The Thallium Diagnostic Workstation: Learning to Diagnose Heart Imagery from Examples

Rin Saunders

The thallium diagnostic workstation (TDW) is an integrated workstation for the learning and application of diagnostic rules for thallium heart imagery. TDW learns diagnostic rules from training sets of images, using a symbolic induction algorithm developed specifically for this application. TDW uses machine vision to identify image features of diagnostic significance (findings), which are described to the physician in a write-up. The physician can perform automated diagnosis by applying a rule set to the findings, selected from TDW's catalog of learned rules. The physician can also view imagery, record his(her) own findings, and enter a diagnosis. TDW informs the physician of its concurrence or non-concurrence with the physician's diagnosis.

TDW is deployed at the United States Air Force School of Aerospace Medicine (USAFSAM). Physicians at USAFSAM qualify fliers for aeromedical fitness. Significant coronary artery disease, causing narrowing of the arteries supplying blood to the heart muscle, is grounds for disqualification. Because Air Force standards are strict, fliers can be

disqualified even if there are no overt symptoms of coronary artery disease (such as angina pectoris). The Air Force can disqualify a flier for the loss of 30 percent of the diameter of a coronary artery. However, coronary artery disease often produces no severe symptoms until close to 90 percent of the diameter is lost, making the diagnosis of aeromedically significant coronary artery disease a harder proposition than diagnosing coronary artery disease in a conventional clinical setting.

USAFSAM uses thallium imagery to screen fliers suspected of having coronary artery disease. A thallium scan is graded normal, borderline, or abnormal. Patients with borderline or abnormal scans are subject to angiography, the definitive technique for diagnosing the disease. In *angiography*, puffs of radiopaque dye are released from a catheter that is passed through the femoral artery to just above the coronary circulation. A moving X-ray image is taken that shows the course of blood through the coronary arteries. Angiography is invasive, requiring cardiac catheterization and the use of a dye sometimes associated with allergic reactions. Using thallium imagery lowers patient risk by reducing unnecessary angiography.

Diagnosing asymptomatic coronary artery disease from thallium imagery is difficult and subjective. Physicians at USAFSAM show considerable variation in their thallium-reading skills (evaluated as the percentage of thallium-based diagnoses confirmed by angiography). The physicians' skill level correlates well with time on the job at USAFSAM (Kay 1989). Physicians become skilled in the technique through on-the-job experience, then leave after a three- to four-year rotation.

Benefits

TDW provides the following benefits to USAFSAM physicians: First, TDW leverages expertise. It enables physicians to perform as well as the best available diagnostician. TDW can also be trained on angiographic diagnoses; in which case, it performs comparably with USAFSAM's best diagnosticians. Second, TDW conserves expertise. USAFSAM physicians become experts in thallium interpretation and then leave. TDW preserves the expertise of physicians for use by their successors. Third, TDW objectifies expertise. TDW's learned rule sets enable experts to compare and evaluate objective criteria for classifying thallium images. Fourth, TDW provides consistent and reproducible diagnoses. Although thallium interpretation is somewhat subjective, precise numbers underlie the image. TDW provides a diagnostic standard that is consistent and reproducible across both doctors and patients. This approach benefits longitudinal studies that involve multiple doctors and patients.

These benefits are provided by traditional expert systems. TDW's machine vision capability provides an additional benefit: It can train incoming physicians. TDW is not primarily intended as a tutoring system. In general, expert systems do not perform well as tutors. However, TDW is useful in honing incoming physicians' *expert vision*, the ability to detect features of diagnostic significance. Because TDW's findings can be displayed on a screen with imagery, the novice thallium reader can use the findings as a guide.

TDW's machine-learning capability provides the following benefits: First, it reconstructs expertise. A longitudinal study of coronary artery disease might include patients diagnosed by physicians who left the study years ago. TDW can reconstruct the physicians' diagnostic criteria from examples. The reconstructed criteria can be used to evaluate puzzling diagnostic calls. The call can be consistent with TDW's learned criteria for this physician, justifying the call. If the call is inconsistent with the learned criteria, the call can be treated as an anomaly. Second, TDW supports studies. USAFSAM physicians are concerned with diagnosing low-grade coronary artery disease. What is the least degree of this disease that can reliably be distinguished from normality? Training TDW on specially selected image sets can help answer this question. Third, TDW is a self-maintaining knowledge base in the sense that users can update the diagnostic rules by training TDW on new cases.

Innovative Application of Machine Vision & Machine Learning

Medical imagery is potentially a high-payoff domain for expert systems. For chest films, Garland (1950) showed that radiologists routinely miss about 30 percent of all abnormalities. However, building an expert system to classify images poses challenges beyond those normally associated with expert systems.

An expert system for medical imagery must incorporate machine vision. Human diagnostic reasoning proceeds from high-level image features rather than pixel intensities. A nuclear medicine specialist learns to see features of diagnostic significance in an image—an expert vision based on pattern recognition rather than deductive reasoning.

TDW emulates the physician's expert vision. Previous automated thallium diagnostic systems used statistical profiles to classify images (Burow et al. 1979; Garcia et al. 1981; Watson et al. 1981). TDW's AI approach is an improvement because it reasons from features that physicians can see in the imagery and presents its findings in terms used by physicians. This approach helps build physicians' confidence in the correctness of TDW's conclusions.

Knowledge acquisition for image classification presents another challenge. Knowledge acquisition is difficult enough when the expert reasons from facts and assumptions that are objectively true or false. However, physicians classify thallium imagery using subjectively defined image features that are construed differently by different physicians or even by the same physician over time. Further, in becoming an expert, the physician fuses feature extraction and diagnostic reasoning into a single process called judgment that the physician cannot explain. Knowledge elicitation produces textbook explanations that bear little relationship to how the physician actually interprets images.

TDW uses machine-learning techniques to automate knowledge acquisition. It deals with the problem of inconsistency by seeking rules that replicate the physician's judgments in as many cases as possible. Because TDW learns from examples rather than explanations, it is not confused by a physician's misunderstanding of his(her) own decision processes.

System Architecture and Function

TDW is implemented on a Science Applications International Corporation SIGMA-1 microcomputer incorporating a 25-megahertz 80386 processor, an 80387 numeric coprocessor, a floating-point coprocessor rated at 22 MFlops, and a VGA monitor. Microsoft WINDOWS provides a windowed user interface. Much work went into the interface design. As a result, TDW has only two menu bars (one for viewing images and one for all other functions) and no submenus. Although TDW has separate programs for learning and diagnosis, this separation is transparent to the user because WINDOWS provides a seamless interface. The Nexpert Object expert system shell maintains information about cases, image features, diagnoses, and diagnostic rules. Nexpert Object was selected because it offers a class-object–oriented frame system and flexible agenda control during inferencing, and it can be embedded under WINDOWS.

TDW performs two main functions: learning diagnostic rules from sets of diagnosed images and applying rules to produce diagnoses. TDW has a subsystem for each function. The subsystems are constelled around a central knowledge base. The learning subsystem inserts rule sets into the knowledge base and the application subsystem runs them.

Processing begins when TDW receives an image file from USAFSAM's medical imaging system over an ethernet. The image file contains the patient's name, the physician's name, and other identifying information as well as raw pixel intensities. TDW enters the patient information into the knowledge base and performs image processing to improve the quality of the raw data. TDW's feature extractor identifies features of

diagnostic significance in the image. These features are entered into the knowledge base and associated with the patient data.

The physician can view imagery using the TDW image viewer. The viewer can display six images belonging to a case (three anatomical views taken at two intervals after thallium injection) either simultaneously or separately for higher resolution. Grey-scale, full-color, and a quantized color-scale views are available. Both raw and processed images can be displayed. A mouse-driven pixel meter enables the physician to evaluate individual pixel intensities. The physician can view TDW's findings by opening a window that displays the feature descriptions in narrative English.

TDW presents the physician with a list of current cases. The patient's name and number, the physician's name, the image date, angiographic results (if available) and diagnoses from thallium reading by physicians, angiography data, and TDW's rules are displayed for each case. The physician can view imagery, record findings and diagnoses, view TDW's findings, and apply rule sets to perform automated diagnosis. TDW lists the available rule sets and displays English translations of the rules. When system capacity is reached, cases are automatically archived to disk.

TDW learns diagnostic rules from training sets of diagnosed images. It enables physicians to construct training sets by grouping cases that meet user-specified screening criteria. The criteria enable the physician to include diagnoses based on the source—specific physicians or angiography. Cases can also be selected based on image data range, concurrence or nonconcurrence of the angiographic result with the thallium scan, or the presence or absence of key words that physicians can associate with cases. A physician can edit the training set to enter or remove individual cases.

The learning subsystem uses a symbolic induction technique called METARULE, which was developed specifically for TDW. As learning proceeds, the current best rule is displayed in English translation on the screen. TDW requires approximately 10 minutes to learn rules covering a training set of 100 cases. When learning is complete, the rule set is translated into Nexpert syntax and compiled into the knowledge base. The English translation is also stored for review by the physician.

How the Thallium Diagnostic Workstation Works

I now take a closer look at how thallium imagery is produced and interpreted by physicians, how TDW extracts features of diagnostic significance from digitized imagery, and how learning is done.

Thallium Image Interpretation

In the thallium technique, patients are run on a treadmill until the electrocardiogram shows physiologic stress. At this point, the heart muscle is low in potassium ions. Depletion occurs most prominently in the left ventricle, which is the main pumping engine of the heart. The patient is injected with thallous-201 chloride as a bolus through an intravenous line. The thallium ion chemically resembles potassium, so it is absorbed by the heart muscle where the muscle is perfused with blood. If perfusion is normal, thallium is rapidly absorbed, then gradually washes out, typically attaining the halfway point in 84 minutes (Gerson 1987).

The patient is imaged under a gamma camera, providing an array of 128 x 128 pixels, within 6 minutes of injection. Three views are taken: the anterior, the 45-degree left anterior oblique, and the 70-degree left anterior oblique. The images are repeated after 4 hours of rest. The two sets of images are called the *stress images* and the *rest images*, respectively.

I studied how physicians grade images by asking the clinical staff members to describe their interpretation techniques. I also obtained talking protocols in which physicians described how they evaluate images while the evaluations are being performed.

The physicians stated that thallium imagery indirectly reveals the presence of coronary artery disease by depicting the perfusion of blood into the left ventricular myocardium. *Arterial narrowing* (stenosis) can delay the uptake and the washout of thallium. An image region that washes out too slowly is a *washout abnormality.* More severe stenosis can delay uptake so that the rest images show thallium is washing in rather than out. An image region showing washin is called a *reperfusion defect.* A reperfusion defect, combined with a washout abnormality, is called a *matched defect.* A still greater degree of stenosis can produce a region with little uptake on either set of images, which is called a *perfusion defect.* A perfusion defect can indicate scar tissue in the myocardium from a prior infarct. According to the clinical staff members, washout abnormalities have the weakest evidentiary strength for coronary artery disease, followed by reperfusion defects, matched defects, and perfusion defects.

The talking protocols revealed additional feature types. When large regions of an image show washin, but the intensity is never sufficient to warrant being called a reperfusion defect, physicians call the defect *low-grade pervasive reperfusion.* A *reversing horseshoe* is a pattern in which the valve plane is hotter than the apex on the stress image but colder on the rest image. This pattern, more formally known as *reversing apical hypoperfusion*, is viewed as highly diagnostic for coronary artery disease.

In general, physicians grade an image with sufficiently intense perfusion defects, horseshoes, or matched defects as abnormal. Reperfusion defects can result in a grade of abnormal or borderline. Washout abnormalities or low-grade pervasive reperfusion warrant a grade of borderline. An image without defects or with low-intensity defects is graded normal.

Feature Extraction

Thallium image features have their visual basis in regions that appear hot or cold relative to the rest of the image and their evolution over time. To automatically extract features, we must quantify the terms by which physicians describe features, for example, hotness, coldness, size, intensity, and location. We must decide how hot, big, intense, and so on, an image region must be to qualify as a feature. We must also decide how to quantify the time evolution of a feature. The beating of the heart makes it impossible to subtract the rest image from the stress image by superposition. The two sets of images do not coincide in shape.

Image Preprocessing. Thallium imagery based on raw pixel intensities is coarse and noisy. Automated and human interpretation alike are greatly facilitated by preprocessing to smooth the image and reduce noise.

TDW applies Watson's (1981) method of bilinear subtraction and Laplacian filtering to reduce noise from background radiation, tissue cross-talk, and absorption of thallium by organs other than the heart. Then, TDW identifies the boundaries of the myocardium. Because the edges are not sharp, I use thresholding rather than gradients or related methods. Pixels that make up the myocardial image are normalized to lie in the range from 0 to 55. The image is smoothed with a Gaussian convolution. The image centroid and cardiac apex are located, and the image is partitioned among the heart muscle walls.

Feature Extraction. The feature extractor uses template-based methods to identify image features. The image is divided into 12 radial segments about the centroid. Because the image was normalized, regions that are hot or cold relative to the image as a whole can be identified through thresholding. The appropriate values were found through experimentation. I also found that a region must occupy at least 10 percent of the area of the myocardium to be considered a candidate feature.

When physicians assess the magnitude of a defect, they consider its size and intensity and the change in size and intensity over time. A candidate feature is confirmed if the magnitude is sufficiently large. About two dozen measures of magnitude were explored, derived from the defect's maximum pixel intensity, the average intensity and size, and the absolute and percentage changes in these values. For reperfusion de-

fects, I found that two measures of magnitude predicted the physicians' assessments: (1) the average pixel intensity of the region on the stress image times the absolute change in area and (2) the difference in the region's average intensity between the stress and rest images times the difference in area divided by the area of the region on the stress image. Other measures proved useful for the other feature types.

Learning

Learning rules from examples is not part of expert system technology but machine learning. Machine learning is still a research area. Several commercial expert system shells claim to learn rules from examples, but none perform adequately on any but simple problems (Thompson and Thompson 1986). This situation is in part because learning is a difficult problem and in part because different types of learning are appropriate to different domains, making it hard to build a general-purpose program. In addition, efficient learning often requires some domain knowledge beyond the examples themselves. I now turn to the requirements imposed by the thallium image domain.

Requirements for Learning Diagnostic Rules for Thallium Imagery. A machine-learning system for thallium imagery must deal with both categorical and numeric data. The type and anatomical location of a defect is categorical; the intensity is numeric. The learning system must be able to generalize about ranges of values. For example, a defect might warrant a grading of abnormal if it is sufficiently large in area or magnitude.

The learning system must tolerate counterexamples in the data. Many learning algorithms will learn a rule only if the rule is never contradicted. Medicine is not an exact science. Expert judgment will not be 100 percent consistent. A few anomalous cases should not always invalidate a rule.

The learning system should assess the quality of its rules. Because rules might not be valid for all the cases from which the system learns, the user must know how often a rule can be expected to be correct.

All other things being equal, preference should be given to rules that are simple and readable and make sense to human experts. The learning system should be biased toward simple rules.

How TDW Learns. A machine-learning algorithm named METARULE was developed to meet the requirements of the thallium domain. METARULE learns diagnostic rule sets in three phases. First, it learns rules for diagnosing normal cases. Then, it learns rules for diagnosing abnormal cases. The five best rules are retained for normal and abnormal cases. Finally, METARULE selects combinations of these rules to produce a complete rule set.

When learning about normal or abnormal cases, METARULE begins by making simple assertions about what makes a case positive (that is, makes an abnormal case abnormal or a normal case normal). These assertions make up the inductive kernel. Rules in the kernel reference aspects of the image, such as which feature types are present, how many features there are, what numeric values exist for feature attributes, and so on. Rules involving ranges of numeric attributes are formulated by seeking cut points with the greatest discriminatory power.

If the case is positive, METARULE postulates that the presence of its feature types and attributes makes it positive. For example, a positive case with a reperfusion defect having a magnitude of 100 might lead METARULE to postulate that (1) a reperfusion defect makes a case positive, (2) a reperfusion defect with a magnitude of at least 100 makes a case positive, or (3) having at least one feature makes a case positive. For negative cases, METARULE postulates that the feature, attribute value, and number of features prevent the case from being positive.

METARULE builds more complex rules by combining existing rules. Promising combinations are identified using a beam search. The search space is the space of all first-order formulas that can be built from the kernel using the Boolean predicates And, Or, and Not.

A rule can be regarded as a characterizer or a discriminator (Michalski 1983). A *characterizer* lists things that positive cases have in common; a *discriminator* tells what separates positive cases from negative cases. Discriminators are the object of the search. If a rule is not a good discriminator, it might be useful as a characterizer. A characterizer covers the positive cases well but includes too many negative ones. In a sense, a characteristic description is half a solution (few false negatives) in search of its other half (a description that reduces the number of false positives).

For each promising characterizer, METARULE seeks discriminators that preserve the coverage of positive cases but reject any false positives. Promising characterizers are combined with discriminators to form conjunctive rules. Disjunctive rules are formed by combining complementary characterizers to extend their coverage. Rules do not need to be conjunctive or disjunctive exclusively and can grow to an arbitrary level of complexity.

METARULE maintains lists of the best performing characterizers and discriminators. The search begins with the strongest characteristic description. For each characterizer in order of strength, METARULE generates a list of promising disjunctive rules, ranked from strongest to weakest. METARULE forms new rules by joining the most promising discriminators with the current characterizer. For each promising characterizer, METARULE finds other descriptions that extend the coverage but include few or no false positives. These descriptions are separated

to form new rules. As new rules emerge, the lists of characterizers and discriminators are continually updated.

Several elaborations on this basic strategy are intended to speed the search or provide a bias toward simpler rules. If METARULE discovers that two candidate rules are logically equivalent (that is, can be reduced to the same expression by symbolic manipulation), the simpler of the two rules is retained. This approach speeds the search and promotes the use of simple rules. Similarly, if two rules are empirically equivalent (that is, classify each case the same way for the same reasons), only the simpler rule is retained. Finally, a rule's simplicity can partially make up for a small lack of accuracy when selecting a rule to elaborate. The simplest rule is selected for elaboration from rules that rate within three percentage points of the best-scoring rule.

The METARULE algorithm most resembles the AQ family of inductive-learning programs (Michalski 1990). The algorithm evolved during development. The first version could generalize about classes of features—a hierarchical learning capability similar to that of OTIS (Kerber 1988). This capability was dropped from the final version when it became clear that features are best considered individually rather than as members of classes.

METARULE's method of incorporating domain knowledge underwent a fundamental change during development. The core of METARULE is a general learning program and is not tailored to a specific domain. Originally, domain knowledge was incorporated as commonsense rules to guide the search and derived features. The commonsense rules were propositions such as "Defects are associated with abnormal cases." These propositions were used to prevent the search from exploring primrose paths. Derived features included LAD-DISTRIBUTION, which was added to feature descriptions of cases showing a pattern of defects consistent with stenosis of the left anterior descending (LAD) coronary artery. I found that the search spent far too much time elaborating portions of the kernel that were not productive for thallium. The kernel contained many rules that were useful in general learning problems but not specifically for thallium. Tailoring the inductive kernel for the domain proved a much more effective way of creating an efficient search than imposing knowledge-based constraints on the use of a general-purpose kernel.

Development, Deployment, and Evaluation

The development strategy for TDW was a modified version of rapid prototyping. The rapid prototyping strategy is useful when users' require-

ments are not well understood (perhaps even by the users themselves) and must be discovered along the way. I believed, correctly, that I understood the functional requirements well at the outset. However, I needed to manage the technical risk imposed by incorporating machine learning and vision in an expert system.

The machine-learning component involved the greatest technical risk; so, I began by prototyping METARULE. METARULE was prototyped in INTERLISP-D on a Xerox 1186 Lisp machine, then ported to C for delivery. The use of Lisp greatly speeded development because of the power and flexibility of the language and the INTERLISP environment for programming, debugging, evaluating, and changing code. The salutary effect of Lisp on development time can be gauged by the fact that porting METARULE to C took as long as the initial prototyping effort—six months.

I demonstrated TDW to USAFSAM sponsors and clinicians about every four months. This process allowed the future users to evaluate the user interface and function and request changes or additional features. Each new capability was added in two steps: First, the user interface was built and demonstrated to the users to show the concept of operation. I modified the concept and the interface, based on users' comments, then built the capability. Once a capability was built, I never had to modify it. Except for the METARULE prototype, I rarely had to discard code during development. This point is significant because many expert system projects that plan to use rapid prototyping find that there is never enough time or money to throw code away and start over.

This strategy can be called *incremental prototyping* because it falls between rapid prototyping and incremental development. Like rapid prototyping, the users could specify their requirements as changes to a prototype rather than having to write them down cold. As in incremental development, I developed the riskiest parts first and could have aborted the project should TDW have proved infeasible. This strategy worked well for TDW. I recommend it for systems in which the user requirements are largely (but not completely) understood at the beginning, and considerable technical risk needs to be managed.

TDW was researched, developed, and tested during a 2.5-year effort involving about 3.75 person-years of effort at a cost of about $500,000 (including equipment and labor). It was deployed at USAFSAM in December 1990.

I evaluated TDW's ability to learn to read thallium imagery like a USAFSAM physician—that is, learn rules that predict physicians' grading of a thallium image—and learn to predict coronary artery disease from thallium imagery independent of the methods used by physicians—that is, learn rules that predict the outcome of angiography.

TDW was tested on training sets varying from 50 to 115 cases in size.

(Below 50 cases, METARULE began to tailor its rules to the quirks of individual cases rather than generalize.) A rule set's performance on cases not included in the training set always closely mirrored its ability to classify the training set itself.

TDW's best rule for grading images like a physician correctly predicts 82 percent of the gradings. This figure is comparable to the level of consistency among physicians: A study of intercoder reliability in thallium interpretation showed that physicians agree about 87 percent of the time (Trobaugh et al. 1978). TDW successfully learns diagnostic rules that replicate the judgment of expert physicians.

It is noteworthy that TDW's rules are much simpler than the physicians' own explanations of how they diagnose. In the talking protocols, physicians consider many subtleties. For example, the positioning of the heart within the chest could present an extra thickness of muscle wall to the camera, accentuating a hot region. However, the learned rules show that the presence of horseshoes and the characteristics of the largest single reperfusion defect adequately predict the physicians' diagnostic calls.

The best rule for predicting angiographic results correctly predicts 76 percent of the outcomes (62 percent of abnormal cases and 83 percent of normal cases). For comparison, USAFSAM's most skilled physicians correctly predict 74 percent of angiographic outcomes (60 percent for abnormal cases and 82 percent for normal cases). Learning by example as physicians learn on the job, TDW performs comparably with USAFSAM's best diagnosticians.

TDW will improve patient care and reduce cost by reducing the number of cardiac catheterizations. A study of diagnoses by less experienced physicians shows that TDW would have prevented as many as half of the catheterizations. There is no satisfactory method for measuring the number of catheterizations that TDW saves because it is difficult to tell how a physician would have diagnosed a case without TDW's assistance. Based on retrospective data from the 1989 case load, TDW should save 15 catheterizations this year. The cost reduction is difficult to calculate because USAFSAM physicians do not charge their patients for the procedure.

Discussion

There is a growing concern among practitioners of AI in medicine about the paucity of deployed applications. Discussion has centered on four issues: (1) some problems in this field are too hard to solve well; (2) the lack of computerized patient records in most hospitals gives AI

medical systems little to work on unless someone takes the time to enter patient information; (3) medical AI applications might automate the most enjoyable part of the physician's job; and (4) AI medical applications, like most new medical technology, improve patient care but increase costs. If AI medical applications reduced costs, wouldn't we see an explosion of deployed applications?

TDW owes much of its success to the selection of a diagnostic problem that is significant, requiring the use of emerging technologies, but not overwhelming. Feature extraction from thallium imagery can be performed using well-established pattern-matching techniques. If feature extraction had required extensive object reconstruction and scene analysis, TDW would still be where those technologies are—in the lab. The problem posed by learning diagnostic rules was harder than the vision problem, and the technological solution was correspondingly more mature. Machine learning is ready to start migrating from the lab into applications, and TDW is one of the first applied learning systems in medicine.

The patient data required by TDW, digitized imagery, already existed in computerized form. The typing required of the physician is no more than physicians already do in writing up their findings and diagnosis.

Because TDW encourages the physician to make judgments and compare them with TDW's, the workstation appears to be more of a colleague than a replacement. I hope that physicians will find it enjoyable to use.

TDW's development cost was substantial because it included research, development, and production—the system had to be robust and friendly enough for clinical use. Like other AI medical applications, it improves patient care at a cost.

I hope that TDW will encourage other researchers to deploy applications that apply emerging technologies such as machine learning to significant, manageable problems in medicine.

Acknowledgments

References

Burow, R.; Pond, M.; Schafer, A.; and Becker, L. 1979. Circumferential Profiles: A New Method for Computer Analysis of Thallium-201 My-

ocardial Perfusion Images. *Journal of Nuclear Medicine* 20(7): 771–777.

Garcia, E.; Maddahi, J.; Berman, D.; and Waxman, A. 1981. Space/Time Quantitation of Thallium-201 Myocardial Scintigraphy. *Journal of Nuclear Medicine* 20(771): 309–319.

Garland, L. 1950. On the Reliability of Roentgen Survey Procedures. *American Journal of Roentgenology* 64(2): 32–41.

Gerson, M. 1987. *Cardiac Nuclear Medicine*. Englewood Cliffs N.J.: Prentice Hall.

Kay, T. 1989. Personal communication at USAFSAM.

Kerber, R. 1988. Using a Generalization Hierarchy to Learn from Examples. In *Proceedings of the Fifth International Conference on Machine Learning*, 1–7. San Mateo, Calif.: Morgan Kaufmann.

Michalski, R. 1990. Learning Flexible Concepts: Fundamental Ideas and a Method Based on Two-Tiered Representation. In *Machine Learning: An Artificial Intelligence Approach*, volume 3, eds. Y. Kodratoff and R. Michalski, 63–111. San Mateo, Calif.: Morgan Kaufmann.

Michalski, R. 1983. A Theory and Methodology of Inductive Learning. In *Machine Learning: An Artificial Intelligence Approach,* volume 1, eds. R. Michalski, J. Carbonell, and T. Mitchell, 83–134. San Mateo, Calif.: Morgan Kaufmann.

Thompson, B., and Thompson, W. 1986. Finding Rules in Data. *Byte* 11:12.

Trobaugh, G.; Wackers, F.; Sokole, E.; DeRouen, T.; Ritchie, J.; and Hamilton, G. 1978. Thallium-201 Myocardial Imaging: An Interinstitutional Study of Observer Variability. *Journal of Nuclear Medicine* 19(4): 359–369.

Watson, D.; Campbell, N.; Read, E.; Gibson, R.; Teates, C.; and Beller, G. 1981. Spatial and Temporal Quantitation of Plane Thallium Myocardial Images. *Journal of Nuclear Medicine* 22(7): 577–584.