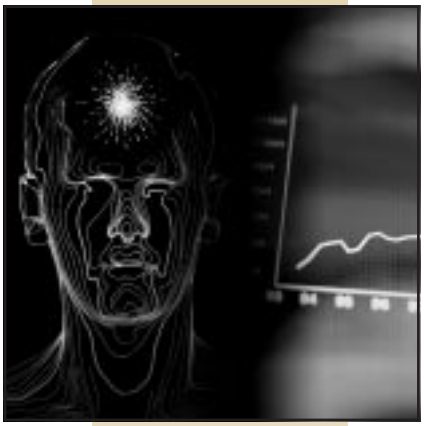


A Knowledge Discovery Approach to Diagnosing Myocardial Perfusion (DRAFT)

Applying a Six-Step Discovery Process to a Database of SPECT Bull's-Eye Maps of the Heart

Krzysztof J. Cios[1], Anna Teresinska[2], Stefania Konieczna[2], Joanna Potocka[2], and Sunil Sharma[1]

[1]University of Toledo, Ohio

[2]Institute of Cardiology, Warsaw

Modern medicine generates huge amounts of data, but at the same time there is an acute and widening gap between data collection and data comprehension. For instance, collection of single-photon emission computed tomography (SPECT), MRI, and positron-emission tomography (PET) images, or ECG and EEG signals can generate gigabytes of data per day. This volume calls for high capacity data storage devices and tools for analysis. Obviously, it is very difficult for a human to make use of thousands of images and to be able to understand basic trends in data to make rational decisions. The information becomes less and less useful as we are faced with difficulties of retrieving it and making it available to the user in an easily comprehensible format. Thus, there is growing pressure to automate methods of data analysis to facilitate the creation of knowledge that can be used for clinical decision-making. Data mining and knowledge discovery are tools that can help in achieving these goals [1].

The knowledge discovery process, however, is still in its infancy, being rather an art than a well-understood process. It is crucial then to have a general framework that will allow for diverse types of medical signals and images. Benefits of a general framework can be measured in terms of development time, reliability, efficiency, and the overall cost. Any such strategy should pay attention to:

- Medical concerns, which are at least equally important as the data mining and knowledge discovery aspects.
- Providing guidelines for reusing the experience gained on previous projects on new projects.

This article first discusses data mining and knowledge discovery issues. It then outlines a six-step knowledge discovery process. This process is applied to the problem of computerizing the process of diagnosing myocardial perfusion defects.

Databases, Data Mining, and Knowledge Discovery

The area of data mining and knowledge discovery is inherently associated with databases. Data mining methods are algorithms that are used on databases, after initial data preparation, for finding patterns or trends in the data. When data mining methods are used in a process and their outcome is evaluated, so that we can think about the product as being a new piece of information, we then talk about a knowledge discovery process. The term itself was defined by Frawley, et al. [2]: "Knowledge discovery in databases is the nontrivial process of identifying valid, novel, potentially useful, and ultimately understandable patterns in data."

A knowledge discovery process is difficult to define, since dozens of data mining methods can be used on any given data, and so far no theory exists of how to go about finding new and ultimately useful knowledge. However, the general steps outlined below for a knowledge discovery process can help in undertaking such a task.

Sometimes, the term knowledge discovery is used interchangeably with the term data mining. However, we shall understand knowledge discovery as a process within which data mining methods are used. When talking about the knowledge discovery process, we need to emphasize not only the iterative character between the data mining methods or tools used but also an interaction with the medical professional who always plays a cru-

cial and indispensable role in a knowledge discovery process.

Fundamental issues in data mining and knowledge discovery arise from the very nature of databases [1], namely:

- **Huge volume of data.** Because of database size, it is unlikely that any of the data mining methods will work with raw data. This calls for using a portion of the database with the hope that the results obtained are representative for the entire database. Reduction of the dimensionality can be done in two ways: sampling in data space, where some records are selected randomly and used afterwards for data mining, and sampling in feature space, where only some features of data are selected. Again when a large number of features exist, the selection can be done in a random manner.
- **Dynamic nature of data.** Databases are constantly updated either by adding new records, say SPECT images (for same or a new patient), or by replacement of the existing records, say when SPECT had to be repeated because of a technical problem. This mutability calls for data mining methods able to incrementally update the knowledge learned thus far.
- **Incomplete or imprecise data.** The information collected in a database can be either incomplete or imprecise. Fuzzy sets [3] and rough sets [4] were developed explicitly for the purpose of dealing with this problem.
- **Noisy data.** It is very difficult to entirely eliminate noise during data collection. This implies that data mining methods should be made less sensitive to noise, or care should be taken that the amount of noise in future data is approximately the same as in the current data.
- **Missing attribute values.** This creates a problem, since most data mining techniques require a fixed dimension of the data record. One approach to remedy this problem is to substitute missing attribute values with the most likely values, or to replace the unknown value with all possible values of that attribute.
- **Redundant or insignificant data.**

We distinguish three types of mining in data:

- **Directed mining.** For instance, the doctor is interested in learning some particular information, such as "is the 40-50% obstruction of the left circumflex functionally important?" Most medical applications probably fall into this category.
- **Hypothesis testing and refinement.** The user provides some hypotheses and expects the system to validate them or to modify them and develop other more refined versions.
- **Undirected or pure mining.** This is the most general case, where there are no constraints on the system, but there are also no indications of what the user expects and what type of discovery could be of interest. It is also the most difficult one to perform.

Preparation of the data is the key step on which the success of the entire knowledge discovery process depends.

Data mining and knowledge discovery are supported by several models that capture the characteristics of data. The most important ones include:

- **Summarization.** The goal is to characterize the data in an aggregated form via a small number of features/attributes.
- **Clustering or Segmentation.** The key objective is to find natural groupings (clusters) in highly dimensional data.
- **Regression Models**
- **Classification.** The classifiers can be regarded as a special case of regression models.
- **Concept Descriptions.** The goal is to come up with understandable descriptions of concepts/categories.
- **Dependency Analysis.** A concern with determination of relationships (dependencies) between fields in a database.
- **Sequence Analysis**. A process geared toward problems of modeling sequential data, such as time series.

Since human comprehension of numbers is very limited, we are interested in building more abstract, higher-level, concepts of, say, high values in some region of the image and low values in another. Thus, when we look at data in an attempt to discern patterns, we look at the data from a certain conceptual distance, formed by pieces of knowledge of a certain granularity. By information granularity we mean a form of encapsulation of numeric data into a single conceptual entity. For this purpose, windows of discovery are used that exhibit different levels of granularity. For instance, a single number descriptor has the highest level of granularity; an interval exhibits a lower granularity level, and then fuzzy set and rough set representations. The lowest level of granularity occurs for a set covering the entire space.

Finally, it is worth mentioning that most of the existing knowledge discovery systems are domain dependent [5], although there is a growing need for general knowledge discovery systems.

Knowledge Discovery Process

Several researchers defined a series of steps that provide a framework for a knowledge discovery process, in an attempt to facilitate diverse data mining applications. These frameworks may range from a few to several steps. For example, four steps (CDG Consulting Group) were defined as: Data Capture, Problem Understanding, Modeling, and Implementation. At the other end, nine very detailed steps [5] were defined as: Understanding Application Domain/Identifying Goal of KDD Process, Creating a Target Data Set, Data Cleaning and Preprocessing, Data Reduction and Projection, Matching Goals to a Particular Data Mining Method, Exploratory Analysis/Model and Hypothesis Selection, Data Mining, Interpreting Mined Patterns, and Acting on the Discovered Knowledge. Below, we outline a six-step knowledge discovery process that is similar to the one defined in [6]. However, its authors do not use the notion of a knowledge discovery process as defined above and in [1] and [5]. The steps are:

1. *Understanding the medical problem domain*.

In this step, we work closely with medical professionals to define the problem and determine medical objectives, identify key people, and learn about current solutions to the problem. That might call

for a high-level description of the problem with its requirements and restrictions as well as determination of success criteria from the medical point of view. After that, a domain-specific terminology is learned. A key subgoal in this step is determination of data mining goals and their success criteria. The goals are obtained by translating medical goals into data mining goals. Finally, a project plan identifying critical steps is usually prepared.

2. *Understanding the data.*

This includes collection of initial data (a sample from existing data). A list is made as to which data will be needed (including format and size), possibly including ranking of data attributes. Next, initial data exploration is performed to answer some of the data mining goals, confirm initial hypotheses, or result in asking for new features/attributes. Finally, verification of data is performed to answer questions about data completeness (all cases should be covered), redundancy, identification of erroneous data, missing attributes, checking plausibility of values of each attribute, etc.

3. *Preparation of the data.*

This is the key step on which the success of the entire knowledge discovery process depends. It usually takes at least half of the entire project effort. In this step, we decide which data will be used as input for data mining methods. Inclusion or exclusion of some data needs to be justified. Significance and correlation tests are performed as well as sampling of the data from the database. Next, the selected data need to be cleaned. This includes correcting, removing or ignoring noise, deciding how to deal with special values, etc. New data can be constructed through operations such as derivation of new attributes, such as discretization, transformation of some attributes, etc. The resulting data need to be integrated by creating new records from the constructed data and aggregation of information (information granularization). Finally, the data can be reformatted to meet requirements of the specific data mining methods.

4. *Data mining.*

This is another key step in the knowledge discovery process. Although it is the data mining methods that discover new information, it usually takes less time than preparation of data. Data mining involves selection of data modeling techniques, deciding on training and test procedures, building the model itself, and assessing its quality.

The best rule is the one that covers the most positive examples and covers the fewest negative examples.

Data mining methods include many types of algorithms, such as rough sets, fuzzy sets, Bayesian methods, evolutionary computing, machine learning, neural networks, clustering, preprocessing techniques, etc. Rough sets and fuzzy sets concentrate on representation of data at the nonnumeric level. Rough sets concentrate on an enriched set theoretical apparatus, while fuzzy sets are concerned with capturing a notion of a continuous transition between complete membership and complete nonmembership. Both deliver some level of summarization (granularization) of data, with intent to place all activities of data mining within the scope implied by hints from the user. Bayesian methods are based on the assumption that a classification problem can be expressed in probabilistic terms. Evolutionary computing can be seen as an optimization method driven by a biological principle of the survival of the fittest. Machine learning is aimed at revealing the relationships in data. The result of such summarization is provided in the form of decision trees or production rules. Neural networks are geared toward processing numeric data and building nonlinear relationships between inputs and outputs. Clustering is a key unsupervised learning technique revealing natural groupings in the data. Finally, the goal of broadly understood preprocessing is to reduce the dimension of the data. Thorough descriptions of all these methods within the framework of the knowledge discovery process can be found in a book by Cios, et al. [1].

5. *Evaluation of the discovered knowledge.*

This includes understanding the results, checking if the new information is novel and interesting, medical interpretation of the results, and checking their impact and the project goal. Approved models (results of different data mining methods) are kept. The entire knowledge discovery process is then revisited to identify failures, misleading steps, and alternative actions that could have been taken. Determination of possible actions can include their ranking, selection of the best ones, and documenting reasons for their choice.

6. *Using the discovered knowledge.*

Use is entirely in the hands of the owner of the database. There should be a plan for implementation and monitoring of the results as well as identification of problems associated with implementation. A final report is written to summarize the project outcome. Lessons learned in the current project should be noted for future applications.

The knowledge discovery process described will be illustrated in the following sections using bull's-eye SPECT images. It is also illustrated in our other work using a much larger database of SPECT images of the heart [7].

Understanding the Medical Problem Domain

Recognition of coronary artery disease is usually based on typical clinical symptoms (like pain) with support of the ECG stress test. There are situations, however, when clinical symptoms are not clear and the ECG stress test is vague or nondiagnostic. A nondiagnostic ECG stress test occurs when, during rest ECG, there are changes that make interpretation impossible; e.g., in patients with bundle branch blocks, pacemakers, co-existing valvular diseases, or patients treated with certain drugs. Additionally, it is recognized that in women there is a high percentage of false positives (low specificity). In those situations, in order to correctly diagnose a patient, it is necessary to perform tests such as echocardiography and myocardial perfusion scintigraphy.

Echocardiography, combined with stress or pharmacological stress, allows for identification of regions with abnormal kinetics caused by underperfusion. Echocardiography, however, is limited by difficulties in obtaining good quality images because of the construction of the patients' chests, obesity, or pulmonary emphysema. In myocardial perfusion

scintigraphy, combined with stress or pharmacological tests, the basis for diagnosing coronary artery disease is to find regions of the left ventricle with decreased radioisotope uptake. If these noninvasive methods fail to allow for a diagnosis, one resorts to coronography. In some rare cases, intravenous ultrasonography is also performed to obtain a detailed assessment of the shape and location of the plaque narrowing the arterial lumen.

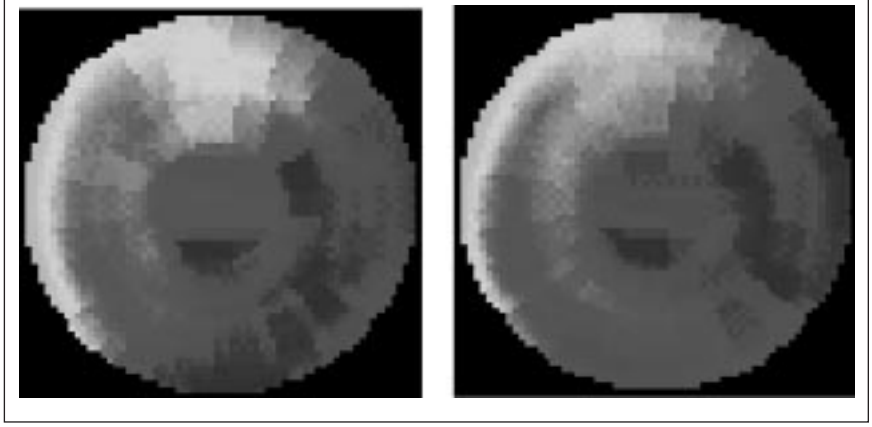
It should be added that the noninvasive methods are used not only in a typical diagnostic process but also for assessment of the severity of known coronary artery disease. The largest clinical problem is posed by patients for whom a decision is to be made about the timing and method of nonpharmacological treatment, depending on the severity of disease. For those patients, SPECT perfusion scintigraphy is an excellent method of evaluation, which precedes or temporally excludes the invasive approach.

Understanding the Bull's-Eye SPECT Data

SPECT is a three-dimensional (3-D) imaging technique used in diagnostic nuclear medicine that deals with single-photon emitters [8]. Another 3-D technique, PET, deals with emitters of positrons, which cause the emission of a pair of photons [9]. In SPECT, the patient is injected with a radioactive tracer that emits gamma rays. The radioactive tracers are made by attaching a radioactive molecule (photon emitter) to some biological molecule of interest. The type of molecule chosen depends on the type of information required. In the case of myocardial perfusion imaging, radioactive tracers that collect in the heart muscle according to regional myocardial blood flow are used. The radioactivity is measured by a gamma camera, which is made up of one-, two-, or three heads that rotate around a center axis (parallel to the long axis of the patient's body) within an angle of 180 or 360 degrees. The information measured by the camera from a number of views around the patient is used to reconstruct a 3-D image of the tracer distribution. The pixel count value in the image is related to the concentration of the radionuclide and thus to the amount of blood flow.

Since no accurate quantitative information about the activity distribution is available in SPECT, because of physical phenomena such as radiation attenuation, scatter of radiation, "partial volume effect," etc. [10], only qualitative (visual) or semi-quantitative data analysis methods are used. The relative values from different regions of interest are compared to the maximum value in the heart. The most common approach relies on qualitative (visual) inspection of tomographic images (slices perpendicular and parallel to the long axis of the left ventricle of the heart), and—if available in medical software—on qualitative or normative evaluation of the bull's-eyes perfusion maps. Bull's-eyes perfusion maps contain the integrated information from all the tomographic slices achieved from rest and stress heart studies and are presented in the standardized form (patient-independent size and orientation). Normative evaluation of these maps relies on comparison of the bull's-eye images of diagnosed patient with the "normal" bull's-eye maps generated from the population of healthy people. "Normal" maps contain the lower limit of the normal perfusion values, usually defined as a mean diminished by 2.5 standard deviations, in values achieved for a given region of the map [11-15].

In this article, we show how to derive quantitative information from regions of interest from bull's-eye SPECT images, and we use it for the purpose of automating the diagnosis. The means of doing this is to extract pertinent data, in terms of pixel counts, and then generate rules for normality and abnormality by using the CLIP3 machine learning algorithm [16, 17].



1. STRESS (left) and REST (right) images of a normal patient.

Preparation of the Bull's-Eye SPECT Data

Radioisotope Data Collection

The following protocol of exercise and rest is used to acquire bull's-eyes SPECT images of the patient. The radiopharmaceutical used is Tc-99m-MIBI [18, 19], applied according to two-day stress-rest protocol. One day at peak exercise, a dose of 0.3 mCi/kg of Tc-99m-MIBI is injected, and exercise is continued for another minute. The acquisition starts approximately 45 to 60 min after injection of the Tc-99m-MIBI. This is the stress condition, which results in STRESS images. The next day, another dose of 0.3 mCi/kg of Tc-99m-MIBI is injected at rest. The acquisition starts approximately 60 to 120 min after injection. This is the rest condition, which results in REST images.

Patients were imaged with a single-head rotating gamma camera. A total of 64 projections (approximately 20 sec per projection) were obtained over a semicircular 180° orbit, from a 45° right anterior oblique (RAO) to a 45° left posterior oblique (LPO) position. All projection images were stored using a 64×64 pixel, 16-bit matrix. Each of the projections was corrected for nonuniformity and the mechanical center of rotation. Then, filtered backprojection was performed using the optimum filter determined for the given computer (1-D Butterworth filter of 7th order, with cutoff frequency equal to 0.4 of Nyquist frequency). Interslice weighting was applied using three-point smoothing with weights 0.5-1-0.5. All transverse tomograms were reconstructed at a thickness of 1 pixel per slice, with this thickness representing 6.2 ± 0.2 mm. After marking left ventricle axis space orientation, the volume of transverse tomograms was reoriented, and sets of slices perpendicular to the long axis ("obliques") and slices parallel to the long axis ("oblique/coronals" and "oblique/sagittals") were created. From "obliques"

the summary polar perfusion maps ("bull's-eyes," in stress and rest) were generated. Figure 1 shows SPECT maps of a normal patient, while Fig. 2 shows maps of an abnormal patient with single-vessel disease.

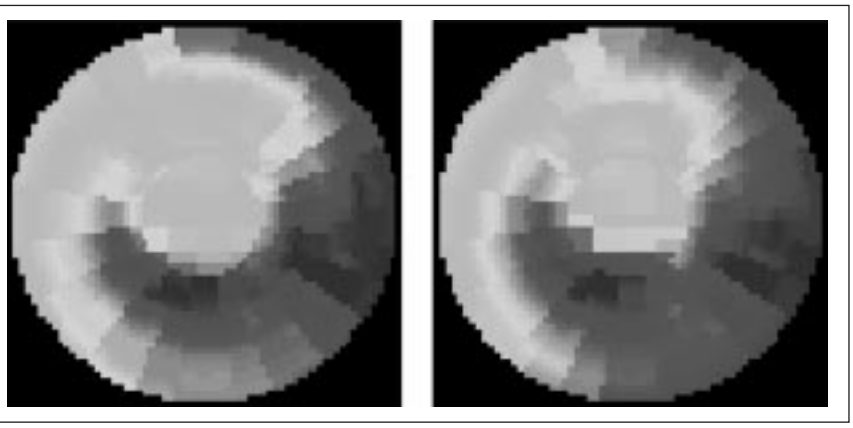
Data Cleaning

Two groups of patients were selected for the study: patients with confirmed significant coronary artery disease (by contrast coronarography) and patients with low probability of coronary artery disease (with or without contrast coronarography performed). The selecting condition for patients was to have the bull's-eyes maps, which visually did not show any discrepancy with short-axis and long-axis tomographic slices.

Vessel narrowing ("disease of the vessel") was defined as significant if at least 70% of the vessel diameter was occluded by visual assessment during contrast coronarography. Only cases of single-vessel disease (SVD) or dual-vessel disease (DVD) were considered. Three-vessel disease, without knowledge of contrast coronarography, is often misinterpreted in perfusion studies as DVD, SVD, or even "normal" (if all three vessels are narrowed similarly), since the maximum in perfusion images corresponds to the best perfused areas, and not to correctly perfused areas. For groups with SVD or DVD, only those patients were selected for whom:

- myocardial infarction was excluded (since the images after infarction usually show significant perfusion defects and are easy to interpret), and
- full protocol of contrast coronarography performed within three months of stress SPECT, was available.

The SVD group consisted of five females and six males and the DVD group consisted of 13 males. The patients were selected from about 7000 stress-rest Tc-99m-MIBI SPECT studies registered in the Institute of Cardiology, Warsaw, during 1991 through 1997 (approximately 20 patients were studied every week). In most cases (about 70%), patients were referred to the Institute after at least one myocardial infarction. The patients without myocardial infarction rarely have the contrast coronarography performed before the SPECT radioisotope studyand hardly ever after normal results of SPECT. Unfortunately, the group of patients of interest had to be selected only from the latter patient data.



2. STRESS (left) and REST (right) images of an abnormal patient with LAD single-vessel disease.

The reference "normal" groups for female and male populations were created from patients with low likelihood of coronary artery disease ($p \leq 10\%$). A post-test likelihood was estimated after an electrocardiographic stress test, according to age, sex, symptoms, and depression of the S-T segment [20]. Additionally, the SPECT perfusion images must be visually "normal" in these patients. Patients after myocardial infarctions, percutaneous transluminal coronary angioplasty (PTCA), or coronary artery bypass grafts (CABG) were excluded. As a result, the female group consisted of 75 patients, age19-64 years (avg. 47 ± 9); 10 of them had earlier coronarography that showed no significant changes in coronary vessels. The male group finally consisted of 86 patients, age 30-67 years (avg. 48 ± 10); seven of them had earlier coronarography that showed no significant changes in coronary vessels. As a result of the above data-cleaning operations, our target database consisted of 185 bull's-eye SPECT images, divided as follows:

- 161 NORMAL (86 males + 75 females)
- 11 SVD (three females [LAD] + four males [RCA] + two females [RCA] + two males [LCX]), where LAD means left anterior descending branch of left coronary artery, LCX is left circumflex branch of left coronary artery, and RCA is right coronary artery
- 13 DVD (eight males [LAD and RCA] + four males [RCA and LCX] + one male [LAD and LCX])

Our goal was to show that a machine learning algorithm can generate rules distinguishing between normal and abnormal patients, in spite of the fact that the ratio of normal to abnormal patients (161/24) was not very favorable.

Creation of a Database

First, a database storing original bull's-eye SPECT image files and other pertinent patient information was created. Next, secondary data was derived in four steps:

1) Reading and processing data files into correct image format, using Matlab's Image Processing Toolbox. All patient data are in the form of Windows bitmap files in MS Access.

2) Creation of a template indicating the boundaries of the coronary vascular territories on the quantitative polar map.

3) Extraction and quantification of the regions of interest (ROIs) obtained by superimposing the template on the images.

4) After quantitative data were extracted from the regions of interest, new discrete data were created through equalization discretization and stored in the database.

The fields in the database were defined as follows:

- Patient ID: Unique key and number field
- Weight: Number field
- Height: Number field
- Hypercholesterolemia: Text-list item (yes/no)
- Smoking: Text-list item (yes/no)
- Lifestyle: Text-list item (active/sedentary)
- Date of Bull's-Eye SPECT: Number
- Age: Number
- Diagnosis: Text-list item (Normal, SVD: LCX, SVD: RCA, SVD: LAD, DVD: RCA + LCX,

DVD: LAD + RCA, DVD: LAD + LCX)
- Sex: Text-list item (male/female)
- Stress Image: OLE object
- Delayed Image: OLE object

The database was designed to interface with the user so that new patient data can be easily entered. It has two queries: one for retrieving female patient data and one for retrieving male patient data.

The second step was to create a template to be superimposed on the images to outline three coronary vascular territories, namely, LAD, LCX, and RCA [21]. Figure 3(a) shows a 64 × 64-bit map image of the template created to outline the coronary vascular territories of the bull's-eye SPECT image. It closely approximates the template shown in Maddahi, et al. [21]. To get the boundaries on the original bull's-eye SPECT images, a program was written in Matlab, that reads the template (bitmap image shown above) and the original image (64 × 64-bit) from a file and converts them into matrix format. The template matrix is such that it has zeros along the black outline of the template and ones on the rest of the image. This matrix, when multiplied with the original image matrix, gives zeros along the boundaries, and the rest of the image is left intact, since it is multiplied by ones. The resulting sample image is shown in Fig. 3(b).

The third step was to extract ROIs, as outlined by the template. The method used to extract the ROI is similar to the one outlined above. Templates are made for each ROI and the template matrix is multiplied (with ones for the ROI and zeros for the rest of the image) with the original image matrix.

Once the ROIs are extracted, the next task is to quantify the regions. To quantify the data, we first calculate the normalized count ratio for each pixel (average value per pixel) in the three regions, using all normal patients. Of interest is the percent of pixels in each region that are more than 2.5 standard deviations from the average for each of the six regions. So for each image, we calculated the percentage of pixels that are more than 2.5 standard deviations from average for that particular region. We assumed that the values this far from the normal region are associated with blockage in the arteries. Only those values will be used for generation of rules. As a result, three features for the STRESS images and three features for the REST images were obtained. The six features were enumerated as follows: F1 = LAD stress, F2 = LCX stress, F3 = RCA stress, F4 = LAD rest, F5 = LCX rest, and F6 = RCA rest.

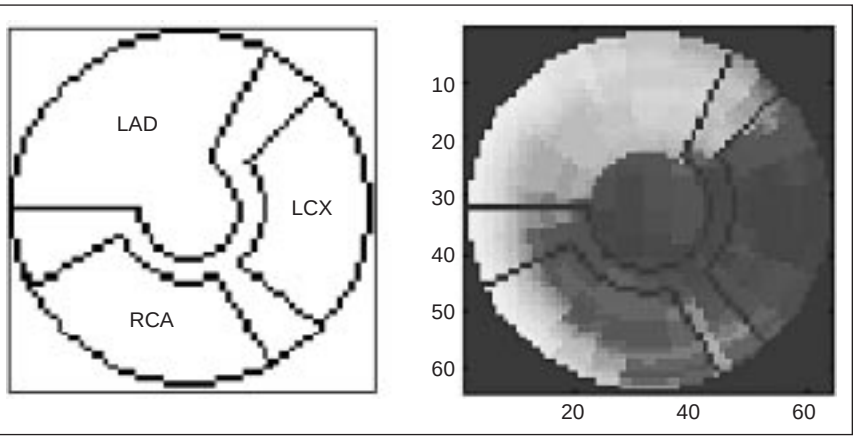
The fourth step was to discretize the six-feature continuous data, since most machine learning algorithms operate on discrete data. To do this, histogram equalization was performed for each region (i.e., LAD, RCA, LCX) for all (normal and abnormal) patients. Although equalization was done in several intervals, the results shown below are for 10 equally spaced intervals/bins. Then, the bins were associated with digits 2 through 11, depending on which bin the value fell into. Since most of the patients in the database were normal, there was a major chunk of the data in the first bin.

Data Mining

Our data mining goal was to come up with an outcome that provides an answer to the "why" question; i.e.: why do images correlate with ischemic regions? The methods that can be used for this purpose include machine learning, fuzzy logic, and genetic algorithms. They can generate, as their output, production rules of the "IF-THEN" form. These rules are in striking contrast to a trained neural network, which provides the user only with the weights associated with the edges interconnecting the neurons. This is important, especially when heuristic features that make sense to the user are selected from the images. To illustrate the point, in our previous work 15 new diagnostic rules were discovered from nuclear medicine data, as opposed to 68 cardiologist-specified rules, to diagnose obstructions in major coronary arteries. These new rules, which make sense to cardiologists, were not previously known to them and were obtained by analyzing planar thallium heart images [22].

As a result, we decided to generate diagnostic rules that are understandable to a physician by using the CLIP3 machine learning algorithm of Cios, et al. [16, 17]. CLIP3 combines the concepts of decision trees and rule-based algorithms and generates multiple hypotheses to increase the chance of capturing the true meaning of the concept. In most machine learning algorithms, the user needs to specify positive and negative sets of examples. Thus, if normal patients are designated as positive, then the rest of the patients are designated as negative, or abnormal (and vice versa).

Briefly, the learning process in the CLIP3 algorithm consists of three phases. In phase 1, the positive examples are partitioned into subsets so that rules can be generated from these smaller subsets. The partitioning is done in a similar manner to growing a search tree; however, the tree is never stored in memory. This enables the algorithm to operate on a large number of examples. The goal of phase 2 is to determine which of the subsets from phase 1 are the best candidates for rule formation as well as what features the rule will use. In phase 3, the best rule, from the rules generated in phase 2, is chosen. The best rule is the one that covers the most positive examples and covers the fewest negative examples. The examples covered by the best rule are then eliminated from the positive training data set, and the remaining positive examples are used to learn the next diagnostic rule by repeating phases 1 and 2. For details see [1], [16], and [17].



3. Template outlining the three regions (left) and a bull's-eye SPECT image with the regions superimposed on it (right).

Discovered Knowledge

Data from 160 normal patients and 24 abnormal patients were used to generate the rules. One normal patient data was dis-

carded at this stage because the image values were very different when compared with the values of the other normal patients. The CLIP3 thresholds were used as follows: Stop threshold was 0, which means the algorithm stops only after all the positive examples are covered. The best rule threshold was 0.5, which means only those rules in phase 2 that have covered at least 50% of remaining positive examples are chosen as best rules. All of the rules shown below cover only positive examples and none of the negative (and vice versa).

First, all normal patient data were treated as positive examples (matrix of size 160 × 6) and abnormal patient data were treated as negative examples (matrix of size 24 × 6). The generated rules were as follows:

R1: F1 = <2> F2 = <2 3 5> F3 = <2 3>
F4 = <2> F6 = <2 3 6>
(covers 101 patients)

Let us repeat the meaning of the six features: F1 represents the LAD region, F2 represents LCX, F3 represents RCA in the stress test, F4 represents LAD region, F5 represents LCX, and F6 represents RCA in the rest test. The rule says that if the value of feature one (F1) is 2 AND the value of F2 is 2 or 3 or 5 AND the value of F3 is 2 or 3 AND the value of F4 is 2 and the value of F6 is 2 or 3 or 6, then a patient can be classified as normal. This is a very "strong" rule since it correctly classifies, or "covers," a large majority (101) of normal patients. Other rules are to be read in a similar way:

R2: F2 = <2 4 5 7> F4 = <2 3> F6 = <2 4 6>
(covers 75 normal patients)
R3: F2 = <3> F3 = <2 4>
(covers 22 normal patients)
R4: F1 = <3> F4 = <2 3> F5 = <3>
(covers 21 normal patients)
R5: F1 = <3> F2 = <2> F3 = <3 4> F6 = <3 6>
(covers 8 normal patients)
R6: F1 = <2> F3 = <2> F4 = <3>
(covers 6 normal patients)
R7: F1 = <3> F2 = <3> F4 = <2> F6 = <2 4 6>
(covers 6 normal patients)
R8: F1 = <2> F2 = <3> F4 = <3>
(covers 3 normal patients)

Next, abnormal patients were used as positive examples (matrix of size 24 × 6) and normal patients as negative examples (matrix of size 160 × 6). The reason for doing this was to see what rules recognizing abnormal patients would be generated. The generated rules were as follows:

We showed that it is possible to differentiate the patients with coronary artery disease from the patients with low probability of the disease.

R1: F1 = <2 4 5 6 7 11> F2 = <5 6 7 8 9 11>
F4 = <2 3 4 5 6 7 8 9 10 11 >
F5 = <3 4 5 6 7 8 9 10 11 >
(covers 10 abnormal patients)
R2: F3 = <3 6 7 8 9 10 11>
F4 = <2 4 5 6 7 8 9 10 11>
F5 = <3 4 5 6 7 8 9 10 11>
F6 = <4 5 7 8 9 10 11>
(covers 7 abnormal patients)
R3: F1 = <3 4 5 6 7 11> F5 = <4 5 6 7 8 9 10 11>
(covers 6 abnormal patients)
R4: F1 = <2 4 5 6 7 11> F2 = <4 6 8 9 11>
F6=<3 7 8 9 10 11>
(covers 5 abnormal patients)
R5: F1 = <3 4 5 6 7 11> F2 = <3 6 8 9 11>
F4 = <2 4 5 6 7 8 9 10 11>
F5 = <2 5 6 7 8 9 10 11>
F6 = <3 7 8 9 10 11>
(covers 2 abnormal patients)
R6: F2 = <2 6 8 9 11> F6 = <5 7 8 9 10 11>
(covers 2 abnormal patients)

Although the algorithm should be run separately on male and female patient data (as normal limits are different), we could not do so because our population of abnormal female patients was too small (only five). Thus, we concentrated on the male patients data set, which consisted of 86 normal patients and 19 abnormal patients. Using normal patients as positive examples and abnormal patients as negative examples, the generated rules were as follows:

R1: F1 = <2> F2 = <2 3 5> F3 = <2 3>
F6 = <2 3>
(covers 53 normal patients)
R2: F2 = <3 4> F6 = <2>
(covers 21 normal patients)
R3: F5<2> F6<2 4>
(covers 27 normal patients)
R4: F4 <3> F5<3>
(covers 16 normal patients)
R5: F1<3> F4<2> F5<3>
(covers 5 normal patients)

When abnormal male patients were used as positive examples (matrix of size 19 × 6) and normal male patients as negative, the generated rules were as follows:

R1: F2 = <3 6 7 8 9 11> F5 = <4 5 6 11>
(covers 8 abnormal patients)
R2: F4 = <2 4 5 6 8 11> F5 = <3 4 5 6 11>
F6 = <4 5 11>
(covers 7 abnormal patients)
R3: F1 = <2 4 5 6 7 11> F2 = <4 6 8 9 11>
F3 = <3 11>
(covers 3 abnormal patients)
R4: F1 = <3 4 5 6 7 11> F3 = <3 4 11>
F4 = <2 4 5 6 8 11> F5=<2 5 6 11>
F6 = <3 5 11>
(covers 3 abnormal patients)

Evaluation of the Discovered Knowledge

Let us look at the rules generated for normal patients (the first experiment where abnormal patients were used as negative examples). Rule R1, which covers the large majority of normal patients, uses five features (all three in stress and LAD and RCA in rest), with all the features values being low—as expected for normal patients. The remaining rules, R2 through R8, also use low values for the few features they use. These rules, in particular R1, use both rest and stress information. The exception is R3, which uses only LCX and RCA in stress information. These rules make sense to an interpreting cardiologist, who also compares stress and rest values for all regions (features F1 through F6).

The rules generated for abnormal vs. normal patients also use information from both stress and rest tests. Now, however, they use a combination of a few low values with many high values of the features. This combination fully agrees with a diagnosing cardiologist, who looks for re-

duced perfused regions of the left ventricle.

Rules generated for normal male patients, R1 through R5, are more interesting. The strongest rule, R1, which covers 53 out of 86 normal patients, uses three stress and one rest feature (similar to what rule R1 found for all the normal patients. However, two rules, R3 and R4, use only features from the rest test. This set does not have an easy clinical interpretation, as it is difficult to accept that patients with normal rest perfusion in two territories can be classified as normal (compare all the totally reversible defects, which are characteristic for CAD without infarction).

The last set of rules, covering abnormal male patients, R1 through R4, also has a strong rule (R2), which uses only information from the rest test. This is not surprising, as a rest image with abnormalities in all three territories must correspond to abnormal perfusion.

The above method of classification of the myocardial perfusion SPECT images seems to be an interesting alternative to the existing "quantitative" (normative) methods of assessment of the bull's-eye maps [11-15].

The clinical value of any diagnostic method is usually evaluated on the basis of sensitivity and specificity in relation to a reference method. For detection of coronary artery disease by radioisotope myocardial perfusion studies, it is contrast angiography that usually serves as a reference method, despite the fact that it is based on anatomy, as opposed to the functional character of radioisotope methods. All of the generated rules, both for recognizing normal and abnormal patients, will be clinically tested on new patient data at the Institute of Cardiology in Warsaw. We will establish the sensitivity and specificity of the rules, since we decided that our data set was relatively too small to divide it into training and test data sets at the present time.

The knowledge discovered about myocardial perfusion in the form of the rules presented above allows for two simple complementary approaches to influence the sensitivity and specificity of diagnosis. The first approach is similar to the one used in available "quantitative" myocardial perfusion SPECT software: the lower threshold of normal limit (accepted by us as average diminished by 2.5 standard deviation) can be defined "freely" as a product of standard deviation and a number from the interval 0–3; the larger the number, the lower the specificity and the higher sensitivity is achieved. The second approach is unique: it lies in the possibility of free configuration of the discovered rules to reach the desired values of sensitivity and specificity.

Discussion and Conclusions

Visual assessment of clinical diagnostic images is observer-dependent. Thus, much effort is expended to computerize the process of diagnosis so it is less dependent on the observer, especially when the observer is not experienced. A large number of images to be evaluated (as in SPECT myocardial perfusion studies: approximately 15 oblique "slices," 15 oblique/sagittal, and 15 oblique/coronal, both in stress and rest, which comes to nearly 100 2-D images per patient) forced the creation of more "comprehensive" images; namely, the bull's-eye perfusion maps. Using these maps, we showed that it is possible to differentiate the patients with coronary artery disease (one- or two-vessel) from the patients with low probability of the disease (normals). In the future, features other than those used in this work will be used; for instance, a feature representing the area of "abnormal" myocardium, available in most previously mentioned algorithms for "normative" evaluation of bull's-eye maps. In the course of this work, we also came up with methods that can accurately extract the ROIs from an image where a thresholding method cannot be used.

Acknowledgments

Dr. Cios kindly acknowledges his support from the Fulbright Foundation to the Institute of Cardiology in Warsaw. The authors also thank Mr. W. Konieczny for using his data-conversion program.

Krzysztof J. Cios is a professor and interim chair of the Department of Bioengineering and Professor of Electrical Engineering and Computer Science at the University of Toledo. He is also an adjunct professor of medicine at the Medical College of Ohio. His research interests are in the areas of biologically inspired learning algorithms, in particular artificial neural networks and machine learning, and in data mining and knowledge discovery. NASA, NSF, American Heart Association, Ohio Aerospace Institute, and NATO have funded Dr. Cios's research. He has published two books, 35 journal articles, nine book chapters, and 48 peer-reviewed conference papers. He serves on the editorial boards of *Neurocomputing*; *IEEE Transactions on Systems, Man and Cybernetics*; *IEEE Engineering in Medicine and Biology Magazine*; and *Mathematical Neuroscience*. Dr. Cios has been the recipient of the Norbert Wiener Outstanding Paper Award, the Neurocomputing Best Paper Award, the University of Toledo Outstanding Faculty Research Award, and the Fulbright Senior Scholar Award. He received his Ph.D. from the University of Mining and Metallurgy, Krakow; his M.B.A. from the University of Toledo; and his D.Sc. from the Systems Research Institute, Polish Academy of Sciences. Dr. Cios is a member of the IEEE, AAAI, and Sigma Xi.

Anna Teresinska received the master's degree in physics from the University of Warsaw and the Ph.D. in medical sciences from the Medical Academy of Lódz. From 1981 to 1986, she was an assistant in the Department of Nuclear Medicine at the Postgraduate Medical Center in Warsaw. In 1986, she started working in the Department of Nuclear Medicine at the Institute of Cardiology in Warsaw, and since 1994 she has served as the head of the department. She is a member of the Board of the Polish Society of Nuclear Medicine, a member of the European Association of Nuclear Medicine, and a member of the Polish Cardiac Society. In 1998, she was appointed editor-in-chief of *Problems of Nuclear Medicine* (a journal of the Polish Society of Nuclear Medicine). Since 1995, she has been a chairman of the Commission of Nuclear Medicine of the Committee of Medical Physics and Radiobiology at the Polish Academy of Sciences. Her current research activity concentrates on optimization of nuclear cardiology methods, with main clinical application in evaluation of the viability of myocardium.

Stefania Konieczna received her M.S. degree in electronics engineering from the Technical University in Warsaw. She has worked in the Department of Nuclear Medicine at the Institute

of Cardiology in Warsaw as a specialist for informatics since 1986. She is a member of the Polish Society of Nuclear Medicine and the Association of Polish Electricians. Her current research interests include methods of registration and processing of medical images.

Joanna Potocka received her M.D. from the Medical Academy in Warsaw, and later the licence for clinical practice. Between 1983 and 1996, she worked in cardiology clinics at two large hospitals in Warsaw (including intensive care units). During those years she completed first and second degrees of specialization in internal medicine and specialization in cardiology. Dr. Potocka is interested in different methods of diagnostics used in cardiovascular diseases, including heart scintigraphy. In 1997, she joined the Institute of Cardiology in Warsaw (Department of Nuclear Medicine) where she performs heart investigations using radioisotopic techniques that are most commonly SPECT procedures. Her research interest concentrates on coronary heart disease in women before menopause and diagnostics of atypical symptoms of heart diseases. She is a member of the Polish Cardiac Society and the Polish Society of Nuclear Medicine.

Sunil Sharma is an M.S. student in the Department of Bioengineering at the University of Toledo.

Address for Correspondence: Krzysztof J. Cios, Department of Bioengineering, University of Toledo, Toledo, OH 43606-3390. E-mail: kcios@eng.utoledo.edu

References

1. **Cios KJ, Pedrycz W, and Swiniarski R:** *Data Mining Methods for Knowledge Discovery*. Norwell, MA: Kluwer, 1998.

2. **Frawley WJ, Piatetsky-Shapiro G, and Matheus C.J:** Knowledge discovery in databases: An overview. In: Piatetsky-Shapiro G, Frawley WJ (Eds): *Knowledge Discovery in Databases*. Cambridge, MA: AAAI/MIT Press, pp. 1-27, 1991.

3. **Zadeh LA:** Fuzzy sets and information granularity. In: Gupta MM, et al. (Eds): *Advances in Fuzzy Set Theory and Applications*. Amsterdam: North Holland/Elsevier, pp. 3-18, 1979.

4. **Pawlak Z:** Rough classification. *Int J Man-Machine Studies*, 20: 469-483, 1984.

5. **Fayyad UM, Piatetsky-Shapiro G, Smyth P, and Uthurusamy R:** *Advances in Knowledge Discovery and Data Mining*. Cambridge, MA: MIT Press, 1996.

6. **Chapman P, Clinton J, Khobaza T, Reinartz T, and Wirth R:** The CRISP-DM process model. CRISP-DM Consortium, 1999.

7. **Sacha JP, Cios KJ, and Goodenday LS:** Issues in automating cardiac SPECT diagnosis. This issue, pp. 78-88.

8. **Jaszczak RJ and Coleman RE:** Single photon emission computed tomography (SPECT): Principles and instrumentation. *Invest Radiol* 20: 897-910, 1985.

9. **Jacobson HG:** Instrumentation in positron emission tomography. *JAMA* 259: 1531-1536, 1988.

10. **Blokland JAK, Reiber JHC, and Pauwels EKJ:** Quantitative analysis in single photon emission tomography (SPET). *Eur J Nucl Med* 19: 47-61, 1992.

11. **Garcia EV, Van Train K, Maddahi J, et al.:** Quantification of rotational thallium-201 myocardial tomography. *J Nucl Med* 26: 17-26, 1985.

12. **Van Train K, Berman DS, Garcia E, et al.:** Quantitative analysis of stress Tl-201 myocardial scintigrams: A multicenter trial validation utilizing standard normal limits. *J Nucl Med* 2717-2725, 1986.

13. **DePasquale E, Nody A, DePuey G, et al.:** Quantitative rotational thallium-201 tomography for identifying and localizing coronary artery disease. *Circulation* 77(2): 316-27, 1988.

14. **Van Train KF, Maddahi J, Berman DS, et al.:** Quantitative analysis of tomographic stress thallium-201 myocardial scintigrams: A multicenter trial. *J Nucl Med* 31:1168-79, 1990.

15. **Van Train K, Areeda J, Garcia E, et al.:** Quantitative same day rest-stress Tc-99m sestamibi SPECT: Definition and validation of normal limits and criteria for abnormality. *J Nucl Med* 34: 1494-1502, 1993.

16. **Cios KJ and Liu N:** An algorithm which learns multiple covers via integer linear programming, Part I - the CLILP2 algorithm. *Kybernetes* 24(2): 29-50, 1995. (www.mcb.co.uk/literati/outst97.htm#k)

17. **Cios KJ, Wedding DK, and Liu N:** CLIP3: Cover learning using integer programming. *Kybernetes* 26(4-5): 513-536, 1997.

18. **Wackers FJ, Berman DS, Maddahi J, et al.:** Technetium-99m hexakis 2-methoxyisobutyl isonitrile: Human biodistribution, dosimetry, safety, and preliminary comparison to thallium-201 for myocardial perfusion imaging. *J Nucl Med* 30: 301-11, 1989.

19. **Maddahi J, Kiat H, Van Train KF, et al.:** Myocardial perfusion imaging with technetium-99m sestamibi SPECT in the evaluation of coronary artery disease. *Am J Cardiol* 66: 55E-62E, 1990.

20. **Diamond GA and Forrester JS:** Analysis of probability as an aid in the clinical diagnosis of coronary-artery disease. *N Engl J Med* 300: 1350-8, 1979.

21. **Maddahi J, Van Train KF, Prigent F, et al.:** Quantitative single photon emission computerized thallium-201 tomography for the evaluation of coronary artery disease optimization and prospective validation of a new technique. *J Am Coll Cardiol* 14: 1689-99, 1989.

22. **Cios KJ, Shin I, and Goodenday LS:** Using fuzzy sets to diagnose coronary artery stenosis. *IEEE Computer Mag* 24(3): 57-63, 1991.