

SCAR 2004 TRIP Session: CAD in Breast Imaging

Maryellen L. Giger, Ph.D.

**Professor of Radiology, the College, & the Committee
on Medical Physics**

Director, Graduate Programs in Medical Physics

University of Chicago

(m-giger@uchicago.edu)

**Supported in parts from grants from NIH-National Cancer
Institute, U.S. Army, American Cancer Society, and Whitaker
Foundation.**

COI Statement: M. Giger is a shareholder in R2 Technology



SCAR TRIP & CAD

- Explore how CAD can help deal with large medical image data sets
- This will be quite evident especially in screening programs such as screening mammography

The benefit of a medical imaging examination in terms of its ability to yield an accurate diagnosis depends on:

- **Quality of the image acquisition
and**
- **Quality of the image interpretation**

Image interpretation has been provided by the human

Limitations of Human Eye-Brain System

- Search & perception problems
- Image noise (which includes confounding normal structures, i.e., **structure noise**)



From C. Vyborny

AND vast amount of data for human interpretation

- Especially in cancer screening programs
- Makes lesion detection a burdensome task
- Impacts radiologists' workloads
- Causes oversight errors
- For example, in mammography:
 - Screening of asymptomatic women
 - Currently, the best method for early detection
 - Radiologists, however, do not detect all cancers on mammograms.

Retrospective Analysis of Screen Detected Breast Cancers

	<u>Year</u>	<u>% visible on prior (blinded)</u>
Harvey et al.	1993	41 %
Vitak	1998	25 %
Burhenne, et al.	2000	27 %
Yankaskas, et al.	2001	29 %

Variability in the interpretation of screening mammograms by U.S. radiologists

Beam C, Layde PM , Sullivan D

Archives of Internal Medicine 156: 209-213, 1996

- Compared interpretation of a set of 79 screening mammograms by a 1992 random sample of 108 radiologists from ACR-accredited mammography centers across the United States.
- **53 percent range of variability between the minimum and maximum sensitivity among radiologists in the sample.**
- While some radiologists referred 100 percent of women with cancer for biopsy, others referred only 47 percent.

The Solution: CAD

- **Incorporation of computer technology into the image interpretation component.**
- **Computer-aided diagnosis is a diagnosis made by a radiologist who takes into consideration the output from a computer analysis of an image.**
- **The final decision is made by the human.**
- **Goal is to reduce search errors, reduce interpretation errors, & reduce variation between and within observers**

Generic Flowchart for the Computerized Analysis of Breast Images

Digital Image Data

Segmentation of Breast Border

PreProcessing

Lesion Extraction

Feature Extraction

Classifier

Computer Output

Examples

- Peripheral equalization
- Noise reduction
- Feature filter, e.g., spiculation
- Region growing
- Lesion segmentation

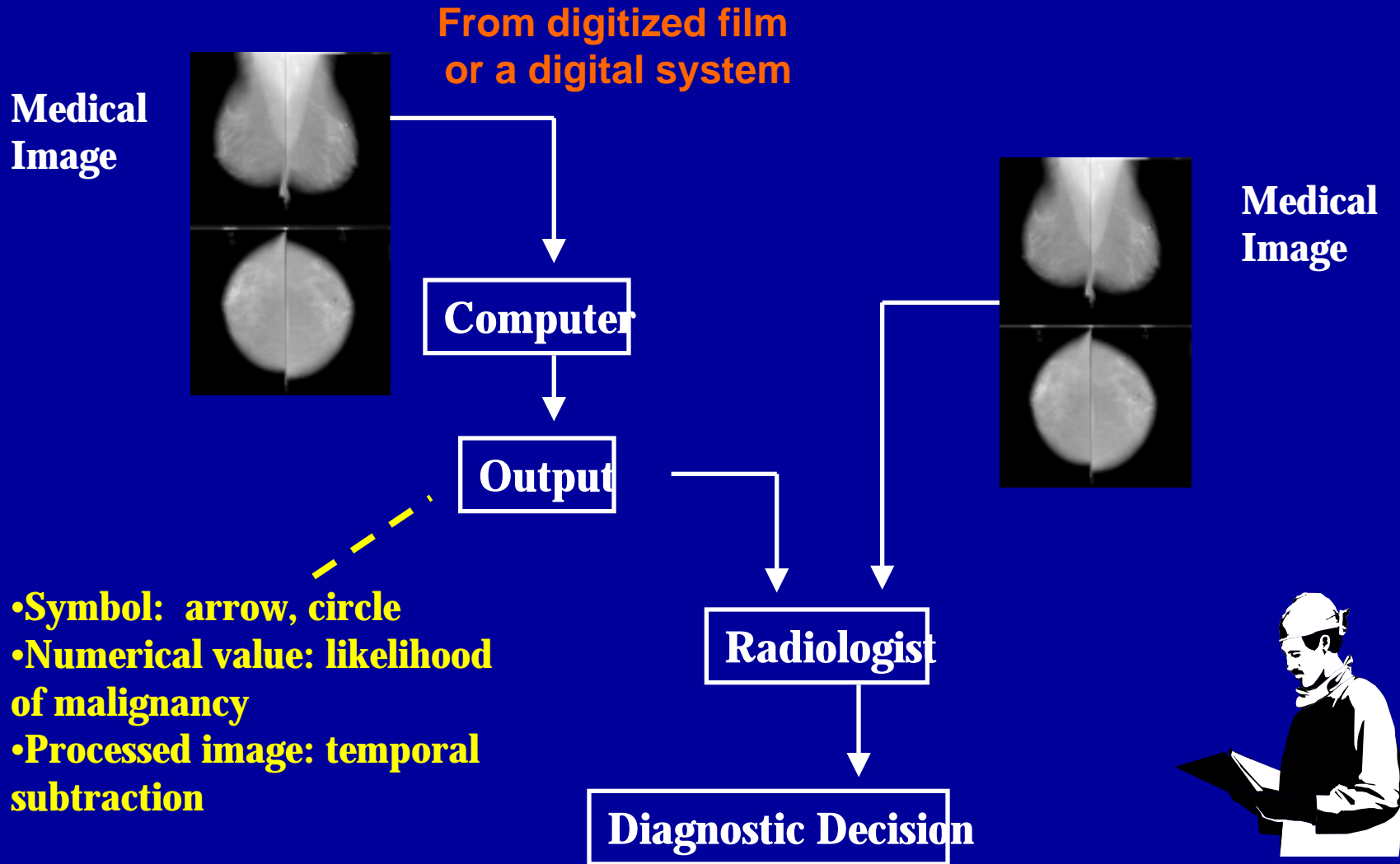
- Masses: spiculation, shape, asymmetry
- Calcifications: size, contrast, clustering

- LDA, ANN, rules, hybrid

- Location of lesion
- Likelihood of malignancy

Mathematical
descriptors

Computer-Aided Diagnosis

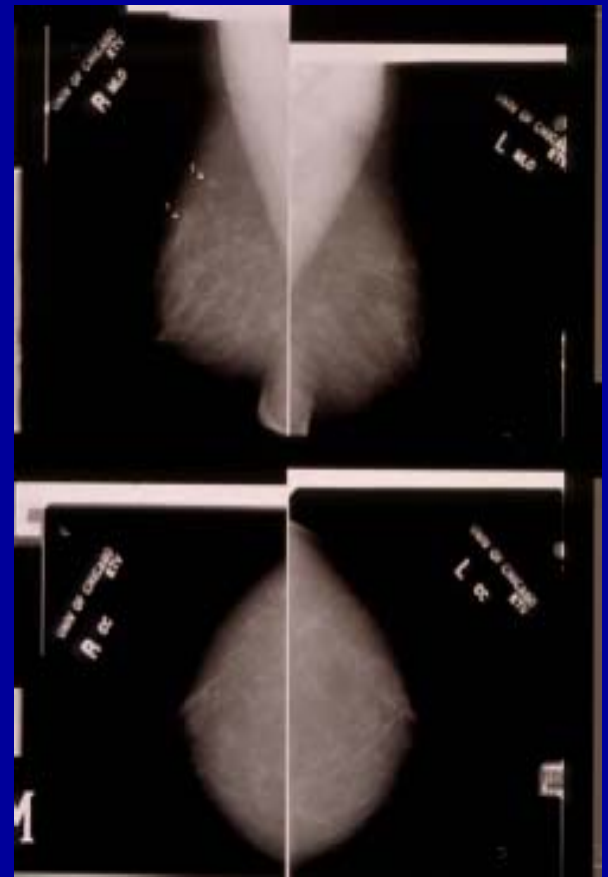


Computer-Aided Detection

- Many cancers on mammograms are missed because they are not NOTICED by the radiologist.
- Use computer output to direct a radiologists' attention to regions on a medical image that the computer deems to have features associated with cancer. (like a spell checker)
- The diagnostic workup and assessment, along with the final decision on patient management, is made by the radiologist.
- Note -- In order to be useful, the computer need not be "better" than the radiologist, but rather be able to detect at least some lesions that the radiologist might miss.

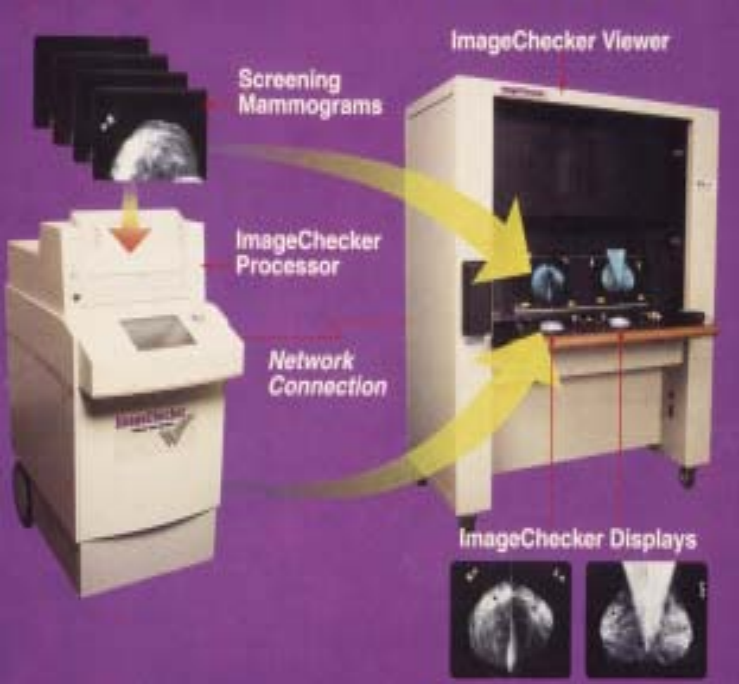
University of Chicago 1994 Prototype System for Computer- Aided Detection

Detection of Mass Lesions & Clustered Microcalcifications



Research has lead to Commercial Systems for Computer-Aided Detection in Mammography

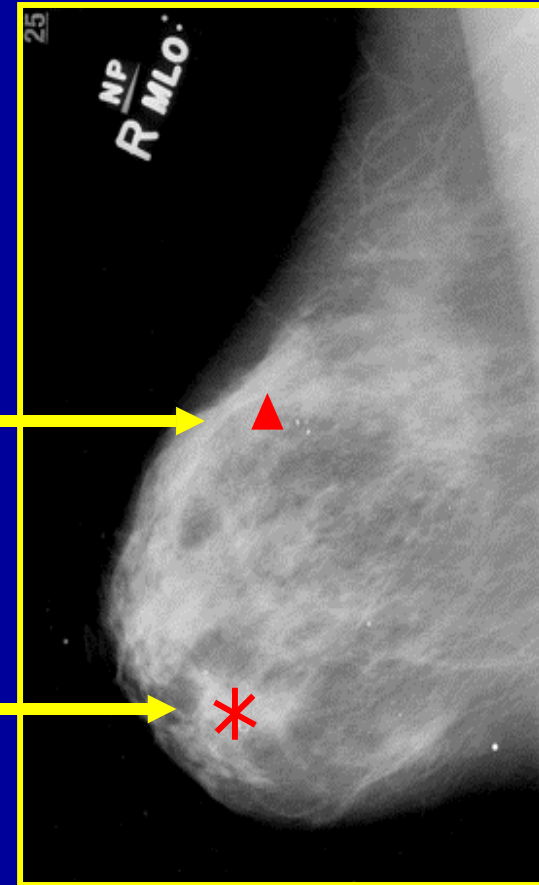
- R2 Technology (FDA approval 1998)
 - CADx (FDA approval 2002)
 - iCAD (ISSI) (FDA approval 2002)
 - Others (Kodak ...)
- Merged
In 2003
- See their websites for specs. In general, similar in sensitivity, however, the number of false marks differ. (Note, cannot consider TP without FP)
 - Differ in terms of display, floor space needed, integration with FFDM, integration with PACS, etc.



R2

A "▲" marks calcifications

A "*" marks masses or distortions.



Can be stored in DICOM fields or printed for patient folder



MammoReader™ is as easy as
1, 2, 3 to use:

1

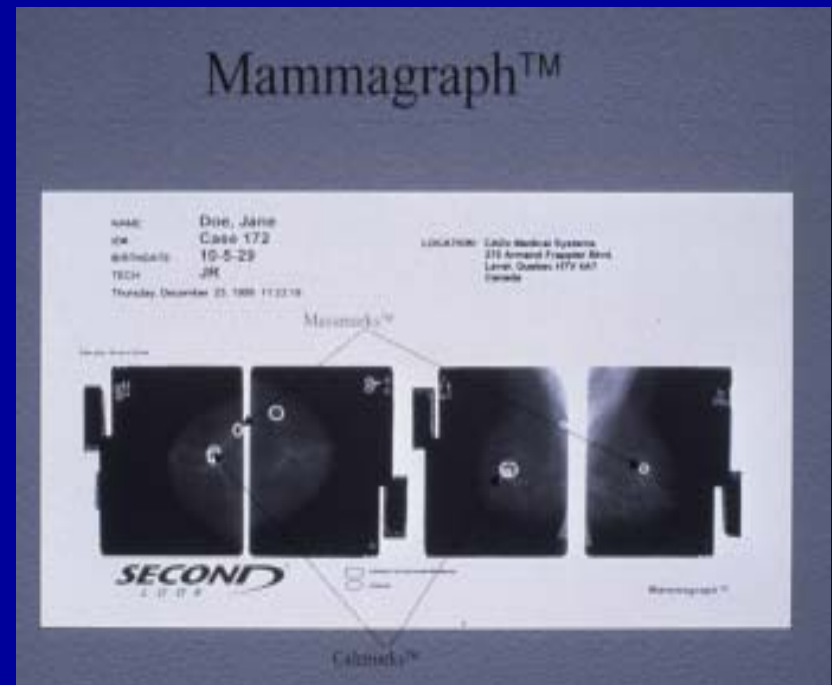
Load existing mammogram
film into the ISSI X-Ray
Scanner



**Example of paper output --
Can be put into patient folder**

iCAD

**iCAD resulted from the merge
of Howtex and ISSI. Now iCAD
has merged with CADx.**

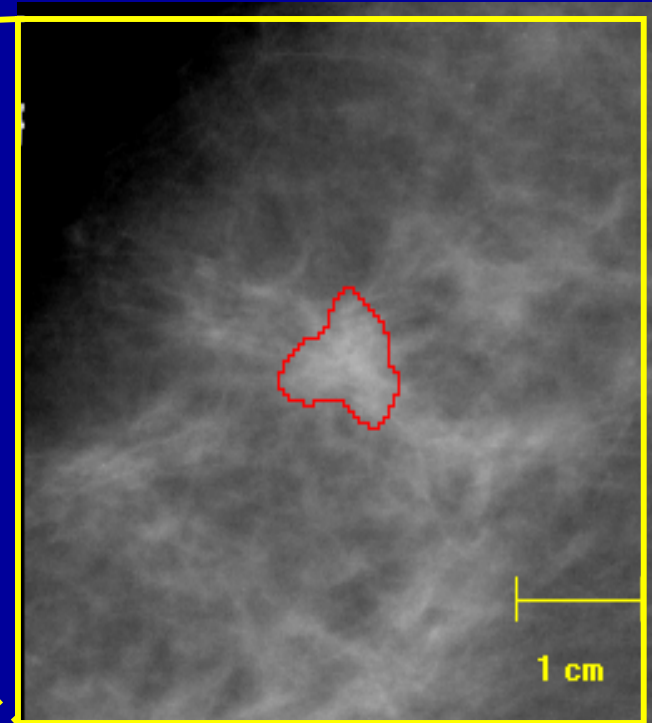
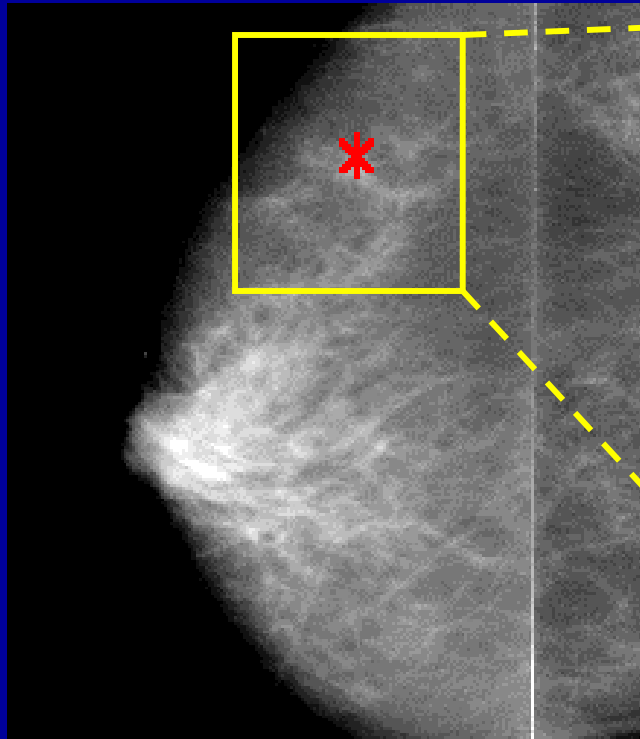


R2

Improvements in CAD Display

current CAD output

touch screen CAD output



 = area of a "mass"

displays outline of the "mass"
detected ("seen") by the code

More information being given to users

Evaluation of CAD Systems

- Performance of the computer detection method alone in terms of sensitivity and false-positive marks per image.
- Investigation of the effect that a CAD system has on radiologists' performance when it is used as an aid is the ultimate test.

Potential Contribution of Computer-aided Detection to the Sensitivity of Screening Mammography

Burhenne LJ, Wood SA, D'Orsi CJ, Feig SA, Kopans DB, O'Shaughnessy KF, Sickles EA, Tabar L, Vyborny CJ and Castellino RA (*Radiology* 2000; 215 : 554-562)

- 13 institutions contributed all consecutive, biopsy proven cancers detected by screening mammography (2 years)

CAD sensitivity (1,083 consecutive breast cancers)

microcalcifications	98.3%*	(399/406)
masses	85.7%*	(580/677)
all cases	90.4%*	(979/1083)

FP CAD marks
(347 normal exams)

0.5* marks / image

* updated in Castellino et al, *Radiology* 2000;217 (P), 400

Computer aided detection (CAD) in screening mammography: Sensitivity of commercial systems in detecting architectural distortions

Baker et al, AJR 2003; 181, 1083-1088

Malignant Architectural Distortions

<u>Sensitivity</u>	<u>System X</u>	<u>System Y</u>	<u>p value</u>
case (n = 27)	48%	19%	p = 0.027
image (n = 51)	31%	10%	p = 0.01
FP marks/image	0.70	1.27	p = <0.0001

Evaluation of CAD Systems

- Performance of the method alone in terms of sensitivity and false-positive marks per image.
- Investigation of the effect that a CAD system has on radiologists' performance when it is used as an aid is the ultimate test.
 - Observer studies that measure performance of radiologists interpreting alone and also interpreting with the use of computer assistance.
 - Prospective clinical studies

Computer aided detection in screening mammography: prospective study of 12,860 patients in a community breast center

Freer TW and Ulissey MJ - *Radiology* 2001; 220: 781-786

12,860 consecutive screening mammograms (12 months)
prospectively interpreted sequentially as follows:

- initial interpretation without CAD prompting
- immediate re-evaluation of CAD prompted areas
- tracked cases where CAD changed from “-” to “+”
- tracked all biopsies performed, cancer detection rates etc

Cancer detection rate -- increased 19.5%
(from 3.2 to 3.8 cancers/1,000 screens)

Recall rate -- increased by 18.5%
(from 6.5% to 7.7%)

Changes in Breast Cancer Detection and mammography Recall Rates after the Introduction of a Computer-Aided Detection System

Gur et al. *J. NCI* 96: 185 - 190, 2004

- Prospective study with **separate** mammographic cases for w/o and w/ CAD
 - **56,432** screening mammograms w/o CAD
 - **59,139** screening mammograms w/ CAD
- **The introduction of CAD into the practice was not associated with statistically significant changes in recall and breast cancer detection rates.**
- Editorial questioned whether or not the increased clinical use of CAD was due to allowed reimbursement.

Computer-Assisted Detection for Digital Mammography (FFDM)

- Commercial FFDM systems with CAD already exist (end of 2003 - 250 FFDM systems with CAD now)
- FFDM will expedite use of CAD since now is a “push button” operation only
- CAD output already incorporated in DICOM headers
- Performances are similar though may increase with optimizations

	<u>Sensitivity</u>	
	<u>Screen/Film</u>	<u>FFDM</u>
m/calcs (97%)	399/406 (98%)	33/34
masses (84%)	580/677 (86%)	47/56
Total (89%)	979/1083 (90%)	80/90
FP marks/image	0.55	0.5

Note

- Need **many** cases for a **DETECTION** study since number of cancers in a screening population is small.
- Need careful design of the study.

Computer-Aided Detection Systems in Clinical Practice

- Approximately **1,200** units (including FFDM)
- In 2003, it was estimated that **8,000,000** screening mammograms were read with CAD assistance (United States)
- Represents approximately 20 – 25 % of all screening mammograms done yearly in the United States

Challenges -- Computer-Aided Detection

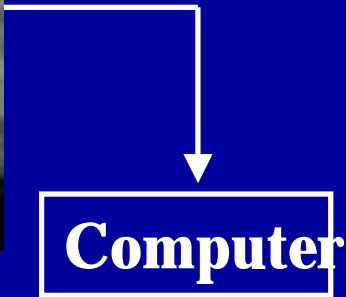
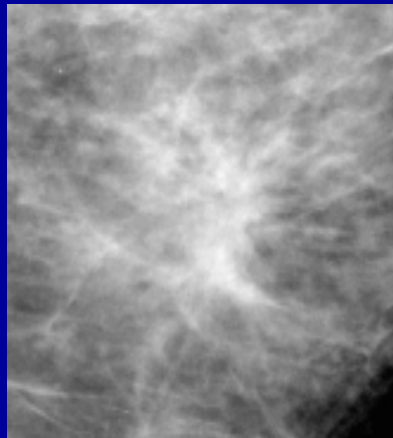
- Need more prospective clinical studies with different patient populations and different radiologists.
- Mass detection algorithms need to improve to a higher sensitivity without increase in FP rate (currently it appears that some radiologists may ignore mass detection output).
- Need to include and improve the detection for other radiographic signs of cancer, e.g., architectural distortion.
- Need to assess the rising practice of radiologists using the computer-detected microcalcification cluster output as a primary reader.
- Need to assess the potential role of CAD for detection as a pre-screener (identify “normal” cases that do not need to be looked at by a mammographer)
- Need to assess computer-aided detection on other modalities if they are used in screening of high-risk patients (e.g., sonography, MRI)

Explore Future Opportunities -- Beyond Detection

Computer-aided diagnosis in the work-up of suspect lesions: malignant vs. benign lesions

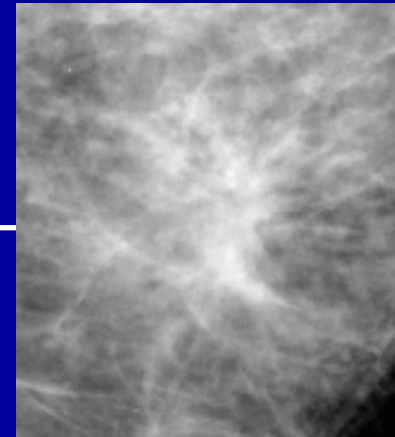
- **Use computer output to help characterize and potentially indicate a computer-determined probability of malignancy of a found lesion**
 - **Currently characterizing clustered microcalcifications and mass lesions**
- **The final decision on patient management is still made by the radiologist**

Computer-Aided Diagnosis



Output

97%



Suspect lesion

Note that the computer is analyzing a lesion that has already been found - either by human or computer

Radiologist



Diagnostic Decision
(Biopsy Recommendation)

Computer performances are high

Improving breast cancer diagnosis with computer-aided diagnosis

Jiang Y, Nishikawa RM, et al.; *Acad Radiol* 6: 22-33, 1999

A potential of CAD is to improve the sensitivity and specificity of interpretation in the task of lesion classification.

Effectiveness of Computer-aided Diagnosis -- Observer Study with independent Database of Mammograms

Huo Z, Giger ML, Vyborny CJ, Metz CE
Radiology 224: 560-568, 2002

A potential of CAD is to move the performance level of general-practice radiologists up to that of expert mammographers.

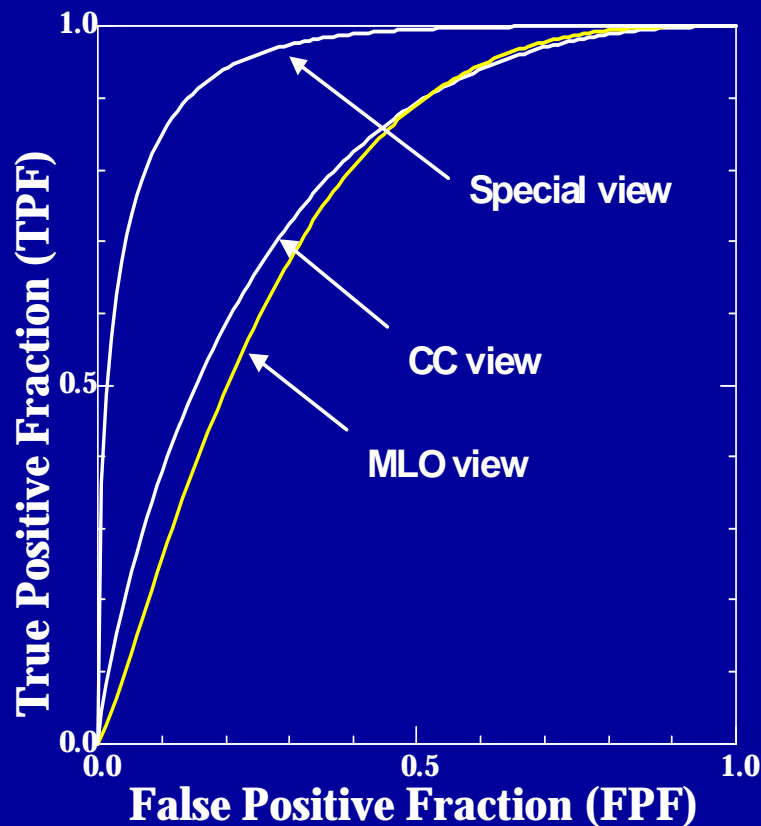
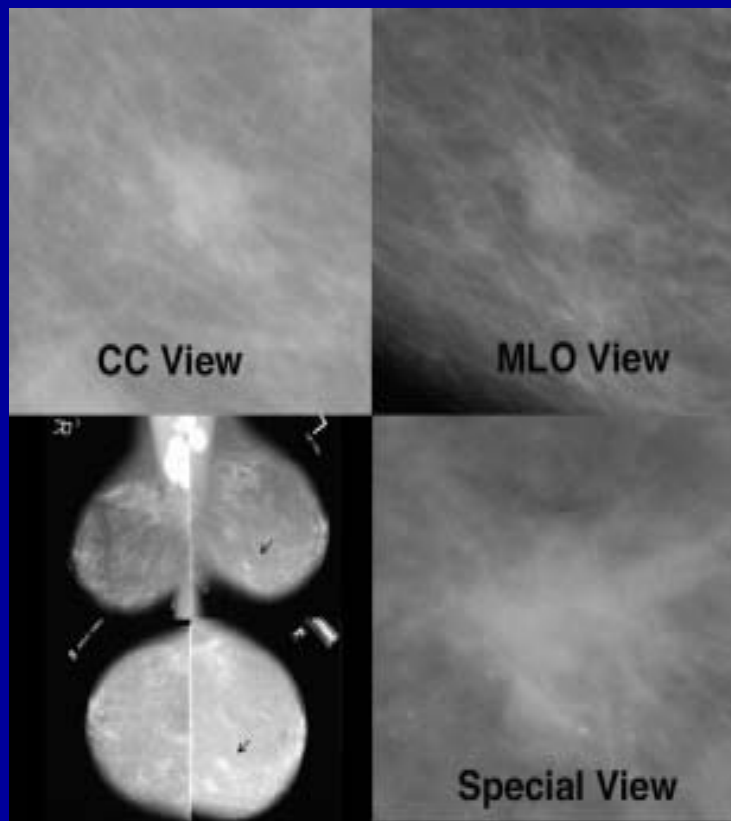
Further Developments in Computerized Diagnosis Algorithms

-- Use of Multiple Images

- Information from special views, as well as CC and MLO views
- Information from images obtained over time
- Information from multi-modality images
- Use of 3D information, e.g., 3D ultrasound & MRI
- Display methods to relate computer output to end user

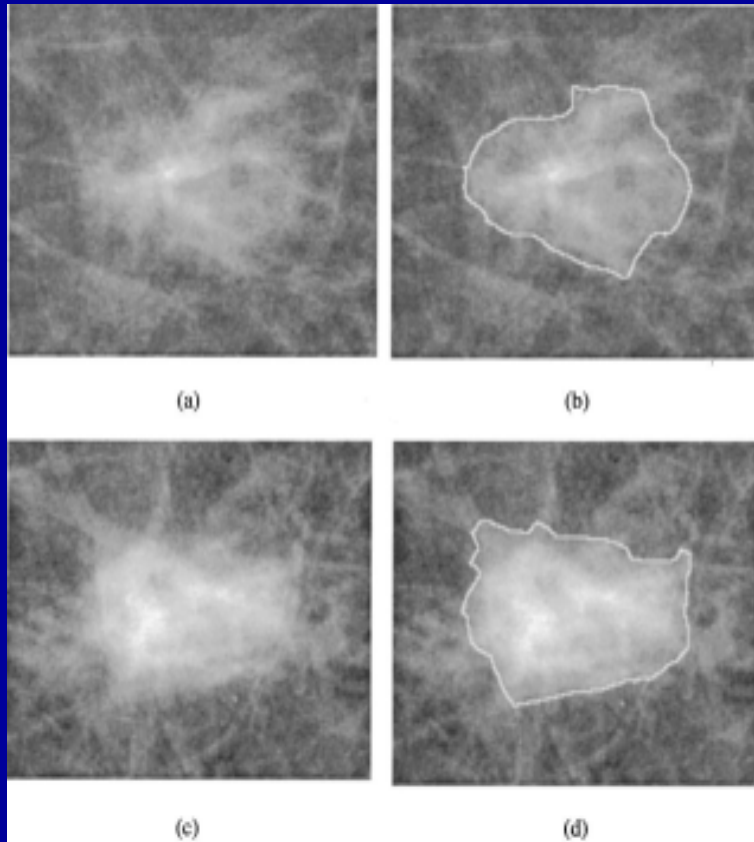
Comparison of Computer Analysis Special View Mammography to Conventional View Mammography

Huo Z, Giger ML, Vyborny CJ, *IEEE Trans on Medical Imaging* 20: 1285-1292, 2001

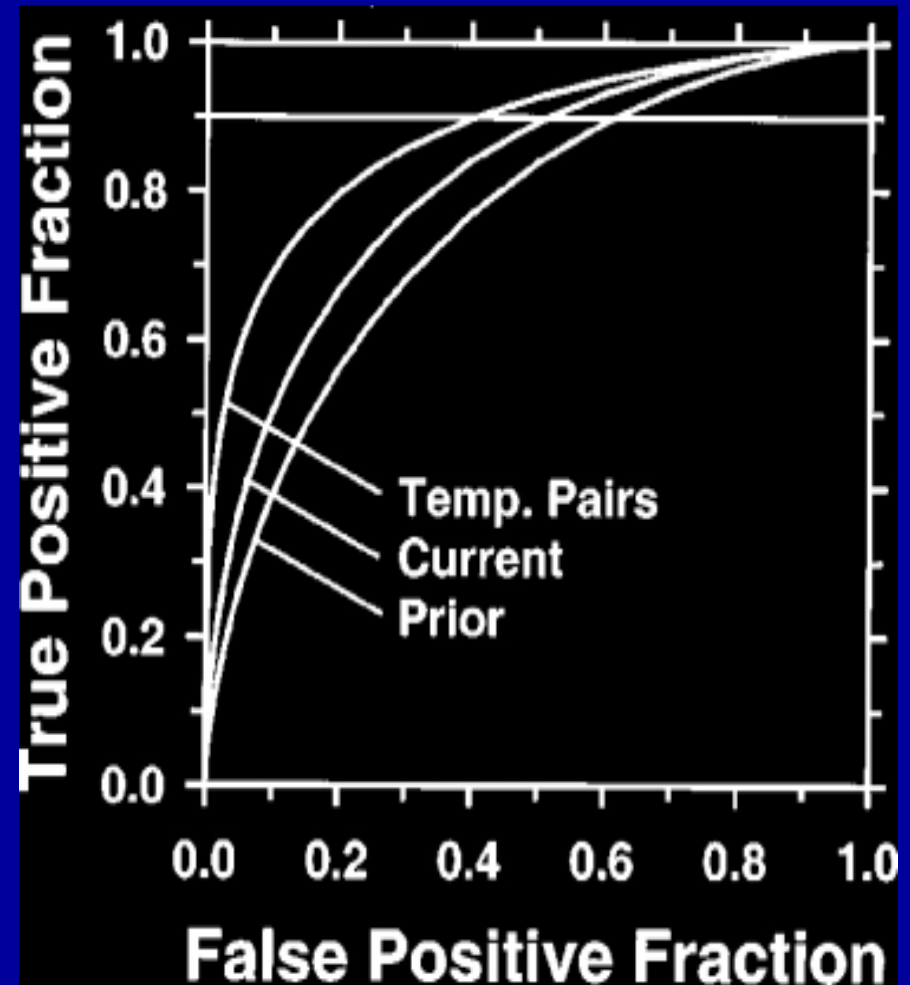


Task of distinguishing between malignant and benign lesions; Independent test set

Use of Temporal Information in the Classification of Masses

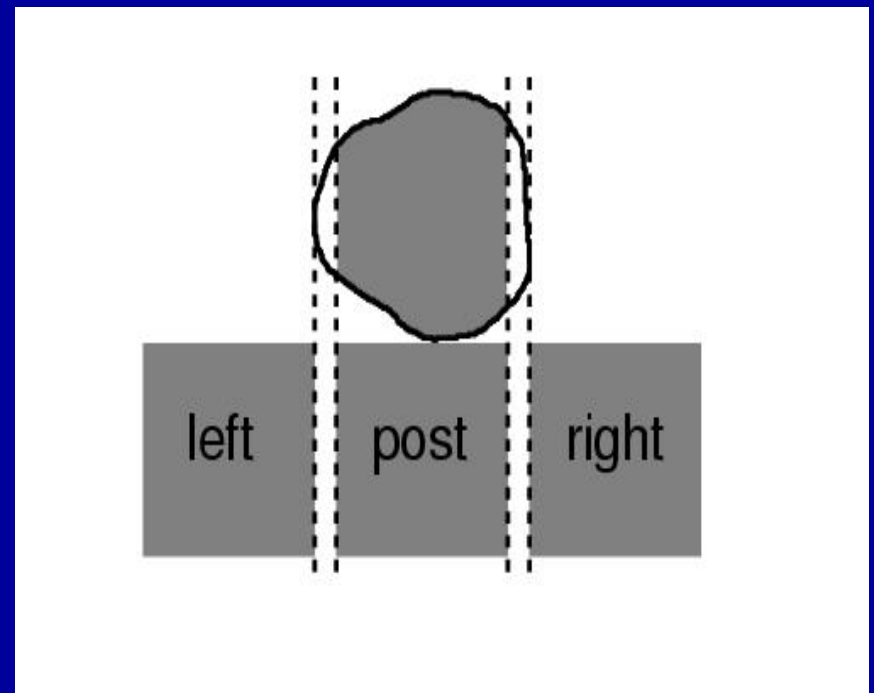


Features --> LDA



CAD in Breast Ultrasound

Ultrasound Feature: Posterior Acoustic Behavior



Breast Lesions on Sonograms: CAD in Nearly Setting-Independent Features and Artificial Neural Networks

Chen CM, Chou YH, Han KC, et al.
Radiology 226: 504-514, 2003

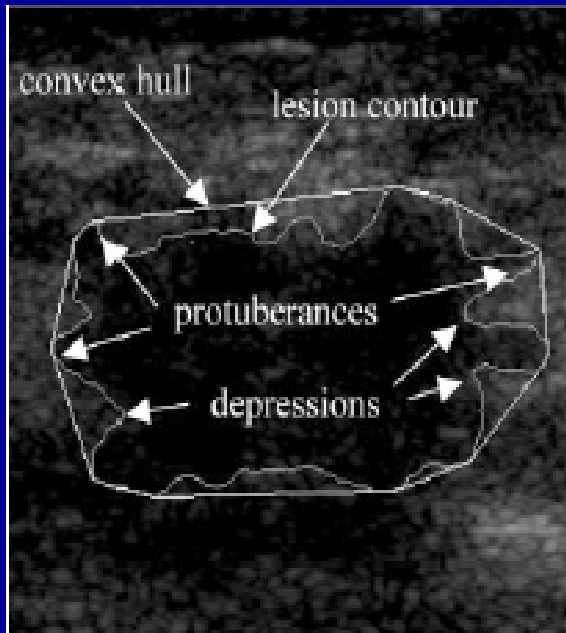


Figure 1. Sonogram shows the inner gray contour as the lesion border and the outer white polygon as the convex hull of the lesion. Protuberances and depressions in the malignant breast lesion are indicated.

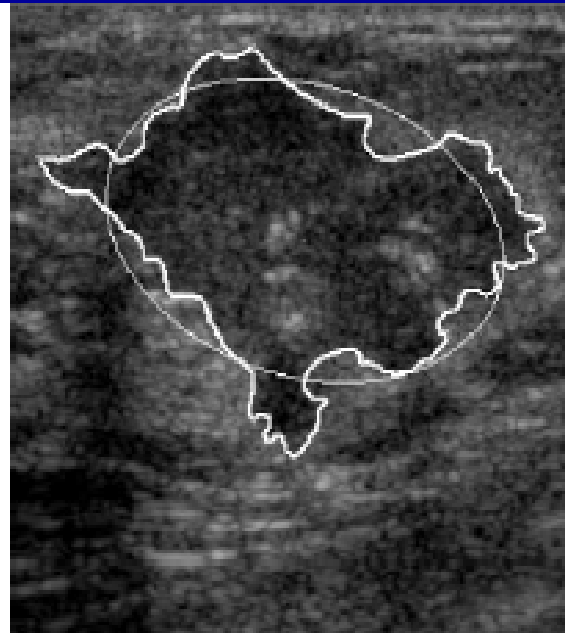
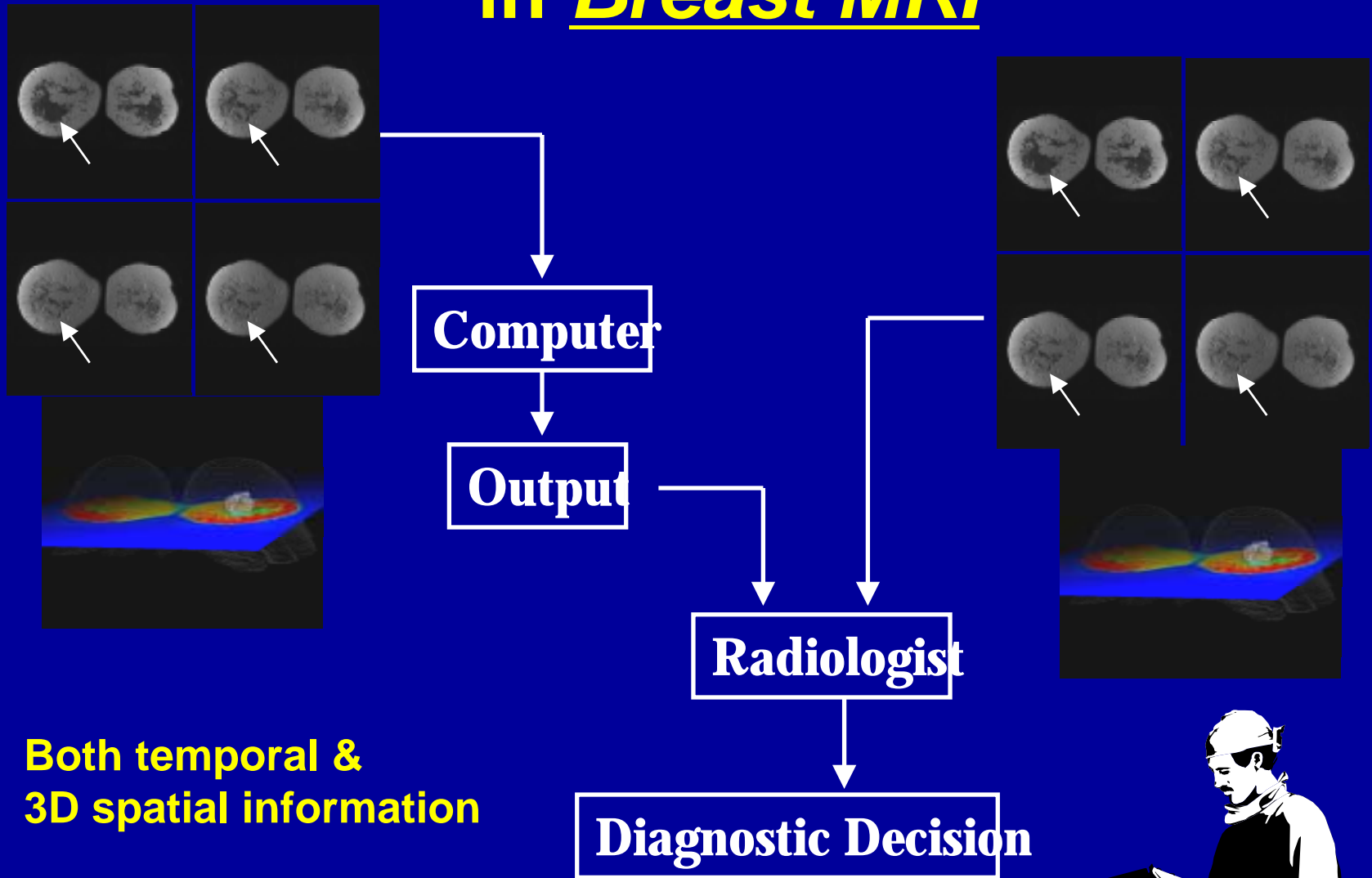


Figure 3. Sonogram shows the equivalent ellipse (thin line) of a malignant breast lesion, the boundary of which is marked by the thick line.

$A_z = 0.95 - 0.98$ in task
of distinguishing
between malignant and
benign lesions

Computer-Aided Diagnosis in Breast MRI

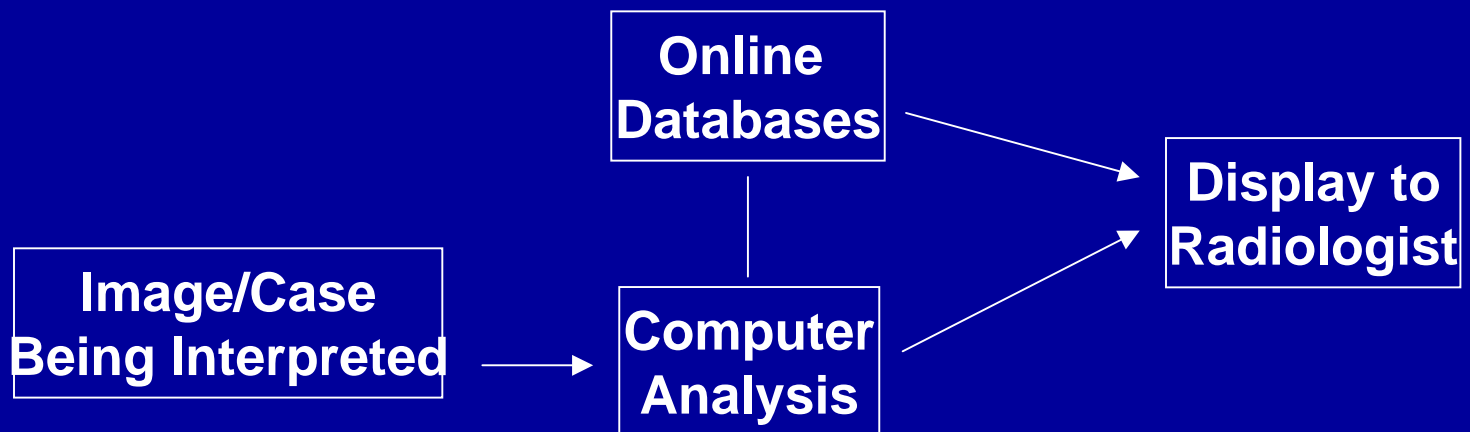


**Both temporal &
3D spatial information**



CAD & Use of Known Databases and Indices of Similarity

- Online databases with known truth on the disease state
- Online searches of the databases based on lesion characteristics or estimated probabilities of malignancy



Intelligent Workstation For Breast CAD

CARS 2000
RSNA 2001
IWDM 2002
RSNA 2002

Giger et al.

Benign
Case

Estimated probability of malignancy:
 19.49% (AID by case) 8.05% (AID by case)

ML CC (current) TRV Other view

Reference Library

Benign (Green)
Malignant (Red)

1) malignancy=0.9996 2) malignancy=0.9996 3) malignancy=0.9996 4) malignancy=0.9994

5) malignancy=0.9994 6) malignancy=0.9984 7) malignancy=0.9984 8) malignancy=0.9983

9) malignancy=0.9983 10) malignancy=0.9984 11) malignancy=0.9983 12) malignancy=0.9983

Case Relative to Reference Library

MA

Occurrence

Estimated Probability of Malignancy

Benign (Green)
Malignant (Red)

Training case: 4/18

Re-Input probability of malignancy(with AID):
 0% 100% No change

Re-Select Assessment Category(with AID):
 Benign Finding No change
 Probably Benign Finding
 Suspicious Abnormality
 Highly Suggestive of Malignancy

Re-Input patient management(with AID):
 biopsy follow-up No change

Go to next case

Intelligent Workstation For Breast CAD

Displays
For both
Mammography
&
Ultrasound

Benign
Case

The screenshot displays the 'Intelligent Search workstation' interface. On the left, a mammogram image is shown with a vertical scale on the right. Below it are 'Contrast' and 'Brightness' sliders. A 'Training case: 4/18' indicator is present. The main area shows a large finding labeled 'Spiculation' with a probability of '8.05% (AID by case)'. Below this are four image thumbnails: 'ML' (0.34), 'CC' (0.36), 'TRV (current)', and 'Other view'. A 'Reference Library' section contains 12 thumbnails, each with a sensitivity value (e.g., 1) sensitivity=0.9906). A legend indicates 'Benign' (green) and 'Malignant' (red). A 'Case Relative to Reference Library' bar chart shows the occurrence of Cyst (yellow), Benign (green), and Malignant (red) findings across estimated probability bins from 10% to 100%. The chart shows a high occurrence of Benign findings at 10% probability, which decreases as the probability increases, with Malignant findings appearing more frequently at higher probability bins.

Large medical databases

- **Need to collect with appropriate truth**
- **Need archival storage for database**
- **Need infrastructure to interact with database**
- **Need means to accommodate online reference libraries**

Challenges -- Computer-Aided Diagnosis

- Potential inherent road blocks to clinical acceptance (FDA) of computer-aided diagnosis -- too close to having a computer declare that a lesion is cancerous -- Examine gradual introductions:
 - Use of similar images from an online atlas
 - Graphical distributions of databases and unknown case
 - Presentation of feature information to characterize lesion
 - Ultimately, computer-estimated probability of malignancy
- Need prospective (pre-)clinical studies beyond ROC analysis -- looking at improvements in sensitivity, specificity, and positive predictive value
- Algorithms need to be tested on databases of cases that have never been biopsied as well as on lesions other than calcs and masses
- Need to examine performance with human detected lesions, computer-detected lesions, and computer-detected false positives
 - integration with computer-detection algorithms
- Need large databases that simulate the real clinical environment, i.e., multimodality, missing data, etc. displayed on a practical workstation
- Need to assess the potential as a primary reader -- computers may see more cancers in their lifetime than a typical mammographer!

**Recent Priorities Discussed at
BIROW II**

&

**is One of the Main topics at
BECON/BISTI 2004**

**(Symposium on Biomedical Informatics for
Clinical Decision Support)**

Databases

- Need database sharing culture change
- Nationally funded infrastructure (including user-friendly storing, user-friendly mining tools, means for quality control for truth)
- NCI funded research means NCI owns database, especially from clinical trials (ACRIN has agreed to put their databases into the shared database after the initial publication goes out)
- Prospectively investigators can increase their budget to pay for cost of database collection

Journal Publications

- **Checklist for those submitting papers to journals**
 - Educate investigators on training and testing
 - Evaluation methods - ROC analysis, sens, spec, etc.
 - Clarify software version
 - Clarify database used
 - Database -- independent? Round-robin?
 - Scoring
 - Establishment and clarification of truth

Potential Checklist for Publication

1. Description of the database in terms of disease type, lesion size, etc.
2. Appropriateness of database relative to the general population.
3. Method for classifier feature selection.
4. Method of CAD training (types of classifier, number of cases, training data, cross-validation, etc.).
5. Methods for truth determination
6. Method for scoring (i.e., determining true and false positives and negatives).
7. Method of assessing performance (e.g., ROC analysis) and the associated statistical analyses.
8. Results, in terms of not only sensitivity, but also specificity or false positive rate. Ideally, ROC or FROC, where appropriate, should be used to characterize performance at a range of operating points.
9. Results should also be related to the characteristics of the database (e.g., stratified appropriately by size or some other features) into agreed upon ranges (varies by CAD application).
10. For observer performance studies utilizing CAD, detailed study design, how CAD results were presented to observers, and what observers were told about CAD performance.
11. Discussion should relate results to clinical relevance where possible. For example, how many increased actionable lesions would be found?

CAD Evaluation Methods

- **Evaluation methodology for CAD -- need to develop and then educate users**
 - What is the appropriate way to measure sensitivity and specificity for individuals
 - Extension of ROC analysis
 - Case enrichment in database used in the evaluation (prevalence issue)
 - Who should be the target observers in the study
 - How to understand and educate radiologists who use CAD
 - What do you tell the observer on the observer study database and on the performance of the computer

- **This is just the beginning ...**
- **Much research is still needed in the application of computer vision and AI to the interpretation of medical images**
- **In the future, all medical images will have some form of computer analysis performed on them in order to benefit the diagnosis.**

- **Thank you.**