljana,
Slovenia

1. Introduction

According to the World Health Report (World Health Organization, 2000), a health-care system (HCS) is a system composed of organizations, institutions and resources that are devoted to producing a health action. Human resources are one of the main parts of this system.

This paper is focused on the model for monitoring and planning of human resources in the Slovenian public HCS. The HCS of Slovenia is divided into the primary, secondary and tertiary health-care levels. The primary health-care (PHC) is the patients' first entry point into the HCS. It is composed of four sub-systems: general practice, gynaecology paediatrics and dentistry.

We have developed a model for monitoring the network of physicians at the PHC level, taking into the account the physicians' specializations, their geographic and work-time dispersion, time capacity constraints and their availability for patients. The motivation for this development came from the Ministry of Health of the Republic of Slovenia, who need a holistic overview of the PHC network in order to make short- and long-term management decisions and apply appropriate management actions, as well as evaluate PHC target achievements.

The first step was to form a data warehouse; a corresponding Entity Relationship Diagram data model was composed of unique database entries from the following existing sources:

- Slovenian Social Security databases: the data about health-care providers together with assigned patients per individual general practitioner, assigned patients with social security, and the data about health-care centres,
- the database of Slovenian physicians and dentists (Slovenian Medical Chamber),
- the database of the National Institute of Public Health containing data about Slovenian health centres, and
- the database of the Slovenian Statistics Bureau concerning the demographic and geographic distribution of citizens and communities in Slovenia.

The next steps involved the development of the model for monitoring the network of primary-care professionals based on the established data warehouse. The model was used

on real HCS data for the year 2006, which forms a part of a larger model for monitoring the responsiveness of the HCS for population, developed for the Ministry of Health of the Republic of Slovenia.

2. Methodology

Despite many frameworks related to performance and activity monitoring (Data-driven Decision Support System (DSS) (Power, 2002), Performance Monitoring, Business Performance Management (BPM), Business Activity Monitoring (BAM) (Dresner, 2003) etc.), there is a lack of methodologies for representing the concept of monitoring based on different data analysis methods. A similar methodology is addressing Performance Monitoring Protocols presented by the Working Party on Performance Monitoring in the Public Services (Bird, 2005). In this section, we present our approach to performance monitoring in modelling the PHC network in Slovenia.

2.1 Approach to human resources monitoring

Our model for monitoring the network of primary-care professionals in the Slovenian HCS consists of hierarchically connected modules. Each *module* is aimed at monitoring some aspect of the PHC network, which is of interest for decision-makers and managers of the network. Typical aspects about physicians are, for example: age and qualification of physicians, their workload and geographical distribution.

Each module involves a number of monitoring processes, which are gathered according to a given monitoring goal. Each *monitoring process* is characterised by: monitoring objectives, input data, data collecting methods, constraints on data, data dimensions, data analysis methods, output data, target criteria or target values of outputs, output data representation and visualisation methods, security requirements and the users of the monitoring system. Among these components, the *data analysis methods* transform the *input data* to *output data* represented using some *data representation formalism* according to the given *monitoring objectives*. The *target* is a level of performance that the organization aims to achieve for a particular activity. Information about *data collection* shows how and how often the data has been collected or needs to be collected (e.g., data can be collected by representative surveys or by standard procedures in organizations according to some refreshment rate). The *constraints* define the valid input and output data. *Security requirements* define the use and management of the monitoring processes and of the data.

This approach is not limited to any particular *data analysis method*. In principle, any methods can be used, such as Structured Query Language (SQL) procedures, On Line Analytical Process (OLAP) techniques for interactive knowledge discovering, as well as knowledge discovery in data (KDD) and data mining methods (Han, 2001) for discovering important but previously unknown knowledge. The same holds for *data representation* methods, which can include pivot tables, charts, network graphs and maps.

With respect to monitoring goals, *output variables* can be classified in different categories like *lead* and *lag* (Niven, 2003). The *lead* ones measure the performances that have influence on achieving the goals, whereas the *lag* are only related to the degree of achieving the goals.

In order to improve the comprehensibility of the model, its modules are hierarchically structured. The modules at the top level represent the main objectives. Usually all the main objectives can be incorporated in a single top-level module. The modules at a lower level are

connected to the one at a higher level. Each connection represents a data channel that connects outputs of the lower level module with the inputs of a higher-level module. In principle, the hierarchy is constructed so that the results of lower-level processes could help to explain the results of monitoring processes at a higher level. For example, the module for the assessment of HCS responsiveness could be composed of the physical accessibility of Health Services, availability of resources of Health Services and the rate of visits of population to health-care provider.

2.2 The model of monitoring human resources in a HCS

The model of monitoring the network of physicians at the PHC level is made in accordance with the above described methodology. The real challenge was to improve the monitoring of human resources in the HCS by different KDD methods..

The main concept is described by the hierarchically connected modules, shown in Fig. 1. The module *human resources* represents the main aspect of monitoring. The included monitoring processes are aimed at the monitoring of anomalies, outliers and other interesting events related to the network of physicians. The lower-level modules intend to provide detailed explanations of these events.
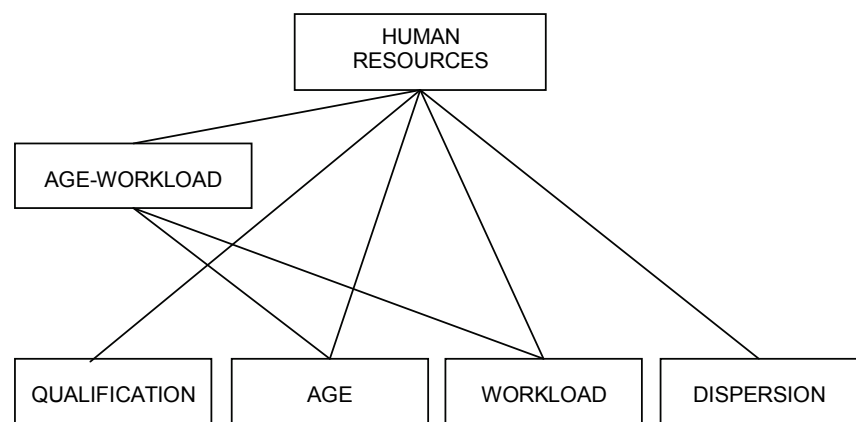

Fig. 1. The structure of the model for monitoring the network of physicians at a primary health-care level, represented by hierarchically connected modules.

3. Description of individual HCS modules

3.1 Monitoring of human resources

The main module *human resources* is aimed at a holistic monitoring of physicians' performance. In principle, the monitoring processes in this module must be able to process several input parameters. Therefore, different methods of KDD and multi-criteria decision models can be used for data analysis. The module includes a monitoring process that intends to represent the main aspects of physicians characterised by their *qualification*, *age*, *gender*, *workload* and *dispersion* (the number of locations where the physician works). The monitoring process is based on the OLAP model, which uses the dimensions: *time*, *location* where physicians work, *specialisation* and *gender*. The results based on a prototype application are presented by pivot tables and multidimensional charts, as shown in Fig. 2.

The scatterplot on the right side of Fig. 2 shows the average age, average workload, and average dispersion of physicians in communities for different specializations. The x-axis shows the average age of physicians, while the average workload is shown along the y-axis.

The communities are shown by shapes and colours of data points as explained in the legend. The size of these points is proportional to the average dispersion of physicians in the community. The specialization and gender of physicians could be selected using combo boxes in the top left corner. Thus, the scatterplot shows these average values for a gynaecologist working in some Slovenian community. For example, the workload and age of physicians shown on the top right corner are above the average. This presentation clearly shows outliers and anomalies in the HCS. Detailed information about these interesting aspects could be found in lower level modules. On the left side of Fig. 2, the same data is represented by the pivot table.

This module also includes a monitoring process aimed at discovering relations between the main aspects of physicians using the methods of association rules discovery (Srikant, 1996). The monitoring process tries to find the outliers overlooked by previous OLAP analyses. Table 1, for example, includes some rules focused on the relations between communities and physicians working in general practice. For example, for the community Škofja Loka it is characteristic that physicians younger than 40 years are underloaded (see the rules 2, 3). This module could also include the monitoring processes based on multi-criteria decision models (Bohanec, 2006).
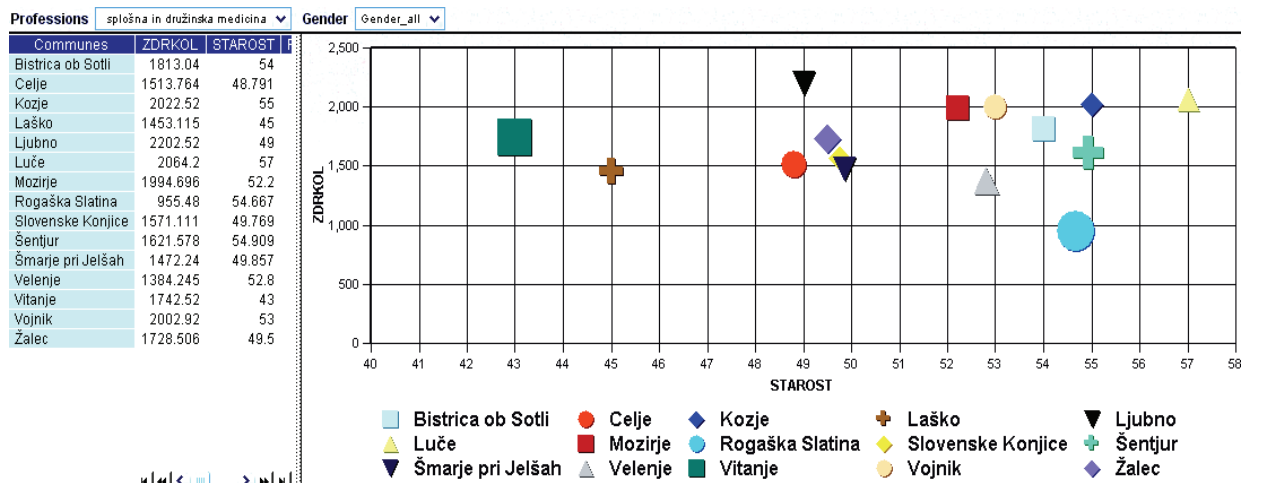

Fig. 2. The holistic aspect of physicians presented by OLAP techniques.

Rule	Supp.	Conf.	Lift
[age:60+]+[workload: middle] ==>Ptuj	0.51%	14.63%	5.99
[dispersion:1]+[age:to40]+[workload: small] ==>Škofja Loka	0.34%	5.97%	4.42
[age:to40]+[workload: small] ==> Škofja Loka]	0.34%	5.19%	3.85
[workload: large]+[age:40-60]+[gender] ==>Domžale	0.42%	7.04%	3.63
[age:to40]+[workload: middle] ==> Novo mesto	0.42%	9.62%	3.46
[gender:z]+[age:to40]+[workload: small]==> Domžale	0.34%	6.35%	3.27
[gender:z]+[age:to40]+[workload: middle]==> Novo mesto]	0.34%	8.89%	3.19
[workload: large]+[gender:m] ==>Domžale	0.51%	6.00%	3.09
[age:60+]+[workload: large] ==> Maribor	0.67%	25.81%	3.00
[age:60+]+[workload: large]+[gender:m]==> Maribor	0.59%	25.00%	2.90

Table 1. Association rules showing relations between communities and general practice physicians.

3.2 Qualification of physicians

The aim of this module is to enable monitoring of physicians' and dentists' qualification for the job they actually perform. The main performance indicator is the physician's specialization degree, granted by the Slovenian Medical Chamber, which must be verified every 7 years. The specialization degree is a prerequisite for getting a license for employment in a certain area of medicine.

To monitor the suitability of physicians for the job they are performing we have used social network visualization technique available in the social network analysis program named Pajek ("Spider" (Batagelj, 2006)). The monitoring of physicians' suitability is achieved by the monitoring of three variables: SPEC (specialization), LIC (luicence), and OPR (the type of patients that the physician is in charge of, categorized by patient type). The motivation for this analysis is based on the observation that physicians with a certain specialization may get different luicenses, and while a certain licence assumes that the physician will only deal with patients of a certain patient category, in reality she may be in charge of different types of patients (e.g., a paediatrician may provide health services to grown up patients, although she has a specialization in paediatrics and a licence in paediatrics).

It has also been observed that an individual physician may have several specializations, and several licences, and hence patients of different types. The Pajek diagram (Fig. 3) shows well the typical (thick lines – a high number of physicians) and atypical (thin lines – a low number of physicians) cases, which enable abnormality detection and further analysis of individual discovered anomalies.
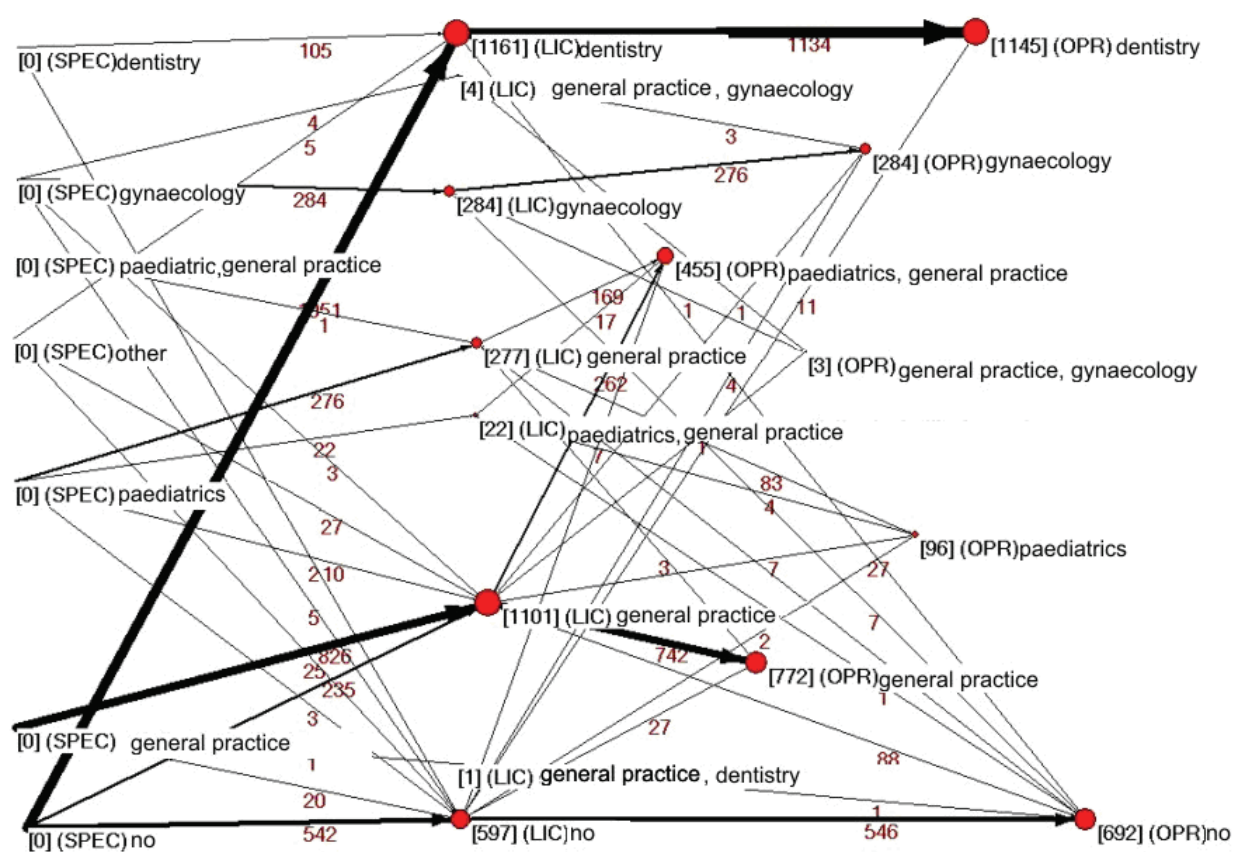

Fig. 3. The qualifications of physicians for the job they are performing.

3.3 Age of physicians

The monitoring processes in this module are aimed at age and gender analyses of physicians. The module includes a process based on the OLAP model. The dimensions of this model are *age*, *gender*, *specializations* and *locations* where they work. The main monitored quantity in the facts table is the number of physicians. The results are presented by pivot tables and different charts. For example, the number of gynaecologists by age and gender is presented in the chart in Fig. 4. The x-axis shows the age variable, while the number of doctors is shown along the y-axis. The gender is shown in the legend. Although relatively simple, this chart clearly shows a decline of the number of young male gynaecologists that have accomplished the studies in the last twenty years. Generally, the number of physicians is related to the number of students that have accomplished the studies at the Faculty of Medicine in Ljubljana that is the main source of doctors in Slovenia. Thus, this presentation can help planning the education system. The other process in this module provides a list of gynaecologists that are near retiring age. The list can be used to help planning missing human resources in the near future.
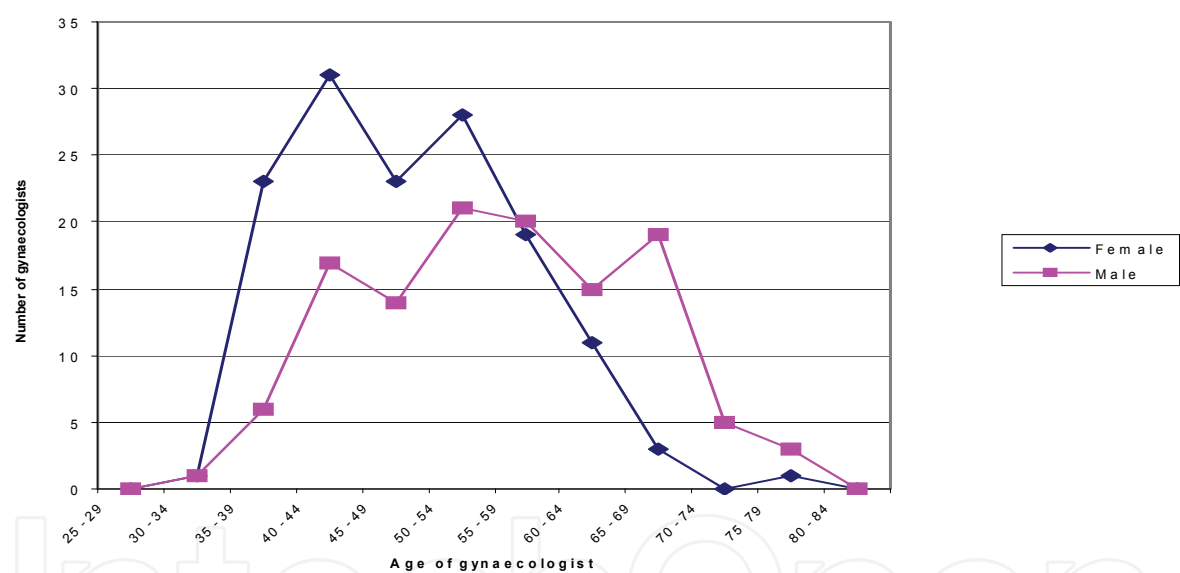

Fig. 4. The numbers of gynaecologists by age and gender.

3.4 Workload analysis

This module is aimed at monitoring the physicians' workload. Considering the available data, the assessment of workload is based on age-adjusted listed patients per physician. Each age group of listed patients is weighted according to use of health-care resources, e.g. the number of visits per physician. The physicians without registered patients are excluded from this analysis.

The monitoring process is based on the OLAP model. The dimensions of this model are: *time*, *specializations*, *locations* where the physicians work, the ratio between the number of

listed patients and physicians, and the ratio between the age-adjusted number of listed patients and physicians. The measure in the fact table is the sum of physicians. Again, results are presented in pivot tables and different charts.

3.5 Age-workload analyses

The monitoring processes in this module are aimed at the combined analyses of physicians' age and their workload (number of patients per physician). The considered dimensions are: *time*, *locations* and *professions*. The monitoring process provides the state of each physician regarding their age and the number of patients. For example, Fig. 5 shows the age and workload of general practitioners. The x-axis shows the physicians' age, while the y-axis shows the number of their patients. Physicians aged between 40 and 50 have the most patients, but some of them have a large number of patients also after 50. The retirement of physicians having a large number of patients has an important impact on the PHC network. This impact is more precisely described by the next monitoring process, which is based on GIS.
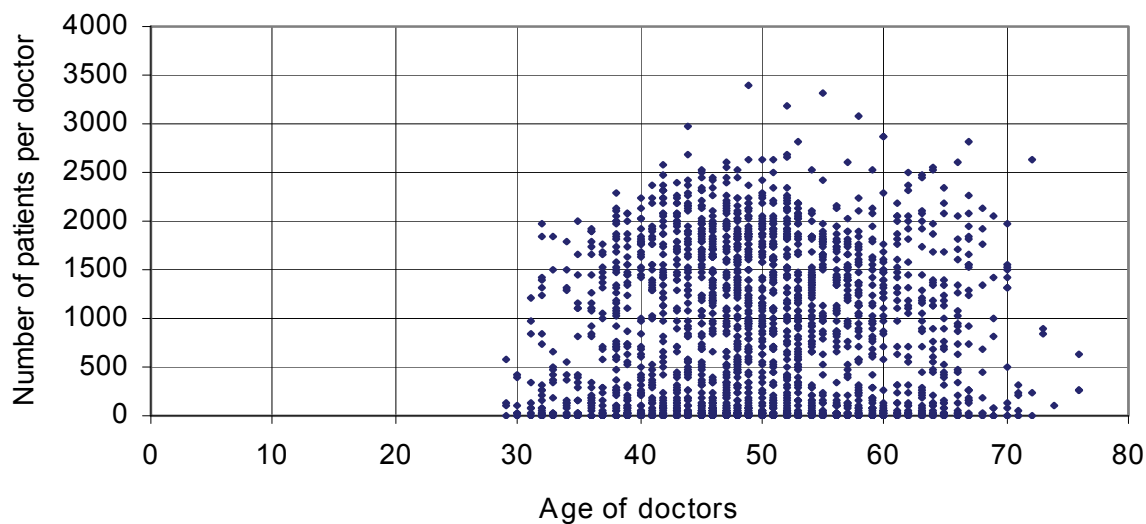

Fig. 5. The physicians in GP by age and listed patients.

The map in Fig. 6 shows the number and share of patients affected by the retirement of physicians in the next five years (until 2011). The assumption is that physicians will retire at the age of 65. In this map, each polygon represents a community. Their shade is proportional to the number of patients whose doctors will retire. The pie charts represent the ratio between the patients unaffected and affected by the retirement until 2011. Thus, this analysis provides information on the number of physicians and regions where they have to be replaced in the next five years.

From the implementation viewpoint, the latter module is composed of data about physicians' age, their registered patients and geographic data. The detailed information about physicians' age and patients is provided by subordinate modules *age* and *workload* (Fig. 1).
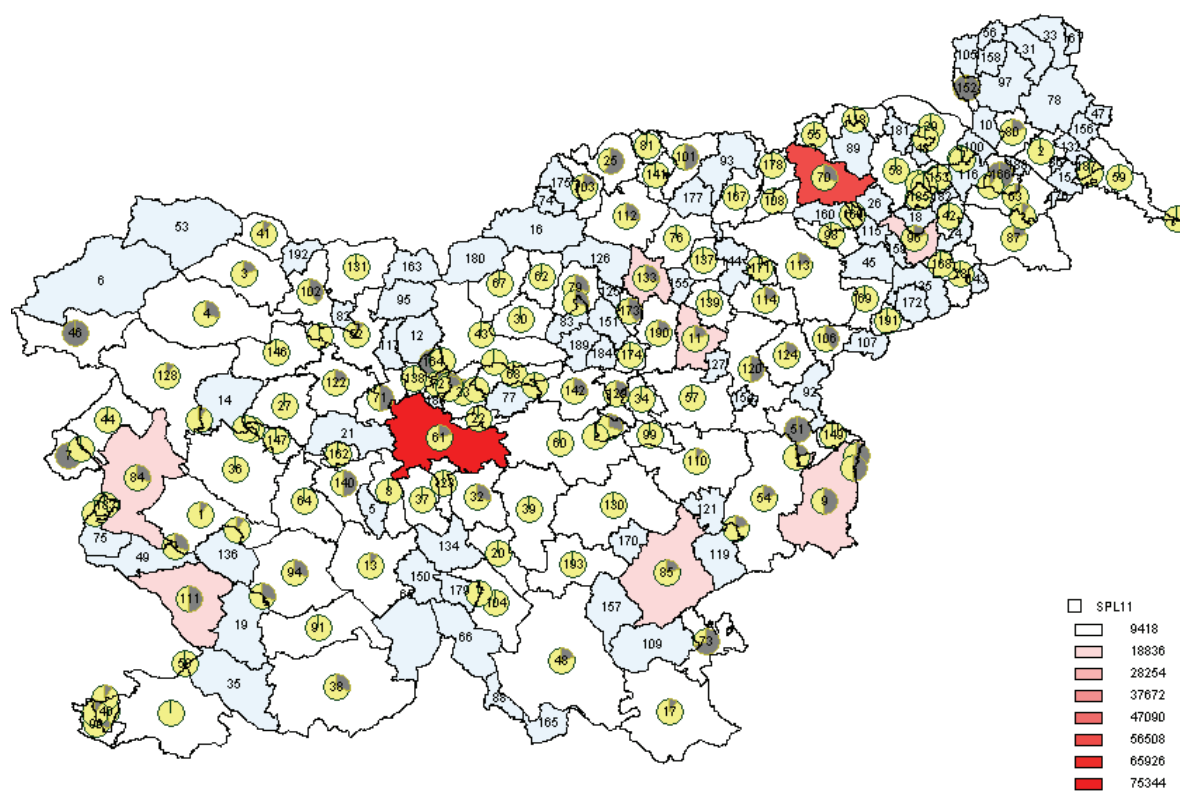
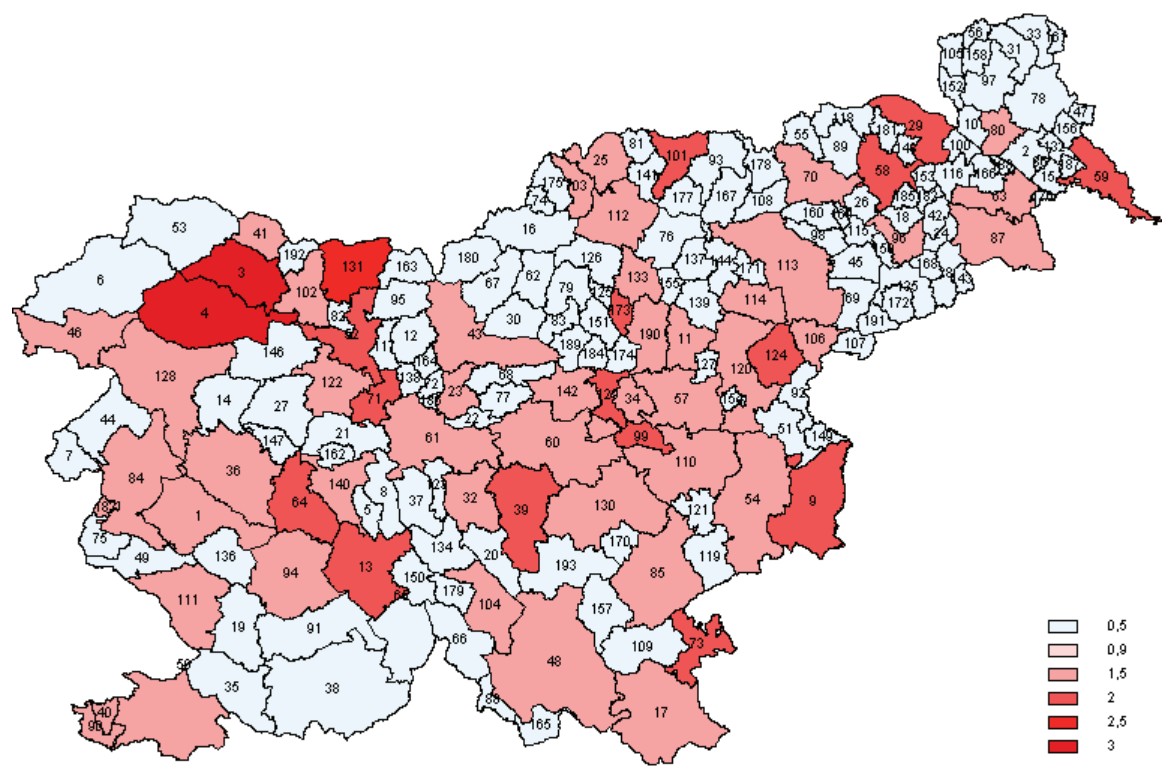Fig. 6. The impact of physicians retiring until 2011.

Fig. 7. Average dispersion of working locations.

3.6 Dispersion of working location

This module is aimed at monitoring the dispersion of locations where physicians work. Depending on the requirements, a physician may work on more than one location, but this dispersion usually means additional workload for physicians and their lower availability for patients at some location. The monitoring process provides the number of locations where physicians work, which are shown using the GIS techniques. Fig. 7, for example, shows the dispersion in 2005 for gynaecology. The darker the community are, the larger the average number of locations where physicians from this community work.

4. Conclusion

The aim of the presented human resource monitoring system is to assess performance and provide information necessary for the planning and management of primary health-care network in Slovenia. At the general level, the approach is based on a carefuly designed hierarchy of modules, each monitoring a specific aspect of the network – in particular physicians' age, qualifications, workload and dispersion. At the implementation level, the system uses a number of different techniques, including OLAP, KDD and data analysis (such as association rule mining). In most cases, the results are presented visually by charts, network graphs and maps. In this way, the monitoring system provides an information-rich picture of the network and its performance, and also helps detecting its critical aspects that require short- or long-term management actions. In principle, the higher levels of the model provide holistic information, while the lower levels provide more details that are useful for the explanation of observed phenomena.

The monitoring system has been developed in collaboration with the Ministry of Health of the Republic of Slovenia. Currently, it is implemented as a prototype and has been tested with real data for the year 2006. For the future, we wish that it becomes a regular tool for monitoring the health-care network in Slovenia. We also envision the application of the same methodology in other public networks, such as education and police.

5. References

World Health Organization: World Health Report 2000: Health Systems. Improving Performance, http://www.who.int/whr/2000/en/whr00_en.pdf, Accessed January 26, 2007 (2000)

Niven, R., P.: Balanced Scorecard for Government and Nonprofit Agencies, John Wiley and Sons, Inc (2003), ISBN 0-471-42328-9

Bird, M., S. (ed.): Performance indicators good, bad, and ugly, Working Party on Performance Monitoring in the Public Services, J. R. Statist. Soc. A (2005) 1-26

Bohanec, M.: DEXi, A Program for Multi-Attribute Decision Making. http://www-ai.ijs.si/MarkoBohanec/dexi.html, Accessed January 26, 2007 (2006)

Batagelj, V., Mrvar, A.: Program for Analysis and Visualization of Large Networks. Reference Manual, University of Ljubljana, Ljubljana (2006)

Dresner, H.: Business Activity Monitoring, BAN Architecture, Gartner Symposium ITXPO, Cannes, France (2003)

Power, J., D.: Decision Support Systems. Concepts and Resources for Managers, Quorum Books division Greenwood Publishing, ISBN: 156720497X (2002)

Han, J., Kamber, M.: Data Mining: Concepts and Techniques, Morgan Kaufmann Publishers (2001)

Srikant, R., Agrawal, R.: Mining Quantitative Association Rules in Large Relation Tables, IBM Almaden Research Center, San Jose (1996)

Data Mining in Medical and Biological Research

Edited by Eugenia G. Giannopoulou

This book intends to bring together the most recent advances and applications of data mining research in the promising areas of medicine and biology from around the world. It consists of seventeen chapters, twelve related to medical research and five focused on the biological domain, which describe interesting applications, motivating progress and worthwhile results. We hope that the readers will benefit from this book and consider it as an excellent way to keep pace with the vast and diverse advances of new research efforts.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Aleksander Pur, Marko Bohanec, Bojan Cestnik, and Nada Lavrač (2008). Monitoring Human Resources of a Public Health-Care System through Intelligent Data Analysis and Visualization, Data Mining in Medical and Biological Research, Eugenia G. Giannopoulou (Ed.), ISBN: 978-953-7619-30-5, InTech, Available from: http://www.intechopen.com/books/data_mining_in_medical_and_biological_research/monitoring_human_reso urces_of_a_public_health-care_system_through_intelligent_data_analysis_and_visu

INTECH
open science | open minds