Lessons Learned from Computer-Aided Detection in Medical Imaging

Robert M. Nishikawa, Ph.D., FAAPM

Carl J. Vyborny Translational Laboratory for Breast Imaging Research,

Department of Radiology and

Committee on Medical Physics

The University of Chicago

Lessons Learned

- Most important factor in developing a CADe system is a high quality, large database
- Most important aspect of clinical implementation is the psychology of radiologists using CADe
- How CADe output is presented to the radiologist can affect radiologists' performance

Financial Disclosure

Robert Nishikawa:

- **shareholder in and receives royalties & research funding from Hologic, Inc.**
- **Paid consultant to Hologic, Inc and iCAD, Inc.**

Outline

1. Need for CAD
2. Commercial offerings
3. How a CAD system is developed from a clinical and technical point of view
4. Technical description of one application
5. Regulator approval
6. Clinical findings

1. Need for CAD in Mammography

- In mammographic screening:
  - FN rate is ~50%
  - FP rate is ~10%
- Cancer prevalence is 0.5%
- Nevertheless, screening mammography can reduce breast cancer mortality by up to 40%

1. Need for CAD

- Interpretation of an image is subjective
- Intra- and inter-reader variability
- Breast cancer screening is a dichotomy:
  - detection of microcalcifications
    - » small high contrast
    - » need to zoom image
  - detection of masses
    - » large low contrast
    - » masked or obscured by normal breast tissue
    - » pseudo-lesions

2. Commercial Systems

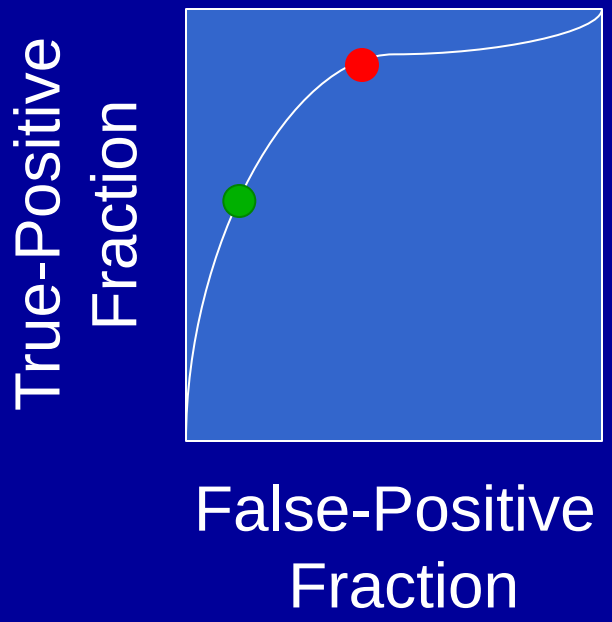
- CADe mammography
  - 4 approved systems in the USA
  - >75% of mammograms read with CADe
- CADe lung cancer
  - chest x-ray
  - chest CT
- CADe colon cancer
  - CT colonography

3. CADe System Development

- Develop database
  - ~1000 abnormal, ~1000 normal
  - Establishing truth can be difficult
    - » biopsy or follow-up
    - » consensus of experts
  - divide into 3 sets: development, training, testing
- Separate evaluation database
  - <~1000 cases

3. CADe System Development

- **Develop algorithm**
- **Train classifier (ROC analysis)**
- **Test (ROC analysis)**
- **Select operating point on ROC curve**



4. Technical Description of One CADe Application

- **Omitting**

5. Regulatory Approval

- FDA ensures safety and effectiveness
- CADe requires FDA PMA
- Changes to an approved system requires 510K approval
- PMA requires an observer study
  - 300 cases (new set of cases)
  - 15 radiologists
  - >$1,000,000
  - >1 year to complete study

6. Clinical Findings

- 7 clinical studies found 9.3% increase in sensitivity and a 12.4% increase in recall rate
- study design to evaluate CADe can be tricky
  - 4 clinical studies with flawed design
  - bias in estimating sensitivity

Clinical Issues

Medical

- indolent cancers

- benign lesions
- FN on aggressive cancer can be fatal
- FP adds cost and affect workflow

Parallels: CADe to ATR

Medical	Security
• indolent cancers	• guns carried by non-terrorists
• benign lesions	• water bottles
• FN on aggressive cancer can be fatal	• FN on targets can be fatal
• FP adds cost and affect workflow	• FP adds cost and affect workflow

Differences

- Mammography has 2 views of each breast and temporal comparisons
- Need to be concerned about radiation dose
  - retakes for ambiguous findings are not done

CADe as a Second Reader

0 radiologists detected without CADe

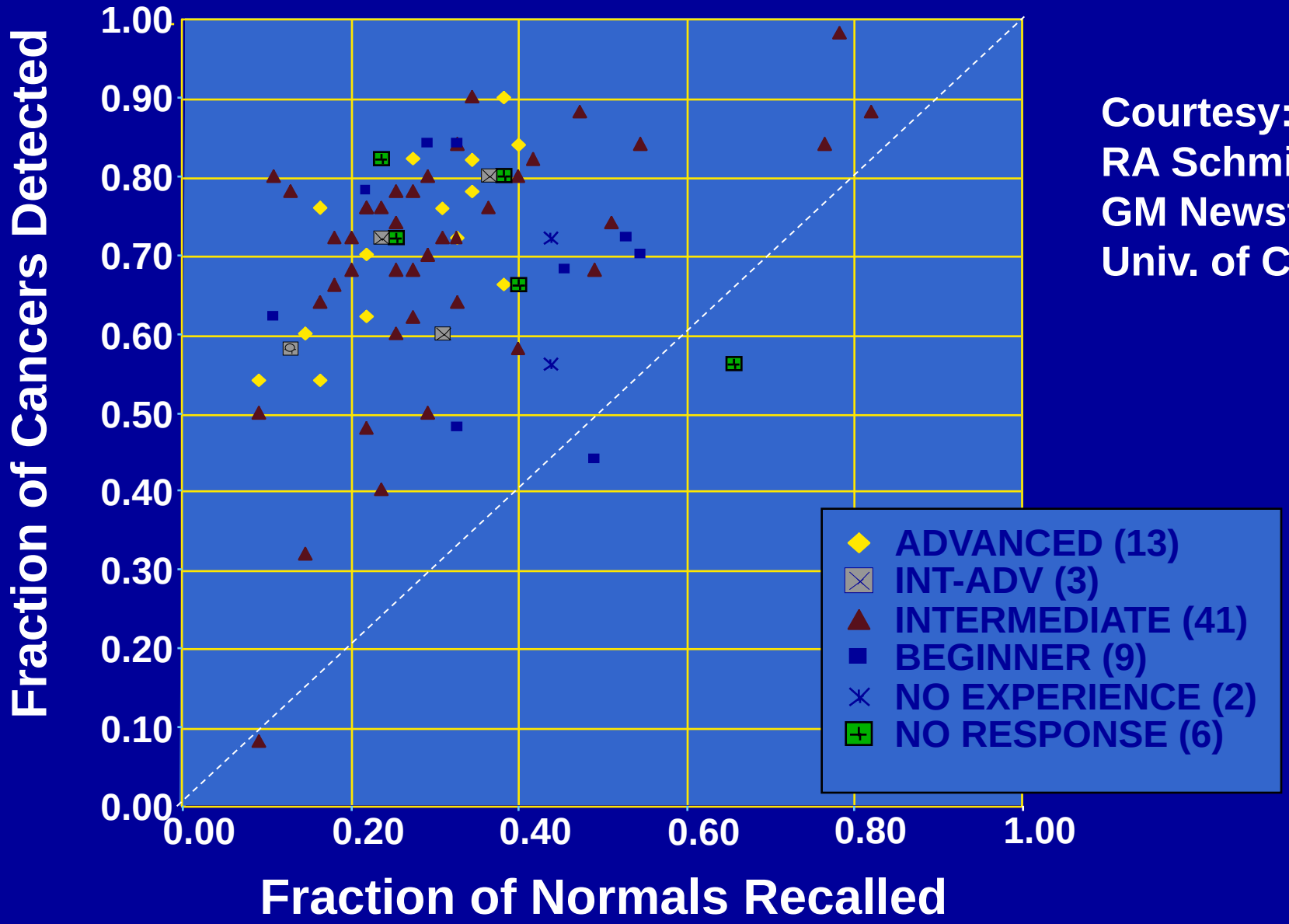
3 radiologists detected with CADe

5 radiologist ignored the correct CADe mark (lower asterisk)

Observer Study

- 8 radiologists reading 300 screening exams
- 69 cancers (all missed clinically)
- reading without CADe sensitivity = 0.549
- reading with CADe sensitivity = 0.603
- 9.9% in sensitivity (12.4% increase in recall rate)
- radiologists ignored 70% of TP marks

Radiologists' Variation in Screening Mammography



Courtesy:
RA Schmidt,
GM Newstead,
Univ. of Chicago

Psychology of Using CADe

- **Radiologist need to believe that CADe will be helpful**
  - **missed caner prevalence is 2 in 1000**
  - **CADe may mark 50% or 2 TP marks in 1000 cases**
  - **CADe FP marks will be 2000 marks**
  - **1 true mark for every 999 false marks**
  - **no feedback when you correctly found cancer or when you missed a cancer**

Human Detection Performance at Low Cancer Prevalence

- **Jeremy Wolfe et al.**

Prevalence	Miss Rate
50%	12%
1%	30%

- **“cognitively impenetrable"**

The CADe Learning Curve

Dean et al. (AJR 2006)

Time Period	Recall Rate	% Increase
Before CADe	6.2% (65/1047)	---
Months 1 - 2	13.4% (50/374)	116%
Months 3 - 21	7.8% (326/4157)	25%
Months 22 - 26	6.75% (59/874)	10%

(Increase in sensitivity was 7.6%)

Concurrent Reading with CADe

- CADe microcalcification detection is 98%
- Concurrent reading with CADe may reduce reading times
- Higher likelihood of a radiologist FN, if CADe did not mark the cancer
  - CADe mass detection is ~85%

Interactive CADe

- Karssemeijer has proposed using CADe interactively
- Radiologist queries suspicious lesions and is shown the CADe output
- Can reduce interpretation errors by radiologist
- Can improve radiologists' performance more than 2nd reader method

Lessons Learned

- Most important factor in developing a CADe system is a high quality, large database
- Most important aspect of clinical implementation is the psychology of radiologists using CADe
- How CADe output is presented to the radiologist can affect radiologists' performance