Medical Informatics Europe '99
P. Kokol et al. (Eds.)
IOS Press, 1999

TeCoMed
Online Consulting of Acute Health Risks

G. Greiner, G. Kundt, L. Gierl
University of Rostock
Institute for Medical Informatics and Biometrics
Rembrandtstr. 16/17
18055 Rostock
Germany

Abstract. Geomedical systems are information systems used to monitor the health situation in geographically distributed populations. Both are used to detect and to forecast health risks. In this paper, we present the system TeCoMed, which fulfills the following tasks: (1) discovering health risks; (2) forecasting temporal and spatial spread of epidemics; (3) estimating consequences of an epidemic; (4) presentation of the results via Internet.

Keywords: medical informatics, information system, case-based reasoning

1 Introduction

Physicians, hospitals, and health authorities need sufficient information on current health risks including outbreak and spreading of diseases with a high contagion level to ensure a high level of human health protection. The TeCoMed Project provides an application for medical informatics, which empowers MDs, those heading medical institutions, health insurance companies, and health authorities as future users to engage in a decision-making process regarding the early combat of contagious diseases. It is geared to accomplish the following tasks: Reflecting the general state of health, early discovery of acute health risks and outbreaks of infectious epidemics, detection of hitherto unknown diseases, forecasting of the temporal and spatial spread of epidemics as well as the economic and constitutional consequences of a given epidemic.

Our Online Consulting System consists of three fundamental, mutually engaged components:

- Data capturing
- Information processing
- Presentation and visualization on the Internet

2 Data capturing

In our system digitized health data of the AOK in Mecklenburg-Vorpommern* can be used. The AOK sends patient records, which contain information about medical treatment of insurants. More precisely, the following items are contained: decade of birth, sex, encoded profession, postcode of the domicile, date of fall ill, emergency, convalescence, death, and the diagnose given by ICD (International Statistical Classification of Diseases and Related Health Problems). A comprehensive data-pool is available, because about 40% of the Mecklenburgian

* AOK: General Health Insurance Company; Mecklenburg-Vorpommern: state in north-east Germany

population are registered under the roof of the AOK Health Insurance Company. The data is rendered anonymous in accordance with the law of the land, and the capturing process takes place largely on a self-induced basis via ISDN transfer. It is kept on a database for further processing.

In addition, data of the state health authorities is included. We receive weekly reports about notifiable diseases, which include microbiological data.

3 Information processing

The considerable quantity of acquired health data provides material for sequences of health situations to discover health risks. Although we have a large and rich data source, we have to notice that these data can not be considered as a representative sample within the meaning of statistics. The social structure as well as the age distribution of the insurants and of the whole population does not correspond. Consequently parametric statistical methods are not applicable. Therefor we combine various geostatistical methods and case-based reasoning to evaluate the collected data.

3.1 Case-based Reasoning

Case-based reasoning (CBR) is a cyclical paradigm of artificial intelligence for solving problems by analogical reasoning and for learning new problem-solution pairs through experiences. CBR is an approach based on the presupposition: similar problem instances have similar solutions. The CBR-cycle consists of four stages (due to Aamodt and Plaza [1]): *retrieve* the most similar case or cases; *reuse* the information and knowledge in that case to solve the problem; *revise* the proposed solution; *retain* the parts of this experience likely to be useful for future problem solving.

We apply the CBR-techniques to answer mainly the following question: What could be the further development of a conspicuous health situation?

By a *case* we understand a concept which describes the public health situation in the considered area (state Mecklenburg-Vorpommern) during a time period, such a description is called a *scenario* Σ. Because we examine the health situation during a time period a scenario is a sequence of separate *situations*, $\Sigma=(\sigma_1,\ldots,\sigma_n)$. Here a *situation* reflects the public health state for a certain week.

For each postcode unit we accumulate the information about new cases of disease and construct scenarios. All scenarios are stored in a case-base Σ^*. Given a current scenario Σ, the system retrieves all scenarios $\Sigma'\sigma'_{n+1} = (\sigma'_1,\ldots, \sigma'_n,\sigma'_{n+1})$ from the case-base Σ^* so that the differences between Σ and Σ' are minimal. By using the retrieved scenarios $\Sigma'\sigma'_{n+1}$ the system adapts the situation σ'_{n+1} to a forecast situation σ_f. The result will be stored in the case-base and will be available for graphical presentation. After a time period we know the situation σ_{n+1}. In comparing the forecast situation σ_f with the situation σ_{n+1} the system evaluate the forecast mechanism. If there are too many differences between these situations the system has to integrate the new scenario $\Sigma\sigma_{n+1}$ in the case-base or the inference mechanism must be changed.

3.2 *Geostatistical Methods*

A given situation σ is analyzed with geostatistical techniques, such as the G Test developed by Getis and Ord (see [2], [3]), to diagnose constitutional risks. The G-statistic is based on a permutation test, which measures the concentration or the lack of concentration of the sum of considered values in an area. The area (state Mecklenburg-Vorpommern) is subdivided into

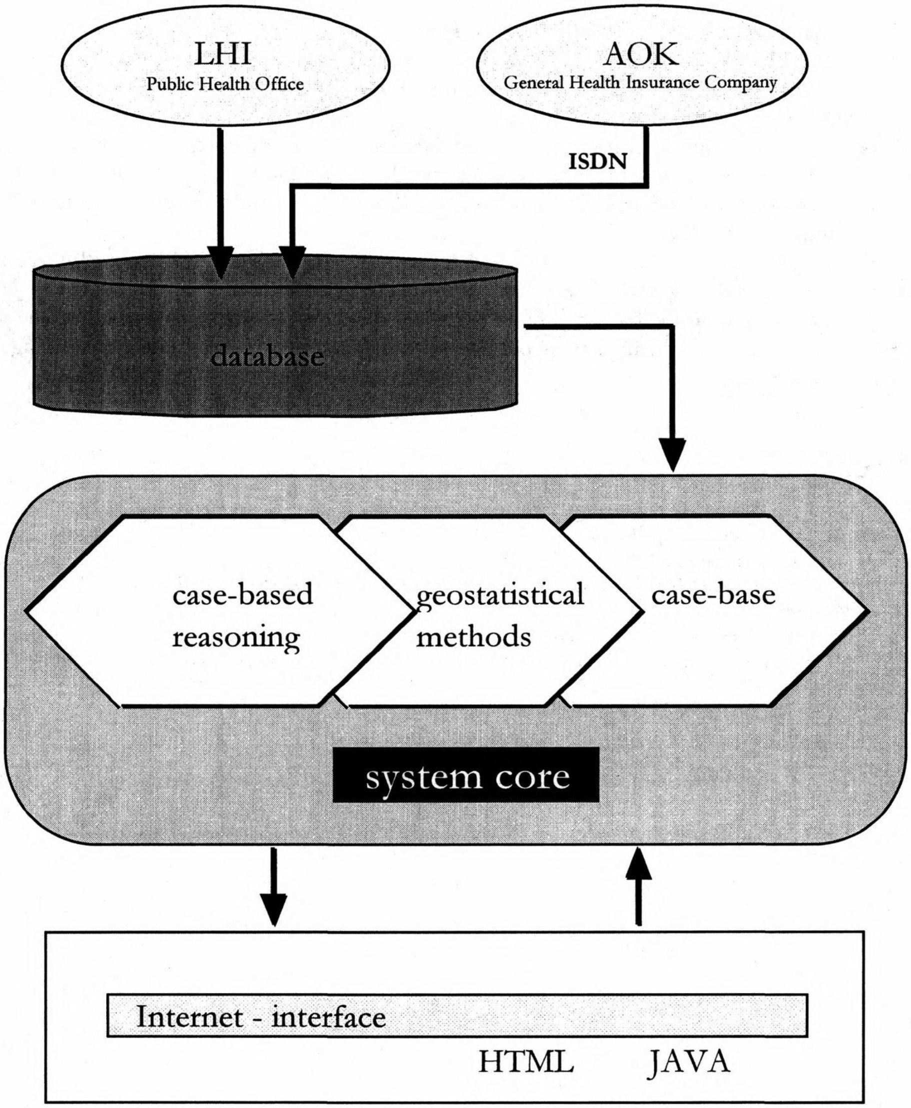

Figure 1. Structure of the System TeCoMed

regions (postcode units). Each region is associated with an incidence value for every disease. We fix the value for a certain region and consider all permutations of the remaining values. The statistic is subsequently calculated as the proportion of the sum of all values that are within a certain neighborhood of the region to the total sum of the area. Under the null hypothesis of spatial independence, the permutations are equally likely, and the statistic is approximately normal distributed. By testing the null hypothesis we detect regions with a conspicuous concentration of either high or low values in the environment.

4 Presentation and Visualization on the Internet

The informational content of the health data captured will be visually represented on the TeCoMed Web Site. Depiction of graphics allows a quick and intuitive recognition of the data results for the respective user groups. In particular, geomedical maps allow, in the given case, an individual evaluation by the user without need to presuppose knowledge of the data structure. In an initial phase of the TeCoMed Project, two independent prototypes for visualization were designed, which can be used as well.

Visualized health data and constitutional scenarios are made available on the Internet to conform to individual needs of respective user classes. In addition, HTML pages and Java applets (for instance on the basis of one of the above prototypes) will be presented on the net page. The access of the system shall be possible without prior knowledge of data processing on a WWW interface.

References

[1] Aamodt, A., Plaza, E. (1994) *Case-Based Reasoning: Foundational Issues, Methodological Variations, and System Approaches.* AI Communications, IOS Press, 7:1, 39 – 59.

[2] Ord, J.K., Getis, A. (1995) *Local spatial autocorrelation statistics: Distributional Issues and an application.* Geographical Analysis, 27, 286 – 306.

[3] Ord, J.K., Getis, A. (1995) *Local spatial autocorrelation statistics: Distributional Issues and an application.* Geographical Analysis, 27, 286 – 306.