The Evaluation of Computer-Aided Diagnosis Systems: 
An FDA Perspective

David G. Brown, Center for Devices and Radiological Health, FDA (dgb@cdrh.fda.gov)

Abstract
Computer-aided diagnosis (CADx) systems have begun a successful transition from academic research to commercial implementation. FDA approval of CADx for PAP smear reading in 1995 and for breast cancer detection in 1998 were major milestones in this process. As Agency experience with these devices has increased, a consensus is emerging concerning factors to be considered in the evaluation required during the approval process. Key elements determining the nature of the proof of safety and effectiveness required by the Agency include the intrinsic level of risk associated with the device and the medical condition that it is meant to address, the precise claims made for the device, and the degree of oversight exercised over its use. The Agency expects that an increasing number of CADx devices will be submitted to it in the future and that guidelines will have to be formulated to assist manufacturers in navigating the approval process.

Introduction
Research into potential applications of computational intelligence to the medical diagnostic process has blossomed over the last several decades. Early decision tree structures developed by artificial intelligence pioneers, such as the expert system MYCIN [1] for the diagnosis (and suggested treatment) of blood infections, have been followed by the application of ever more sophisticated statistical tools such as neural networks [2]. Medical diagnostic imaging applications date from early in this period, with the work by Kruger [3] on automated diagnosis of pneumoconiosis being of particular significance, and have been the focus of great activity in recent times.

A number of factors are driving the introduction of CADx technology into the imaging area. Foremost of these are probably the quantity of diagnostically relevant information being created and the corresponding speed and facility of use of modern computers. The explosive growth of new imaging modalities over the last few decades shows no signs of abating. It has made vast pools of three dimensional and four dimensional (with time) data available—and all in digital form, ready for further computer analysis. Moreover, multimodality imaging is now of greater importance, whereby a stream of e.g., 3D cine-MRI, can be combined with ultrasound, CT, or nuclear medicine imaging data. This overwhelming crush of data cries out for computer assistance for the human observer. Furthermore, the combination of imaging with nonimaging information such as genomic or proteomic data (or simply patient age) lends itself to computer analysis.

Commercial exploitation of this research has been comparatively slow to be realized. Successful commercial applications for cardiovascular signal processing devices have become very prominent; however, serious imaging applications have only recently begun to appear. There are, after all, significant barriers to commercial introduction of these sophisticated devices. These include the premarketing review process of the Food and Drug Administration’s Center for Devices and Radiological Health (CDRH), the reimbursement approval process of the Centers for Medicare & Medicaid Services (CMS—formerly the Health Care Financing Administration (HCFA)), and of course the “approval process” of the medical community, the potential device purchaser. To achieve these “approvals” typically a considerable body of evidence must be developed in support of the device.

The CDRH seeks to provide “reasonable assurance of the safety and effectiveness of medical devices.” The definition of a device appears in section 201(h) of the Food, Drug, and Cosmetics (FD&C) Act and is quite broad. A device is: “...an instrument, apparatus, implement, machine, contrivance, implant, in vitro reagent, or other similar or related article, including a component, part, or accessory, which is . . .intended for use in the diagnosis of disease or other conditions, or in the cure, mitigation, treatment, or prevention of disease, in man or other animals...” This is interpreted by the Agency to include software as well as “material” products. The CDRH web site http://www.fda.gov/cdrh is a very good source of general information about Agency policy and guidelines used in carrying out its responsibilities as well as news and information concerning approved devices.

A total of five imaging CADx systems have been approved thus far for marketing in the United States by the Food and Drug Administration: 1. NeoPath (now TriPath Imaging) AutoPap (9/95) and 2.
System, Computer Digitizer for Screening Mammograms Image Analysis System

Classification Name: Mammogram Image Analysis System
Generic Name: Mammograms

Applicant: R2 Technology
PMA Number: P970058
Trade Name: IMAGECHECKER

Date Received: 12/16/1997
Decision Date: 06/26/1998
Product Code: MWE
Docket Number: 98M-06-18
Date Notice: 08/12/1998
Advisory Committee: Radiology
Supplement Reason: CHANGE

Expedited Review Granted? No
Approval Order Statement: The Image Checker M1000 is a computer system intended to identify and mark regions of interest on routine screening mammograms to bring them to the attention of the radiologist after the initial reading has been completed. Thus, the system assists the radiologist in minimizing observational oversights by identifying areas on the original mammogram that may warrant a second review.

Figure 1. Sample device approval information available on the FDA web site. "06/26/1998" (Decision Date) linked to additional information.

Neuromedical Systems (Tripath) Papnet (11/95) cervical cytology slide readers, 3. R2 Technology ImageChecker screening mammogram image analysis system (6/98), 4. Logicon RDA Logicon Caries Detector for diagnosis of dental caries (9/98), and 5. Deus Technologies Chest X-ray Medical Image Analyzer (10/2000). Thus, only a limited number of CADx devices have been approved to date, but they cover a fairly wide range of medical specialties. Further, judging from company web sites, there is potential for CADx devices in other areas, for example, dermatology (e.g., http://www.dermalert.com/) and optical in vivo cervical cancer screening (e.g., http://www.polartechnics.com.au/).

Further information is available on the CDRH web site concerning the approved devices. For example, there is a useful search facility available under http://www.fda.gov/cdrh/mda/mda-databases.html. Figure 1 gives the result of a premarket approval (PMA)

search for the manufacturer "R2." This is further linked to more complete documentation, including the "Summary of Safety and Effectiveness" that is issued during the final step of the approval process.

Of course, typically a great deal of detailed product information is available directly from manufacturer web sites. Figure 2 illustrates the type of descriptive material available for the Logicon Caries Detector at the Logicon web site.

Categories of CADx Devices

Computer-aided diagnosis devices may be divided into a number of different categories depending on the nature of their contribution to the diagnostic process. These range from image processing applications affecting the way a medical image is presented to an observer to formulation of a diagnostic “opinion.” Categorization of the device is important as a key step in the determination of the Agency’s approach to review of the device.

Image processing is the most direct form of computer assistance to the human diagnostician. This can take the form of edge enhancement, gray-scale alteration, noise suppression, or any of a large number of standard (or not so standard) image processing operations designed to make the image more “observer friendly.” Note that these may not all have only benign consequences: diagnostic information may be lost or obscured through imprudent processing. On the other hand, the
reconstruction algorithms required for producing usable CT and MRI images (required for the human but not necessarily required for a machine observer) may also be included within this category.

The next category of computer assistance is abnormality detection. Is there a suspect location on a radiograph of interest that should be examined by a radiologist as a potential lesion candidate? Is there a cell on a Pap smear slide that is someway “unusual” and should be examined by a cytotechnologist? In simplest form these devices should have high sensitivity (they will find almost all of the abnormals) but may have low specificity (they may “call out” many normals as abnormal). The original Papnet device, for example, was a very pure detection device. It was only used in the examination of slides that had already been examined and determined to be normal. Thus, it was a “2nd chance” system. After all, there may be several tens of thousands of cells on a single slide, and it is unreasonable to expect that even an experienced cytotechnologist will be able to catch every abnormality in the few minutes allotted for manual review of the slide. Similarly, it has been found that in mammography about half of the lesions missed on first reading of a film were visible in retrospect, i.e., had either been missed through lack of attention to the suspect area or had been seen but misinterpreted as benign.

The R2 Technology ImageChecker system is designed to address the “lapse-of-attention” problem. As illustrated in Figure 3, it identifies potential regions of interest from the detection of clusters of microcalcifications or spiculated masses. As configured at present, the radiologist first reads the film, then views the results of the ImageChecker analysis on a display monitor. The radiologist may then return to the original film to confirm whether or not anything was missed at a location indicated on the monitor.

Detection assists are of obvious potential benefit, but they are also not completely benign. For example, if they are high sensitivity/low specificity devices, they could lower the specificity of an inexperienced human observer, raising the number (and percentage) of patients sent to what could be expensive and/or hazardous follow up procedures. Similarly, if an inexperienced human observer becomes overly reliant on the device, some abnormal conditions may be missed, especially, for example, if the device has low sensitivity for some important but rare configurations (e.g., a cytology detection system that has not been trained adequately on certain rare but virulent cancer cell types). These are the sorts of problem areas, which must be addressed before device approval is possible.

Feature classification or discrimination assists make up the next category of CADx devices. In terms of the sensitivity/specificity distinction made earlier, the emphasis shifts from sensitivity toward increased specificity—and an increase in the “judgmental” quality of the assistance. With the Logicon Caries Detector, for example, a dentist is not seeking a general “tooth scan” to determine suspicious areas for examination, but is “asking the computer’s opinion” about a particular region. Is this an area of decay? And has the decay penetrated into the dentin? The program responds with the probability that this is indeed the case. For CADx systems of this kind probability of disease, possible treatment options, and even prognosis for cure would be possible outputs.

An even more advanced system could represent a full-fledged “second opinion.” Presented with an image, the system might respond “the image is normal, no further action is warranted.” Alternatively, a positive finding could be made. In both cases, a health care provider would have to make the final decision on the validity of the CADx systems’ suggestion. With these devices both the accuracy of the device algorithm and the nature of its relationship to the healthcare professional are of critical importance. Does that professional exercise meaningful oversight, or merely a pro forma administrative function?
Finally, the human may be dispensed with entirely and the computer diagnosis accepted without question. This is commonly the case with (frequently implanted) cardiovascular devices, but seems unusual in an imaging context. There is precedent for this, however: The NeoPath Pap smear reader has been approved for use as a prescreening tool. Up to 25% of the NeoPath reviewed slides may be found "sufficiently" normal to preclude cytotechnologist review. These slides do not undergo any human reading whatsoever.

**Safety and Effectiveness Concerns**

The degree of regulatory attention devoted by the Agency to a particular device is dependent not only on the device characteristics as outlined above but also on more general public health concerns: what are the prevalence and the seriousness of the medical condition being addressed? The greater the impact of possible device "failure", both in terms of numbers of individuals affected and the consequence for each individual patient, the greater the need for scrutiny of the device. Clearly, the CADx applications for screening mammography and Pap smear slide reading are of particular concern on both counts, whereas the dental caries detector operates in a less critical arena.

The characteristic of the particular CADx device under consideration most critical to determining the level of proof required during premarket review is the degree of authoritativeness it exhibits. The CADx categorization scheme outlined in the previous section reflects this concept. The more the CADx system behaves like an independent healthcare professional, the tougher the "Boards" it is required to pass. The more transparent its operation and the more meaningful the oversight exercised over it, the less stringent the review.

A critical role is played by the device claim or "indications for use" submitted by the device manufacturer. In general, a manufacturer’s claim for a certain capability demands evidence to support that claim. For example, if the manufacturer wishes to include a statement of the device sensitivity and specificity, then those must have been measured in a statistically valid study.

**CDRH Regulatory Process**

The CDRH has a broad range of regulatory tools available to assist it in meeting its device approval responsibilities. As previously indicated, these may be ascertained from material on the Agency web site or through direct contact with CDRH staff. In particular, the Center’s Division of Small Manufacturers, International and Consumer Assistance (DSMICA) is specifically charged with providing technical and regulatory assistance to small manufacturers to help them comply with Agency requirements for medical devices. Their assistance is available on the web site, by telephone at 800-638-2041 or 301-443-6597, by fax at 301-443-8818, and by email at dsma@cdrh.fda.gov.

Very briefly, a "device" product may not be an FDA "regulated article" (e.g., if it serves only an administrative or forensic purpose), or it may be a sufficiently "low risk" device that only certain general controls apply to it. These include such measures as facility registration, product listing, good manufacturing practices/quality systems regulations (GMP/QSR), and medical device adverse event reporting (MDR). The next level of regulatory control requires submission of notification of intent to market a device (referred to as a “510(k)" from the enabling provision of the FD&C Act). Finally, for innovative and potentially higher risk devices a Premarket Approval Application (PMA) must be submitted.

The Center’s Office of Device Evaluation (ODE) is charged with carrying out the premarket device review process. Communication and consultation with the relevant review group within ODE should be of the highest priority for anyone contemplating introducing a new (CADx or other) device into commerce in this country. In addition to contacting DSMICA, the information obtained from the web site at http://www.fda.gov/cdrh/premarket.html may be of particular interest and of assistance in identifying the appropriate ODE component.

**Conclusion**

Computer-aided diagnosis devices are making the transition from the research laboratory to the clinical marketplace. The Food and Drug Administration's Center for Devices and Radiological Health seeks to assure the safety and effectiveness of these and all medical devices. It operates within a flexible, risk-based framework, and is currently striving to develop a consistent and reasonable set of guidelines to assist in the review process for these innovative products.

**References**