ELSEVIER

Artificial Intelligence in Medicine 9 (1997) 61–78

Artificial Intelligence in Medicine

Feasibility analysis of a case-based reasoning system for automated detection of coronary heart disease from myocardial scintigrams

Mojgan Haddad[a], Klaus-Peter Adlassnig[b], Gerold Porenta[a,*]

[a]*Department of Cardiology, University of Vienna, Währinger Gürtel 18-20, A-1090 Vienna, Austria*
[b]*Department of Medical Computer Sciences, University of Vienna, Spitalgasse 23, A-1090 Vienna, Austria*

Accepted 29 July 1996

Abstract

Myocardial perfusion scintigraphy is a noninvasive diagnostic method for the evaluation of patients with suspected or proven coronary artery disease (CAD). We utilized case-based reasoning (CBR) methods to develop the computer-based image interpretation system SCINA which automatically derives from a scintigraphic image data set an assessment concerning the presence of CAD. We compiled a case library of 100 patients who underwent both perfusion scintigraphy and coronary angiography to document or exclude the presence of CAD. The angiographic diagnosis of the retrieved nearest neighbor match of a scintigraphic input case was selected as the CBR diagnosis. We examined the effects of input data granularity, case indexing, similarity metric, and adaptation on the diagnostic accuracy of the CBR application SCINA. For the final prototype, sensitivity and specificity for detection of coronary heart disease were 98% and 70% suggesting that CBR systems may achieve a diagnostic accuracy that appears feasible for clinical use. Copyright © 1997 Elsevier Science B.V.

Keywords: Case-based reasoning; Automated image interpretation; Myocardial perfusion scintigraphy; Coronary artery disease

* Corresponding author. Tel.: +43 1 40400 4641; fax: +43 1 4081148; e-mail: gporenta@mail.kard.akh-wien.ac.at.

PII 0933-3657(96)00361-2

1. Introduction

1.1. Application domain

Cardiovascular diseases account for approximately 50% of the total mortality in industrial countries and thus have an enormous economic impact on health care systems. In particular, coronary artery disease (CAD) is one of the most important health problems associated with a high rate of morbidity and mortality. The atherosclerotic narrowing of coronary arteries may impede the blood flow to myocardial tissue and thus cause myocardial ischemia, myocardial infarction, and sudden cardiac death. Therefore, methods and techniques for the diagnosis and treatment of coronary artery disease play a major role in health care today.

Three main coronary arteries supply the heart with nutritive blood flow: the left anterior descending artery (LAD), the left circumflex artery (LCX), and the right coronary artery (RCA). When obstructive disease with a $\geq 70\%$ luminal narrowing develops in any segment of these vessels, a diagnosis of CAD is established.

Diagnosis of coronary artery disease may involve both invasive or noninvasive examination techniques. Selective coronary angiography is the reference standard for determining the presence, severity, and location of CAD but as an invasive examination method is associated with a minor risk of morbidity and mortality. Noninvasive techniques such as echocardiography, stress electrocardiography, and myocardial perfusion scintigraphy have a minimal risk, but are not as accurate as the invasive method.

SPECT (single photon emission computed tomography) perfusion scintigraphy is a diagnostic technique in nuclear medicine that permits measurement of the spatial or temporal distribution of a radioactive perfusion tracer. Myocardial SPECT perfusion imaging is capable of mapping the tracer distribution within myocardial tissue in three dimensional space and provides a 3D description of the regional myocardial perfusion as a sequence of tomographic 2D images. Human experts interpret scintigraphic images mostly by visual analysis with respect to the presence, severity, and location of CAD. While semiquantitative methods of image analysis have helped to reduce interobserver variability [12,18], image interpretation still depends on the expertise of the human observer and thus may vary between different readers. Computer-based methods that provide decision support for the interpretation of myocardial perfusion scintigrams may help to reduce interobserver variability and increase diagnostic accuracy.

1.2. Case-based reasoning

Case-based reasoning is a technique of artificial intelligence that attempts to solve a given problem within a specific domain by adapting established solutions

to similar problems. Case-based reasoning appears particularly promising as a method for building image interpretation systems, because it is simple to implement for pattern recognition tasks and seems to mirror the natural structure of expertise [10,15]. Knowledge is acquired by collecting cases from experienced human experts. Each case is defined by a set of features and values which describe both structure (context) and information (content) of a piece of given experience. The context of a case is created by the developer through definition of a set of features and feature value types, while the actual content of a case is introduced by the expert.

The reasoning process in case-based reasoning systems can be regarded as a combination of guided retrieval and usage of prior experiences to help solving a new problem. A case-based reasoner can either perform the retrieval process and leave the comparison of the retrieved results to the user, or the CBR application can also assist the user in reasoning by providing adaptation. Because old situations are hardly ever exactly identical to a new one, old solutions must usually be adapted to fit the new situation. In the adaptation process, a solution that is not entirely appropriate should be modified to optimize the applicability to the new problem [2].

1.3. Specific problem

The specific aim of this study was to develop and evaluate the prototype of an automated image interpretation system for the classification of tomographic myocardial perfusion scintigrams with respect to the presence of significant coronary artery disease. In particular, we examined the feasibility of using a case-based reasoning system in clinical practice by optimizing diagnostic accuracy through system design and system specifications.

2. Methods

2.1. Image processing

An automated system for the interpretation of myocardial perfusion images critically depends on the capability to depict the myocardial tracer distribution across standardized image planes [1]. As SPECT imaging generates transaxial image planes that are oriented according to the body axis, image reorientation needs to be performed to generate images oriented according to the axis of the heart. Thus, we developed a Macintosh-based image processing program to generate standardized views of tracer activity and to derive an input data set for the case-based reasoner from standardized reoriented images [7]. The OSIRIS software [17] which was developed at the University Hospital of Geneva as a general platform for image display and manipulation in a window-based graphical user interface environment was extended to perform the tasks of image reorientation and data extraction.

The output of the reorientation procedure for the present study was a sequence of 6 equidistant short axis planes as indicated in Fig. 1 (pixel size: 3.4 mm, image matrix: $\sim 60 \times 60 \pm 10 \times 10$ depending on heart size). This data set was too large for further processing, so that data had to be compacted to extract the desired information from the digitally prepared and reoriented images. A frequently used approach for data extraction is to visualize the 3D-distribution of myocardial perfusion as 2D polar maps. These polar maps are constructed by mapping sequential maximal count circumferential profiles from the apex to the base of the left ventricle into successive rings on the polar map [12] as shown in Fig. 1. Six short axis planes were used reaching from the base of the left ventricle to the endocardial border of the apex. The raw polar maps assembled from the reoriented myocardial perfusion images were subsequently normalized to eliminate scaling effects of extracardiac activity and prepare the data for use in the automated interpretation system.

For the case-based reasoner, the polar map data in the form of 6×60 8-bit integer numbers is still too large and needs further simplification. In order to determine the level of detail that needs to be included for the case description, we derived from the original polar maps four different matrices of segmental count

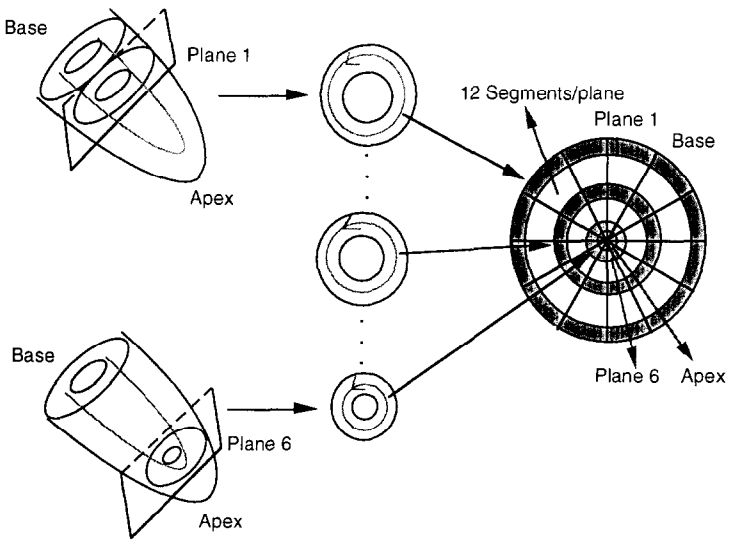

Fig. 1. Schematic drawing illustrating the generation of a polar map from short axis perfusion scintigrams. Circumferential profiles of maximal count activities are obtained from each of six equidistant short axis planes and depicted as concentric rings on the polar map. The apical short axis plane is mapped into the center of the polar map while the most basal slice appears at the outside. Twelve 30° segments are outlined on the polar map.

activities by varying the number of included planes and the number of segments per plane as follows: (number of planes) × (number of segments)...2 × 6, 4 × 6, 6 × 6, 6 × 12. These data matrices were stored in an ASCII file and used to determine the impact of data granularity on diagnostic performance of the CBR.

2.2. Case library

A case library of 100 patients was compiled who underwent both stress perfusion scintigraphy and coronary angiography to document or exclude the presence of significant CAD. The study group was obtained by a retrospective search in databases of our clinic and included 77 men and 23 women with an age of 59 ± 12 years (range 19–85 years). There were 84 patients with significant CAD, defined as the presence of a ≥ 50% diameter stenosis in at least one of the three main coronary vessels (LAD, LCX, RCA) as assessed by visual analysis. Thirty four patients had single vessel disease and 50 patients had multiple vessel disease.

Data in the case library included patient information such as sex and age, segmental values of the relative thallium activity obtained by polar map analysis of the scintigraphic images, and results from coronary angiography specifying presence or absence of CAD. Scintigraphic images of the 100 patients were processed using the OSIRIS software and polar map data were extracted from six short axis images and saved as a 2 × 6, 4 × 6, 6 × 6, and 6 × 12 matrix of integer numbers. Therefore, a case in the case library consists of the following features:

name, sex, and age of the patient

12, 24, 36, or 72 integer values obtained by polar map analysis of the scintigraphic image

one binary value representing the result of the visual analysis of the coronary angiography (1/0...presence/absence of CAD).

2.3. Similarity metric

Case retrieval is the implementation of the process which identifies from the case library a set of cases that are similar to the current problem. Thus, a similarity metric was defined to measure in quantitative terms the similarity between a case and the new problem. This metric was then applied for case retrieval.

In order to define the similarity metric, two observers independently processed scintigraphic images of ten patients using the OSIRIS software. Respective polar map data were compared and three different intervals were used as tolerance ranges: ± 1 standard deviation (S.D.), ± 2 standard deviations, and ± 2.5 standard deviations of the difference in segmental tracer uptake averaged over all 72 numbers in the polar map. The standard deviation of normalized segmental activities (maximum value 250) between the two observers was 11, so that the tolerance range was defined as ± 11, ± 22, and ± 28. Numeric values for the difference between the current case and the case in the library that are within the tolerance range were considered a complete or partial match depending on the similarity metric.

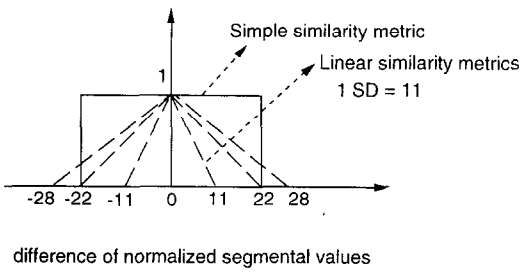

Fig. 2. Schematic illustration of two algorithms for the measurement of similarity: (1) simple similarity measurement (similar: 1, not similar: 0) (2) piece-wise linear similarity measurement (grade of similarity $\in [0, 1]$), using three different values as tolerance range: 1 standard deviation (S.D.), 2 S.D., and 2.5 S.D.

We used two different distributions to determine the degree of similarity, as shown in Fig. 2: a step function with two discrete states (0 and 1), and a linear function. Moreover, the similarity metric was designed to also include weights for each of the 6 planes to model possible differences in the diagnostic importance of each plane. Thus, for the case of 72 segmental values the similarity score is calculated for the step function as follows:

for i: = 1 to 72
  if $|\text{CurrentCase.Segment}_i - \text{CaseInCaseLibrary.Segment}_i| \leq \text{ToleranceRange}$
    $\text{Similar}_i := 1$
else
    $\text{Similar}_i := 0$
    $\text{SimilarityScore} := \text{SimilarityScore} + \dfrac{\text{Weight}_i * \text{Similar}_i}{\Sigma_{i=1}^{72} \text{Weight}_i}$
end for

and for the linear function as follows:

for i: = 1 to 72
  if $|\text{CurrentCase.Segment}_i - \text{CaseInCaseLibrary.Segment}_i| \leq \text{ToleranceRange}$
    $\text{Similar}_i :=$

$$1 - \frac{|\text{CurrentCase.Segment}_i - \text{CaseInCaseLibrary.Segment}_i|}{\text{ToleranceRange}}$$

else
    $\text{Similar}_i := 0$
    $\text{SimilarityScore} := \text{SimilarityScore} + \dfrac{\text{Weight}_i * \text{Similar}_i}{\Sigma_{i=1}^{72} \text{Weight}_i}$
end for

The weights in these formulas are selected according to heuristic values specified by the expert as follows: 1 for planes 1 and 6, 2 for planes 2 and 5, and 3 for planes 3 and 4. The five best matches to the current case were retrieved and their similarity scores were displayed in descending order.

2.4. Case indexing

The case library at present comprises 100 cases. Thus, all cases can be sequentially interrogated for degree of similarity without the use of indices. However, for large case libraries it is conceivable that sequential searching could impose demands on computational resources that may limit system performance. Under these circumstances, indexed searching may offer an alternative approach for similarity assessment. To determine the effect of indexing, we developed different versions of the retrieval process with and without the use of an index to compare the efficiency and accuracy of both strategies. The following formula was used for building an index of the case library:

$$\text{Index:} = \frac{\sum_{i=1}^{72} \text{Weight}_i * \text{SegmentalActivity}_i}{\sum_{i=1}^{72} \text{Weight}_i}$$

2.5. Adaptation process

After a set of similar cases has been retrieved, these cases can be adapted to resemble more closely the case under scrutiny. Adaptation is typically performed based on a set of heuristic rules. These rules can be applied on differences observed between input parameters of the retrieved case and the current case to advise on adaptation of the proposed outcomes according to the current case.

Two different adaptation strategies were implemented and tested in the current study. The first adaptation strategy was developed based on a mapping between the main coronary vessels and the polar map according to a schematic model of the

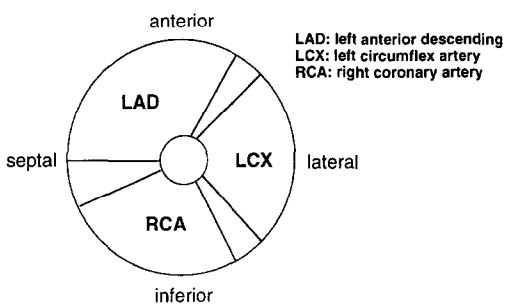

Fig. 3. Assignment of vascular territories corresponding to the three main coronary arteries onto the scintigraphic polar map as used for model-based adaptation.

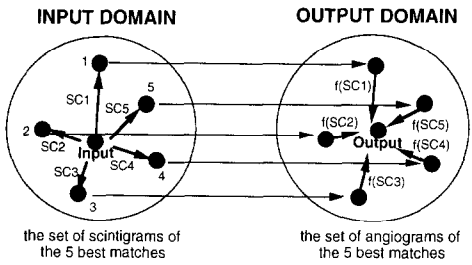

Fig. 4. Adaptation strategy based on the angiography results of the five most similar cases. To account for different grades of similarity, the similarity scores (SC) derived from the input domain are incorporated into weighting factors f(SC) that are applied to the angiographic results in the output domain.

coronary circulation as illustrated in Fig. 3. If a segmental activity value on the polar map of the target case was significantly lower than the corresponding value of the most similar case, the score of the respective vessel was increased to indicate more severe CAD in this vascular territory and vice versa. Thus, the model of the coronary anatomy together with scintigraphic data from the current case and the case library were used to adapt the output of the CBR application.

For the second adaptation strategy, the binary number reflecting the presence or absence of CAD of the five retrieved cases with the highest similarity scores were used to derive an assessment concerning the presence of CAD in the current patient. This adaptation algorithm, which is schematically shown in Fig. 4, uses the differences in the input domain between retrieved cases and the current case to calculate a data point in the output domain that corresponds to the angiographic evidence of CAD. The algorithm calculates a number between 0 and 1 that reflects the likelihood of CAD based on the similarity scores of the retrieved cases according to the formula:

$$\text{Result(CAD)} = \frac{\sum_{i=1}^{5} \text{SimilarityScore}_i * \text{CAD}_i}{\sum_{i=1}^{5} \text{SimilarityScore}_i}$$

Also, the standard error for each result is calculated to inform the user about the homogeneity or heterogeneity among the five retrieved cases:

$$\text{StandardError(CAD)} = \sqrt{\frac{\sum_{i=1}^{5} ((\text{CAD}_i - \text{Result(CAD)})^2 * \text{SimilarityScore}_i}{4 * \sum_{i=1}^{5} \text{SimilarityScore}_i}}$$

At the end of retrieval and adaptation, the results are displayed on the computer screen as shown in Fig. 5, where the presence of CAD was specified as a binary decision for the best match approach and a value between 0 and 1 for the adapted

result. The later value represents the likelihood of disease based on the CAD status of the five most similar images and their similarity scores. The output of SCINA is presented to the user as a recommendation with the best match highlighted to indicate the diagnosis suggested by SCINA.

2.6. Development tools

Programming shells can be regarded as indicators of the degree of generalization that can be achieved in a particular area of knowledge-based systems. Several commercial shells are available that support the development of case-based reasoning systems [16]. ESTEEM [9] is a professional tool that allows developers to design case-based reasoning systems on a Microsoft Windows-based computer. It uses the PC-DOS file system for storing entire applications with the options for export and import of case libraries in dBase, Lotus 1-2-3, ORACLE databases, or ASCII files. ESTEEM allows the creation of 'forms-like' interfaces for users to enter the features of a new case. For developers, ESTEEM allows to specify options for users and permitted modifications to the case library. Based on the choices made by the developer, a user can change the way cases are retrieved, add new cases to the case library, or use rules to assist in adaptation of the retrieved case.

While ESTEEM offers the capability for rapid prototyping, it has only limited flexibility concerning changes of structural components. For example, the definition of the similarity function or its parameters cannot be altered. Thus, the final prototype of the CBR system SCINA had to be implemented and examined using MS-EXCEL 5.0 with VISUAL BASIC as programming language. The case library

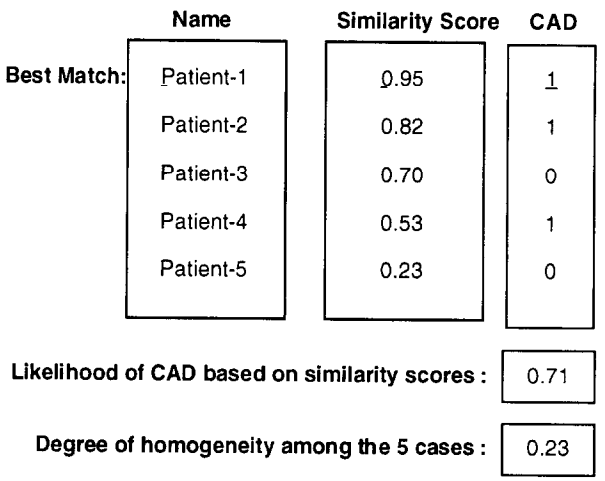

Fig. 5. Sample output of SCINA displaying the suggested diagnosis concerning the presence or absence of coronary artery disease in the current patient. The five cases with the highest similarity scores are displayed as a sorted list with the best match as first entry. Results derived from score-based adaptation are reported at the lower part of the output display.

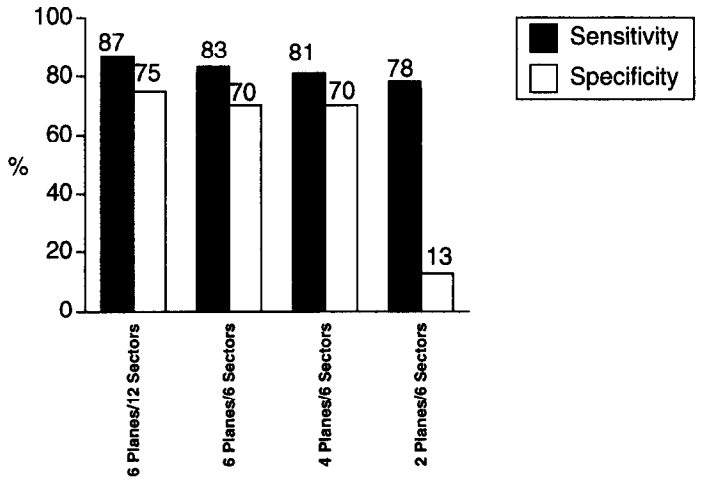

Fig. 6. Sensitivity and specificity for detection of coronary disease using four different specifications of input data granularity.

was organized as a database, and VISUAL BASIC was used to implement the retrieval and adaptation processes.

2.7. System evaluation

The effects of input data granularity, similarity metric, case indexing and adaptation strategies were evaluated based on the diagnostic accuracy of the CBR system with respect to the identification of significant CAD.

To evaluate the diagnostic accuracy of the CBR application SCINA, a 'round robin' evaluation method was applied as follows: each of the 100 cases was temporarily removed from the case library and was then presented to SCINA as a case for evaluation. The output of the CBR system was compared to the true status of the patient based on angiographic results of the patient. In all tests, we measured the sensitivity (true positive rate), specificity (true negative rate), and diagnostic accuracy (number of true positive and true negative cases / number of all cases) of the system for detection of CAD.

3. Results

The effect of input data granularity on sensitivity and specificity for the detection of coronary heart disease was studied using ESTEEM as rapid prototyping tool. As shown in Fig. 6, the use of six planes and 12 30° segments per plane provided the best diagnostic accuracy. There was a slight decrease in diagnostic accuracy when six 60° sectors per plane and six or four planes were used, while a marked reduction in accuracy occurred with only two planes. Therefore, a 6×12 matrix was used for the final prototype system implemented in MS-EXCEL.

The impact of different algorithms for the assessment of similarity is shown in Fig. 7. The highest values of sensitivity and specificity were observed for a tolerance range of ± 2 S.D. with a linear function used for similarity assessment. Without the use of weighted planes, sensitivity and specificity for detection of CAD of SCINA were 95% and 70%. When weights were used in the similarity metric, sensitivity increased to 98% while specificity remained unchanged. The overall diagnostic accuracy of this optimized configuration of SCINA was 93%.

The influence of indexing was assessed by relating diagnostic accuracy to computational efficiency. Fig. 8 depicts sensitivity and specificity for diagnosis of CAD in relation to computation time for both sequential and index-based search strategies. The computation time without indexing was considered 100%. While indexing markedly reduced the computation time, a drop in accuracy was observed at the same time indicating that a restriction of the search space during case retrieval may severely affect the diagnostic accuracy of a case-based reasoning system.

The results obtained by adaptation algorithms following case retrieval are displayed in Table 1. The adaptation strategy based on the mapping of coronary vessels onto the polar map is called the model-based adaptation strategy, while score-based adaptation refers to the adjustment of output value by similarity scores. There was no difference of sensitivity and specificity values obtained with and without the use of the model-based adaptation approach. The score-based adaptation algorithm yields a real number between 0 and 1 which reflects the likelihood of coronary disease and needs to be converted to a binary value to yield a diagnostic statement. Thus, two threshold values were evaluated with scores ≥ 0.6 and ≥ 0.8 assumed to indicate CAD. As with model-based adaptation, score-base adaptation did not improve the diagnostic accuracy of the system. However, selection of a cutoff value permits to construct a diagnostic system with either a predefined specificity or sensitivity according to diagnostic needs.

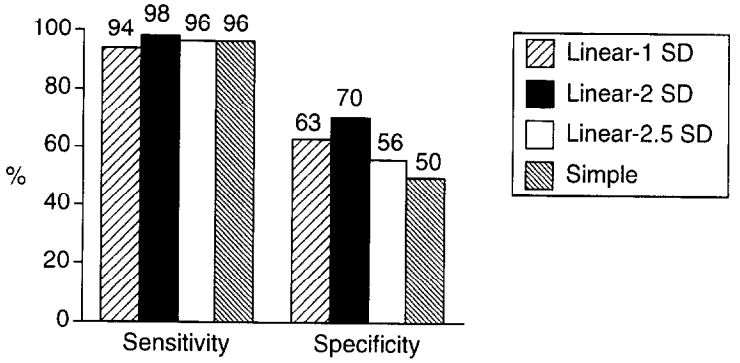

Fig. 7. Sensitivity and specificity for the diagnosis of coronary artery disease obtained from different implementations of a similarity metric using either a piece-wise linear function with three tolerance ranges (1 standard deviation (S.D.), 2 S.D., and 2.5 S.D.), or the step function (simple).

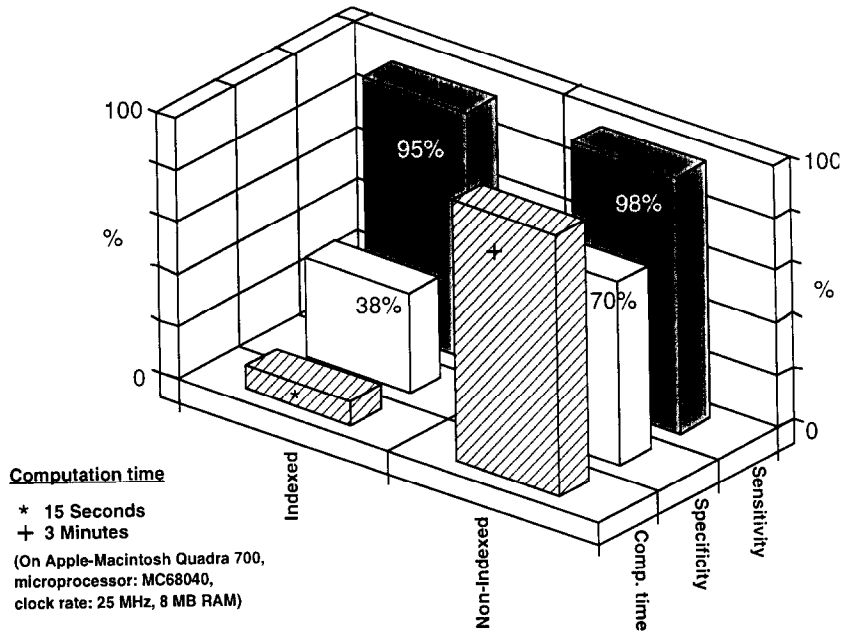

Fig. 8. Comparison of sensitivity and specificity for detection of coronary artery disease, and the computation time of retrieval process with and without the use of indexing.

4. Discussion

This study investigated the feasibility of using a case-based reasoning system for the automated image interpretation of myocardial SPECT perfusion scintigrams. The diagnostic accuracy of a prototype system for detection of significant coronary artery disease was investigated with a focus on the effect of several system design issues. Different configurations of SCINA concerning the input data granularity, definition of a similarity metric, retrieval procedures, and adaptation strategies were developed and evaluated to identify an optimized system configuration.

Table 1
Sensitivity, specificity, and diagnostic accuracy of SCINA for detection of coronary heart disease using different adaptation strategies and development tools. Two threshold values (0.6 and 0.8) were used in score-based adaptation

Adaptation	Tool	Sensitivity	Specificity	Accuracy
No adaptation	ESTEEM	87%	75%	85%
No adaptation	EXCEL	98%	70%	93%
Model-based	ESTEEM	87%	75%	85%
Score-based (≥ 0.6)	EXCEL	77%	81%	78%
Score-based (≥ 0.8)	EXCEL	56%	94%	60%

4.1. Input data granularity

The case context includes data from myocardial perfusion scintigraphy as input data and coronary angiography as output data. For a CBR system, the representation of input data needs to be carefully selected to ascertain a complete definition of every case with minimal redundancy. If input data granularity is too coarse, the similarity assessment may prove to be inadequate. If input data granularity is too dense, case retrieval may be inefficient.

Depending on image reorientation and plane selection, polar map analysis of perfusion scintigrams can generate data matrices of different sizes. Thus, we performed experiments to identify an adequate matrix size for SCINA. Initially, we used a set of 72 integer numbers (6×12) reflecting the relative tracer activity of 12 contiguous 30° segments on six equidistant short axis planes. This matrix provides for an almost homogenous information density in 3D space. As the distance between image planes is approximately 1 cm, the information density along the axial direction is about one sampled value per 1 cm. In comparison, the inplane information density with 12 segments is also about one averaged value per 1 cm, as 12 segmental values are extracted from the relative tracer activities on circumferential profiles with a diameter of approximately 4 cm.

From the initial matrix configuration, we attempted to achieve data reduction by decreasing data granularity. When the number of segments per plane was reduced by 50% to six, the diagnostic accuracy dropped only slightly. Additional experiments were performed to elucidate the relation between the number of planes and the diagnostic accuracy. Consultation with human expert readers revealed that the most apical and most basal planes typically contribute the least for the diagnostic reading. Accordingly, we removed one and two apical and basal planes from the polar map matrix and retained only four and two middle planes. While diagnostic accuracy decreased only slightly when the two outermost planes were removed, a marked change occurred when only the middle two planes were retained. Thus, for the final prototype of SCINA, we selected a 6×12 input data matrix and retained plane weighting to reflect the difference in diagnostic content between planes.

4.2. Similarity metric

The definition of a similarity metric is a critical design issue in the implementation of a CBR system. In the present study, the algorithm of the similarity metric was based on results obtained from a reproducibility analysis for polar map data. In experiments where two human observers processed the polar map data, we measured the interobserver standard deviation of segmental activity values and defined tolerance ranges with ± 1 S.D., ± 2 S.D., and ± 2.5 S.D. to indicate similarity. A step function and a symmetric piece-wise linear function were evaluated as basic algorithms of a similarity metric that permits quantification of the similarity of cases on a scale of 0 to 1. Diagnostic accuracy was best for the piece-wise linear function defined on a tolerance range of ± 2 S.D. corresponding to a 95% confidence interval.

4.3. Case indexing

Indices may be helpful in CBR applications to improve case retrieval. Indices should be predictive and should address the specific purposes for which the case will be applicable. In addition, indices should be more abstract than the details of a particular case to make a case useful in a variety of situations. However, they should be also sufficiently concrete to be easily recognizable and specific for certain situations [2,10]. As index-based retrieval limits the similarity assessment to a subgroup of cases, the use of indexing may also affect the performance of a CBR system.

In the present study, the diagnostic accuracy was lower with the use of an index than without the use of an index. Thus, the index-based search improved computational efficiency, but simultaneously decreased the diagnostic accuracy. As the number of cases in the case library of SCINA allows for the sequential interrogation of all cases within a short time period, sequential searching was employed in the prototype version. The loss of accuracy due to indexing may be explained by the particular algorithm chosen in the present study to generate index values. According to its definition, the indexing scheme does not represent adequately the position or morphology of a defect. For example, scintigrams with a large but slight defect or a small but severe defect will have similar index values.

4.4. Adaptation

Two different adaptation algorithms were evaluated that modify the diagnostic assessment of the retrieved cases to fit the new problem. First, a mapping between vessels and the polar map regions was used to identify the individual vessel that supplies a specific area on the polar map. When an area on the polar map was associated with lower or higher segmental values than the retrieved case, the retrieved score for vessel segments associated with that area was increased or decreased accordingly. However, while conceptually intriguing, in practice this adaptation strategy did not improve the diagnostic performance of the reasoning system.

In a second adaptation strategy, we tried to implement adaptation rules based on the angiography data of the five cases that matched the input case best. This strategy was based on the assumption that information content may be improved by increasing the number of cases that contribute to the final diagnosis. However, the diagnostic accuracy of this reasoning system was also less than the best match approach. Thus, the angiography data of the retrieved case with the highest similarity score was finally selected as the proposed diagnosis for the prototype implementation of SCINA.

4.5. Development tool

While the CBR system SCINA was initially designed using the CBR development tool ESTEEM [9], different pilot versions of the CBR system had to be imple-

Table 2
Sensitivity, specificity, and diagnostic accuracy of the CBR system SCINA in comparison with two previously published studies from the Cedars-Sinai Medical Center (CSMC) [12] and a Multicenter trial [18]

	Sensitivity	Specificity	Accuracy
SCINA	98%	70%	93%
CSMC	95%	56%	89%
Multicenter	94%	44%	84%

mented in a programming environment due to the limited set of design specifications that could be modeled by ESTEEM. Selection of ESTEEM among other CBR shells was based on the reported advantages [16], its availability on common platforms (PC, Windows 3.1), and competitive pricing. However, systems developed with ESTEEM are restricted by several limitations including a limited number of characters for the definition of similarity and adaptation rules, the inability to define a proprietary similarity metric, a limited number of cases ($n = 1$) that can be used for adaptation, and the limited ability to present the retrieved results with an adequate user interface. Therefore, we chose the object-oriented programming language VISUAL BASIC in combination with MS-EXCEL data management to implement the final prototype version of SCINA. However, for rapid prototyping ESTEEM was used when possible to test and evaluate aspects of system design and system specifications.

4.6. Related studies

Quantitative image analysis methods of SPECT perfusion scintigraphy was successfully applied in previous studies for detection and localization of CAD [12,18]. These studies based their assessment of CAD on a normal database by comparing the patient data to gender-matched normal files that were assembled from individuals with a low likelihood of coronary artery disease. In contrast, SCINA uses both normal and pathological cases to obtain an assessment of myocardial perfusion scintigrams. In comparison with these other methods, SCINA shows an improved specificity and a slightly improved diagnostic performance as indicated in Table 2.

The diagnostic accuracy of a CBR system for the interpretation of myocardial perfusion scintigraphy critically depends on the distribution of the study population stored in the case library of the system. As only 16 normal cases were included in the case library of SCINA, the complete spectrum of normal patterns may not be represented adequately. Therefore, the specificity of SCINA for the detection of significant CAD possibly could increase provided additional cases with non-significant CAD are included into the case library.

Recently, methods of artificial intelligence (AI) such as neural networks, rule-based expert systems, and model-based diagnostic systems have been used to derive automated systems for the interpretation of myocardial scintigrams.

Porenta and coworkers [14] developed artificial neural networks to automatically detect the presence of CAD and assess its severity and location based on the segmental analysis of planar perfusion scintigrams. Fujita and coworkers [5] reported on the use of neural networks for a computer-aided diagnosis of CAD from myocardial SPECT polar maps. The diagnostic accuracy of both neural networks (71% [14] and 90% [5]) was lower than the diagnostic accuracy of SCINA (93%).

PERFEX is a rule-based expert system which has been developed at the Georgia Institute of Technology for the interpretation of cardiac SPECT images [4,6]. While a perfect agreement for detection of CAD between the diagnosis of PERFEX and visual analysis by human expert was reported in these studies, the authors also indicated that problems occured in the areas of knowledge acquisition and robustness. The rule-based expert system developed by Cios and coworkers [3] is based on the analysis of planar scintigrams and is limited to patients with single-vessel disease. The predictive accuracy for detection of abnormal scintigrams was 87%.

A model-based classification of cardiac SPECT images [13] has been described that interprets polar maps by integrating an abductive model of coronary myocardial perfusion and coronary arterial branching with heuristics embodying additional domain knowledge (e.g. variations of vessel anatomy) and information related to the image acquisition process. This model-based approach provides a diagnostic accuracy that is similar to SCINA (91%).

CBR systems have also been used successfully for image interpretation in radiology. Macura and coworkers [11] designed a case-based training system to assist radiologists in the diagnosis of brain tumors from X-ray computed tomography and magnetic resonance imaging. A case library of 122 cases with 640 digitized brain images was used to provide radiologists with a tool that supports their visual memory using case-based reasoning. In order to evaluate the accessibility of the cases and the system accuracy, 109 studies from the case library were presented to the residents as a diagnostic problem with the additional task to retrieve comparable cases from the case library. Eighty one percent of these cases were indexed accurately and were accessible to the residents. The authors concluded that a case library can offer reference images for the purpose of a comparison to the case under scrutiny, and thus may be used successfully for teaching applications.

ProtoISIS [8] is a prototype CBR decision support system for the selection of appropriate diagnostic imaging procedures based on the clinical request. ProtoISIS was trained using 200 consecutive cases of actual requests for ultrasound and X-ray computed tomography. The system performance was evaluated by presenting ProtoISIS 100 cases of actual requests for imaging procedure. The system correctly classified 84% of those requests suggesting that CBR can be applied successfully for selecting diagnostic imaging procedures in clinical praxis.

In summary, for image interpretation case-based reasoning appears to offer advantages over other approaches as it attempts to mirror human experts who use cases for explanation, extrapolation, strategic planning, learning, and teaching. Using cases also allows to apply past solutions, to deal with domains that lack a causal model, and to explain the solution by examples. Existing databases can be

used as case library for a case-based reasoner and thus may serve as an automated problem solver. Data acquisition and modification is easily accomplished, because knowledge is acquired through appropriate examples from experiences of the domain expert, and maintenance of a CBR system is similar to editing a database for inaccurate records. In particular, possible advantages of the CBR system SCINA over other AI-based applications include the flexibility in data acquisition and data modification, the capability to explain the reasoning process, improved diagnostic accuracy, and implementation on standard low-cost PC-software with wide availability.

A major limitation of case-based reasoning is that optimal solutions cannot be guaranteed, as a case-based system is by definition restricted to its case library and has an only restricted degree of flexibility. In contrast, a rule-based system produces optimized answers for all problems contained within the knowledge base. Other problems of CBR systems are limitations which are associated with the selection of an indexing scheme and the design of adaptation rules.

4.7. Conclusion

Case-based reasoning as a method of artificial intelligence was successfully used to develop an automated image interpretation system that generates from myocardial perfusion scintigrams an assessment concerning the presence of coronary artery disease with good diagnostic performance. However, further studies are needed to define in more detail the clinical impact of SCINA not only with respect to the detection of CAD but also regarding the assessment of severity and location of CAD.

Acknowledgements

The technical assistance of Brigitte Slama and Regina Wagner is gratefully acknowledged.

References

[1] ACC, AHA, and SNM, Standardization of cardiac tomographic imaging, *J Am Coll Cardiol* 20 (1992) 255–256.

[2] R. Barletta, A hybrid indexing and retrieval strategy for advisory CBR systems built with ReMind®, *Proceeding of Second European Workshop on Case-Based Reasoning* (Chantilly, 1994) 49–58.

[3] K.J. Cios, R.E. Freasier, L.S. Goodenday, and L.T. Andrew, An expert system for diagnosis of coronary artery stenosis based on Tl201 scintigrams using the Dempster Shafer theory of evidence, *CABIOS* 6 (1990) 333–342.

[4] N.F. Ezquerra and E.V. Garcia, Artificial intelligence in nuclear medicine imaging, *Am J Card Imaging* 3 (1989) 130–141.

[5] H. Fujita, T. Katafuchi, T. Uehara and T. Nishimura, Application of artificial neural network to computer-aided diagnosis of coronary artery disease in myocardial SPECT bull's-eye images, *J Nucl Med* 33 (1992) 272–276.

[6] E.V. Garcia, M.D. Herbst, C.D. Cooke, N.F. Ezquerra, B.L. Evans, R.D. Folks and E.G. DePuey, Knowledge-based visualization of myocardial perfusion tomographic images, *Proc Vis Biomed Comput* 1990 (1990) 157–161.

[7] M. Haddad, B. Slama, R. Wagner and G. Porenta, The impact of reorientation algorithms on quantitative myocardial Tl-201 SPECT, *J Nucl Med* 5 (1995) 231.

[8] C. Kahn Jr. and G.M. Anderson, Case-based reasoning and imaging procedure selection, *Invest Radiol* 29 (1994) 643–647.

[9] J. King, *ESTEEM Enabling Solutions Through Experience Modeling* (Esteem Software Inc., USA, 1994).

[10] J. Kolodner, *Case-Based Reasoning* (Morgan Kaufmann Publishers Inc., San Mateo, CA, 1993).

[11] R.T. Macura, K.J. Macura, V.E. Toro, E.F. Binet, J.H. Trueblood and K. Ji, Computerized case-based instructional system for computed tomography and magnetic resonance imaging of brain tumors, *Invest Radiol* 29 (1994) 497–506.

[12] J. Maddahi, K.V. Train, F. Prigent, E.V. Garcia, J. Friedman, E. Ostrzega and D. Berman, Quantitative single photon emission computed thallium-201 tomography for detection and localization of coronary artery disease: optimization and prospective validation of a new technique, *J Am Coll Cardiol* 14 (1989) 1689–1699.

[13] P. Petta, W. Horn and G. Porenta, Model-supported interpretation of cardiac Tl-201 SPECT polar maps, *J Nucl Med* 5 (1995) 12.

[14] G. Porenta, G. Dorffner, S. Kundrat, P. Petta, J. Duit-Schedlmayer and H. Sochor, Automated interpretation of planar thallium-201-dipyridamole stress redistribution scintigrams using artificial neural networks, *J Nucl Med* (1994) 2041–2049.

[15] R.C. Schank and C.K. Riesbeck, *Inside Case-Based Reasoning* (Lawrence Erlbaum Associates, 1989).

[16] T.J. Schult, *Case-Based Expertsystem Shells, Methods and Tools* (Institute of Psychology, University of Freiburg, 1992, in German).

[17] Univ.Geneva, OSIRIS (1.0) *User's Manual* (Digital Imaging Unit Center of Medical Informatics, University Hospital of Geneva, Geneva, 1992).

[18] K.F. Van Train, J. Maddahi, D.S. Berman, H. Kiat, J. Areeda, F. Prigent and J. Friedman, Quantitative analysis of tomographic stress thallium-201 myocardial scintigrams: a multicenter trial, *J Nucl Med* (1990) 1168–1179.