

Decision Support Systems for Cardiovascular Diseases Based on Data Mining and Fuzzy Modelling

Markos G. Tsipouras

University of Ioannina, Greece

Themis P. Exarchos

University of Ioannina, Greece

Dimitrios I. Fotiadis

University of Ioannina, Greece

Michaelideion Cardiology Center, Greece

Biomedical Research Institute, Greece

Aris Bechlioulis

University of Ioannina, Greece

Michaelideion Cardiology Center, Greece

Katerina K. Naka

University of Ioannina, Greece

Michaelideion Cardiology Center, Greece

INTRODUCTION

The widespread availability of new computational methods and tools for data analysis and predictive modelling requires medical informatics researchers and practitioners to systematically select the most appropriate strategy to cope with clinical prediction problems. In particular, data mining techniques offer methodological and technical solutions to deal with the analysis of medical data and construction of decision support systems. Furthermore, fuzzy modelling deals with the ambiguity inherent in all medical problems. These methods can be used to design and develop clinical decision support systems (CDSSs), which, after evaluated from the experts, can be integrated into clinical environments.

Cardiovascular diseases (CVDs) are the leading cause of death in many countries worldwide. According to the World Health Organization, CVDs are the cause of death of 16.6 million people around the globe each year. The multifaceted nature of these diseases, combined with a wide variety of treatments and outcomes, and complex relationships with other diseases, for example, diabetes, have made diagnosis and optimal

treatment of cardiovascular diseases a problem for all but experienced cardiologists.

This article addresses the decision support regarding cardiovascular diseases, using computer-based methods, focusing on the coronary artery disease (CAD) diagnosis and on the prediction of clinical restenosis in patients undergoing angioplasty. Methods reported in the literature are reviewed with respect to (i) the medical information that are employing in order to reach the diagnosis and (ii) the data analysis techniques used for the creation of the CDSSs. In what concerns medical information, easily and noninvasively-obtained data present several advantages compared to other types of data, while data analysis techniques that are characterized by transparency regarding their decisions are more suitable for medical decision making. A recently developed approach that complies with the above requirements is presented. The approach is based on data mining and fuzzy modelling. Using this approach, one CDSS has been developed for each of the two cardiovascular problems mentioned above. These CDSSs are extensively evaluated and comments about the discovered knowledge are provided by medical experts. The later is of great importance in designing

and evaluating CDSSs, since it allows them to be integrated into real clinical environments.

BACKGROUND

Data Mining is the process of discovering patterns and correlations from large amounts of data, using artificial intelligence, statistical, and mathematical techniques (Tan, Steinbach, & Kumar, 2005). Fuzzy logic is the extension of the classical crisp (binary) logic into a multivariate form. Fuzzy logic is closer to the human logic, thus being able to deal with real world noisy and imprecise data (Wang, 1986). CDSSs are computerized tools developed to assist physicians through the process of decision making. A known approach for the development of CDSSs is the use of experts' knowledge combined with an inference engine. However, recent advances in designing CDSSs employ automated knowledge extraction from data, using data mining techniques, while fuzzy logic provides several advantages in designing inference engines, compared to the classical crisp logic. The combination of data mining and fuzzy modelling provides a powerful tool for fully automated creation of CDSSs, experiencing several advantages: (i) transparency in decision making, (ii) addressing the ambiguity inherent in clinical data, and (iii) ability to interpret all decisions in a medical manner. All the above are of great importance for physicians, when performing decision making.

Coronary artery disease (CAD) is the development of atherosclerotic plaques in coronary arteries, resulting in coronary luminal narrowing and subsequently occlusion, and thus leading to myocardial infarction or sudden cardiac death. Coronary angiography (CA) is considered to be the "gold standard" method for the diagnosis of CAD and it is widely used. However, CA is an invasive and costly procedure that needs high level technical experience and technology and cannot be used for screening of large populations or close follow-up of treatment (Escolar, Weigold, Fuisz, & Weissman, 2006). Computer aided methodologies for CAD diagnosis have also been proposed in the literature; in this case the data obtained by some of the above mentioned methods or other sources (i.e., laboratory examinations, demographic, and/or history data, etc.) are evaluated by a computer-based application, leading to a CAD diagnosis. These methodologies can be divided into various categories, based on the type

of data they use for subject characterization. Several methods are based on heart sounds associated with coronary occlusions (Akay & Welkowitz, 1993). Also, methods which employ the resting or exercise electrocardiogram (ECG) of the patient, extracting features from it, like the R wave (Szildgyi, Szildgyi, & David, 1997), the QT interval (Ng, Wong, Mora, Passariello, & Almeida, 1998), the T wave amplitude (Sabry-Rizk et al., 1999), the heart rate variability (Tkacz & Kostka, 2000), and the ST segment (Lewenstein, 2001) have been proposed. Furthermore, there are methods using medical images, such as SPECT (Haddad, Adlassnig, & Porenta, 1997), and methods based on arterio-scillography (Pouladian, Golpayegani, Tehrani-Fard, & Bubvay-Nejad, 2005). There exist also methods that employ demographic, history, and laboratory data (Frossyniotis et al., 2001; Mobley, Schechter, Moore, McKee, & Eichner, 2005; Tsipouras et al., 2006) and methods that combine more than one type of data (Kukar, Kononenko, Groselj, Kralj, & Fettich, 1999; Scott, Aziz, Yasuda, & Gewirtz, 2004).

The evolution and widespread adoption of percutaneous transluminal coronary angioplasty (PTCA) represents a major advance in the management of acute coronary syndromes, resulting in a significant reduction in early and late mortality compared to pharmacologic reperfusion therapy. Coronary artery restenosis remains a major limitation of PTCA and is usually defined as $\geq 50\%$ stenosis in the treated segment at follow up, or at least 50% loss of the original gain in the minimal luminal diameter. Many clinical, angiographic, and procedural features have been studied as predictors for restenosis but it has proven difficult to stratify patients with regard to the risk of restenosis. Knowledge of risk factors for restenosis may help to refine indications of PTCA, reduce the frequency of restenosis, and select optimal candidates for a PTCA procedure. Despite lowering the restenosis rate with the implantation of coronary stents, it occurs in approximately 12-60% of the patients within 6 months after intervention depending mainly on the patients' and procedural characteristics. Computer aided methodologies for the prediction of clinical restenosis have also been proposed in the literature. These can be divided into two categories, regarding the data they use for the analysis: methods that employ only CAD risk factors such as demographic, history, and clinical data (Budde, 1999; Tsipouras et al., 2006) and methods that combine CAD risk factors with angiographic features (Maier, Mini, Antoni, Wischniewski, & Meier, 2001; Resnic, Popma, Ohno-Machado, 2000).

7 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:

www.igi-global.com/chapter/decision-support-systems-cardiovascular-diseases/12965

Related Content

Medical Data Visualization via a Pervasive Multi-Agent Platform

Antonio Coronato, Luigi Galloand Giuseppe De Pietro (2009). *International Journal of Healthcare Delivery Reform Initiatives* (pp. 68-81).

www.irma-international.org/article/medical-data-visualization-via-pervasive/40334

User Driven Learning in Mathematics

Monica Kochar (2011). *User-Driven Healthcare and Narrative Medicine: Utilizing Collaborative Social Networks and Technologies* (pp. 242-254).

www.irma-international.org/chapter/user-driven-learning-mathematics/49257

Wireless Networking Credibility, Device Interoperability & Other Important Issues to Take Into Consideration for the Deployment of a Homecare Service Provision Model

Konstantinos Perakisand Dimitris Koutsouris (2011). *Smart Healthcare Applications and Services: Developments and Practices* (pp. 1-23).

www.irma-international.org/chapter/wireless-networking-credibility-device-interoperability/50653

Conceptual Levels of Information Processing and Information Interpretation in Knowledge Management

Murako Saito (2010). *Redesigning Innovative Healthcare Operation and the Role of Knowledge Management* (pp. 75-90).

www.irma-international.org/chapter/conceptual-levels-information-processing-information/36518

Non-Traditional Data Mining Applications in Taiwan National Health Insurance (NHI) Databases: A Hybrid Mining (HM) Case for the Framing of NHI Decisions

Joseph Tanand Fuchung Wang (2017). *International Journal of Healthcare Information Systems and Informatics* (pp. 31-51).

www.irma-international.org/article/non-traditional-data-mining-applications-in-taiwan-national-health-insurance-nhi-databases/187046