Optimization and performance standards of digital mammography systems

Professor David Dance

NCCPM, Royal Surrey County Hospital, Guildford, United Kingdom
The OPTIMAM2 project: funded by Cancer Research-UK

- Optimising the use of X-rays to detect breast cancers
- Investigating the performance of imaging systems using real and simulated images of breast cancers
## Main Colleagues in OPTIMAM

<table>
<thead>
<tr>
<th>NCCPM, Guildford</th>
<th>CVSSP, University of Surrey</th>
<th>University of Leuven</th>
<th>Addenbrooke’s hospital, Cambridge</th>
<th>St George’s Hospital</th>
<th>Jarvis Centre, Guildford</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ken Young</td>
<td>Kevin Wells</td>
<td>Hilde Bosmans</td>
<td>Mathew Wallis</td>
<td>Ros Given-Wilson</td>
<td>Julie Cooke</td>
</tr>
<tr>
<td>Alistair Mackenzie</td>
<td>Prem Elangovan</td>
<td>Emmy Shaheen</td>
<td>Paula Wilsher</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lucy Warren</td>
<td>Oliver Diaz</td>
<td>Nick Marshall</td>
<td></td>
<td>Charul Patel</td>
<td></td>
</tr>
<tr>
<td>Padraig Looney</td>
<td>Alaleh Rashidnasab</td>
<td>Lesley Cockmartin</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mark Halling-Brown</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>+ Students</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Standards and Guidelines

Commissioning and Routine Testing of Full Field Digital Mammography Systems

European guidelines for quality assurance in breast cancer screening and diagnosis
Fourth edition - Supplements
• Background
• Breast dosimetry and breast dose
• Choice of X-ray spectra
• Image quality assessment
  – Physical image quality measures
  – Quality measures using phantoms
  – Clinical measures of image quality
  – Do phantom measurements and clinical measures correlate?
• Conclusions
Albert Salomon (Berlin) published first radiographs of surgical breast specimens
Mammography 1980 and 2015

Direct exposure
film
High kV
No grid

Digital receptor
Appropriate X-ray spectrum
Grid

Image processing to optimise skin edge and improve contrast and display of detail
Mammography requirements 1

• Visualise:
  – small calcifications down to about 100 micron
    • limited by contrast, noise, system resolution & image processing
  – soft tissue masses – as small as 5mm or less
    • limited mainly by contrast, anatomical clutter (overlying tissue) and image processing
• Control the dose. It must be:
  – high enough so that calcs are visible against noise
  – no higher than necessary to control risk
  – BUT if it is too low, cancer detection will decrease
Risk Benefit Analysis

NHSBSP Report 54

REVIEW OF RADIATION RISK
IN BREAST SCREENING

Report by a joint working party of the NHSBSP
National Coordinating Group for Physics Quality
Assurance and the National Radiological Protection Board
Breast cancer screening
Age 50 to 70 in the UK

Lives saved and lost due to screening

- Lives saved due to screening: 3507
- Lives lost due to radiation: 0

Breast cancer screening for Age 50 to 70 in the UK.
Breast cancer screening
Age 50 to 70 in the UK

Lives saved and lost due to screening

- Lives saved due to screening: 3507
- Lives lost due to radiation: 26

Halve dose to save 13 lives

If halving dose reduces detection by 1%, 35 lives are lost
Conclusion

When the benefit of imaging is much greater than the radiation risk
we should concentrate on achieving sufficient image quality rather than on reducing dose.
Linear attenuation coefficients

Taken from work of Johns and Yaffe

- Carcinoma
- Gland
- Fatty tissue
Mammographic contrast

- 100μ calcification
- 1 mm glandular tissue

Graph showing contrast versus photon energy (keV) for two different tissue types.
Variation of SNR with photon energy (fixed energy deposited in receptor)
Transmission through 5 cm breast

Graph showing the transmission of energy (keV) through fat and glandular tissue.
OPTIMISATION

• A balance between dose and image quality

• Dose is influenced by
  – X-ray spectrum
  – Noise level acceptable
  – Efficiency of image receptor (DQE)

• Image quality is influenced by
  – X-ray spectrum (contrast) and dose (quantum noise)
  – Structure noise and electronic noise
  – Receptor unsharpness (MTF)
  – Focal spot size (geometric unsharpness)
  – Scatter
  – Image processing
  – Image display and viewing conditions
• Background
• Breast dosimetry and breast dose
• Choice of X-ray spectra
• Image quality assessment
  – Physical image quality measures
  – Quality measures using phantoms
  – Clinical measures of image quality
  – Do phantom measurements and clinical measures correlate?
• Conclusions
UK DOSE SURVEY 1981

- 5CM PHANTOM
- Entrance surface dose
- Dose range 0.9 to 45 mGy (!)

- kV range 24- 49 kV
- HVL range 0.2 mm Al to 1.7 mm Al
Breast dose measures

• Incident air kerma is a poor measure of breast dose

• Average glandular dose
  – Karlsson in 1976
  – Recommended by ICRP in 1987
Average glandular dose

- Cannot measure \textbf{AGD} on patients
- Incident air kerma \textbf{K} (without backscatter) can be easily measured or estimated
- Need conversion coefficients which relate \textbf{K} to \textbf{AGD}. In UK and Europe we use:

\begin{center}
\textbf{AGD} = \textbf{Kgecs}
\end{center}

- \textit{g}, \textit{c} and \textit{s} estimated using Monte Carlo modelling and simple breast model
  - tabulated against thickness and HVL
- Monte Carlo also used to design breast equivalent PMMA phantoms

D R Dance et al, PMB, 2000
SIMPLE GEOMETRICAL BREAST MODEL

Fine for quality control of AGD using breast equivalent phantoms or series of patient data

Does not give the true dose for individual patients
Simple breast model

Hammerstein 1979

- 5 mm adipose shield region
- Central region with mixture of glandular and adipose tissues
- Fraction by weight of glandular tissue in central region is known as the glandularity

Hammerstein et al suggested the use of 50% glandularity for dosimetry
$g$-factors

Breast thickness cm

$g$-factor mGy/mGy

- HVL 0.3 mm
- HVL 0.6 mm

D R Dance, PMB, 1990
Breast glandularity (UK)

Ages 50-64

Glandularity %

Breast thickness cm

D R Dance et al, PMB, 2000
PMMA equivalence age 50-64

D R Dance et al, PMB, 2000
EUREF dose standards for digital mammography

- Use PMMA phantoms 20-70 mm thick
- Clinically selected AEC and spectra
- Based on previous standards for screen-film and survey data from Holland, UK and Germany

- The **acceptable level** is the minimum acceptable standard
- However, it is recommended that systems operate as far as possible at a standard equal to or better than the achievable level

- REMEMBER that if the dose is too low, cancer detection will decrease
  - different receptors have different unsharpness and noise properties and will require different doses
Acceptable and achievable AGD levels

![Graph showing AGD levels against PMMA thickness cm]

- **AGD mGy**
  - **Acceptable**
  - **Achievable**

PMMA thickness cm

- 2
- 3
- 4
- 4.5
- 5
- 6
- 7

AGD levels:
- **0**
- **1**
- **2**
- **3**
- **4.5**
- **5**
- **6**
- **7**
AGD: 53 mm breast (45 mm PMMA)

98% < 2.5 mGy
Patient dose

- What dose level is currently used?

DR systems, 50-60mm, OB views

National DRL 3.5mGy

Avg. MGD (mGy)

2000D

2000D

Amulet

DS

Essential

Selenia

Selenia W

Giotto

Sectra

Inspiration

Novation
CR systems, 50-60mm, OB views

National DRL 3.5mGy

Average MGD (mGy)

Agfa 25X  Agfa 35X  Fuji 5000  Fuji Profect  Kodak  Konica

J Oduko, IWDM 2010
Patient dose

• What dose level is currently used?
  Answer: 0.8 – 2.8 mGy AGD for 50-60mm breasts

• Doses for CR are generally higher than for DR

• Different systems have differ noise and unsharpness and require different dose levels

• What dose level should be used?

We can measure how changing the dose changes the results of physical or test phantom measurements BUT how does it affect cancer detection?
Outline

• Background
• Breast dosimetry and breast dose
• **Choice of X-ray spectra**
• Image quality assessment
  – Physical image quality measures
  – Quality measures using phantoms
  – Clinical measures of image quality
  – Do phantom measurements and clinical measures correlate?
• Conclusions
Optimal energies for mammography

SNR for 1mGy breast dose vs. Photon energy keV

- 2 cm calcification
- 4 cm calcification
- 6 cm calcification
- 8 cm calcification

100μ calcification
Optimal energy (simple model)

Varies between approximately 16 - 23 keV

Increases with breast thickness
How can we adjust the spectrum?

• Mo target:
  – K X-rays 17.4 and 19.6 keV

• Rh target:
  – K X-rays 20.2 and 22.7 keV

• W target:
  – K X-rays 59.3 and 67.2 keV
Rh/Rh spectra

Spectrum after 5 cm tissue

Incident spectrum
K-edges

Mo  20.0 keV
Rh  23.3 keV
Ag  25.5 keV
8 cm breast 10% glandularity

Tube voltage (kV)

Dose at fixed CNR

Mo/Mo
Mo/Rh
Rh/Rh
W/Rh

25 26 27 28 29 30 31 32
Spectra for digital mammography

• Optimum spectrum varies with breast thickness and glandularity
• Mo/Mo is optimum only for 2 cm breasts
• Dose saving possible cf Mo/Mo when CNR (not contrast) is main constraint
Spectra for digital receptors:

For the same generating kV:

<table>
<thead>
<tr>
<th>Anode/Filter</th>
<th>Beam Energy</th>
<th>Breast size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mo/Mo</td>
<td>Lowest</td>
<td>Smallest</td>
</tr>
<tr>
<td>Mo/Rh</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rh/Rh</td>
<td></td>
<td></td>
</tr>
<tr>
<td>W/Rh</td>
<td>Highest</td>
<td>Largest</td>
</tr>
<tr>
<td>W/Ag</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Outline

• Background
• Breast dosimetry and breast dose
• Choice of X-ray spectra
• Image quality assessment
  – Physical image quality measures
  – Quality measures using phantoms
  – Clinical measures of image quality
  – Do phantom measurements and clinical measures correlate?
• Conclusions
Modulation transfer functions

MTF for DR is better than for CR
Noise in digital images

Arises from:

Quantum mottle
- fluctuations in no of X-ray photons detected and no of light photons/charge carriers produced per X-ray photon

Structure mottle
- generally small

Electronic sources
Relative noise power spectra contributions

Quantum is generally the largest overall contribution.

Low frequency quantum is largest for CR.

Quantum for amorphous Se decreases less with frequency.
DQE for DR better than for CR especially at higher frequency
How does the difference between MTF and DQE for CR and DR impact on:

Image quality assessed using test phantoms
Outline

• Background
• Breast dosimetry and breast dose
• Choice of X-ray spectra
• Image quality assessment
  – Physical image quality measures
  – Quality measures using phantoms
  – Clinical measures of image quality
  – Do phantom measurements and clinical measures correlate?
• Conclusions
Contrast detail test phantom

CDMAM phantom imaged with 50 mm PMMA
Can be read using CDCOM software
Uses unprocessed images (ideally)
Example: 160 μm detail

Good System

Poor System
How were detectability standards set?

• Minimum gold thickness detectable at diameters 0.1 to 2 mm
• Acceptable set so that 97.5% of systems used in UK breast screening would comply
  – NOT VERY DEMANDING
• Achievable values set as averages of data from established digital systems recognised to have ‘good quality’
Dose and noise

- 40 mAs
- 80 mAs
- 160 mAs
- 320 mAs
Dose required to reach minimum image quality standard (60mm thick breast)

<table>
<thead>
<tr>
<th>SYSTEM</th>
<th>mGy (0.1 mm)</th>
<th>mGy (0.25 mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DR 1</td>
<td>0.60</td>
<td>0.67</td>
</tr>
<tr>
<td>DR 2</td>
<td>0.63</td>
<td>0.52</td>
</tr>
<tr>
<td>DR 3</td>
<td>0.85</td>
<td>0.80</td>
</tr>
<tr>
<td>DR 4</td>
<td>1.01</td>
<td>0.87</td>
</tr>
<tr>
<td>Average film-screen</td>
<td>1.17</td>
<td>1.07</td>
</tr>
<tr>
<td>CR 1</td>
<td>1.67</td>
<td>1.45</td>
</tr>
<tr>
<td>CR 2</td>
<td>3.46</td>
<td>1.49</td>
</tr>
</tbody>
</table>
How well are systems in NHSBSP optimised?

Plot dose v image quality from data for 318 systems
Hologic systems in NHSBSP

- Dimensions
- Selenia (Mo)
- Selenia (W)

- Acceptable
- Achievable
- OPTIMAL

Threshold gold thickness (l^m) at 0.1mm vs. MGD (mGy)
Siemens systems in NHSBSP

![Graph showing threshold gold thickness vs. MGD (mGy)]

- Inspiration
- Acceptable
- Achievable
- OPTIMAL

K C Young, NHSBSP 2014
GE systems in NHSBSP

- Senographe Essential
- Senographe 2000D
- Senographe DS

Threshold gold thickness ($\mu$m) at 0.1nm

MGD (mGy)

Acceptable
Achievable
OPTIMAL
Fuji systems in NHSBSP

- Amulet
- Profect CR

Threshold gold thickness ($t_{mg}$) at 0.1mm

- Acceptable
- Achievable

Achievable
Acceptable

MGD (mGy)

K C Young, NHSBSP 2014
Proportion exceeding achievable IQ

<table>
<thead>
<tr>
<th>Digital system</th>
<th>GE 2000D</th>
<th>Fuji CR</th>
<th>Sectra L30</th>
<th>Giotto</th>
<th>Inspiration</th>
<th>Selenia (Mo)</th>
<th>GE DS</th>
<th>Essential</th>
<th>Dimensions</th>
<th>Selenia (W)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0%</td>
<td>0%</td>
<td>5%</td>
<td>25%</td>
<td>28%</td>
<td>50%</td>
<td>59%</td>
<td>59%</td>
<td>82%</td>
<td>83%</td>
</tr>
</tbody>
</table>

K C Young, NHSBSP 2014
Drawbacks of contrast detail measures

- Standards set using unprocessed images
- Phantoms have no anatomical background
- Details may not be clinically realistic
Outline

• Background
• Breast dosimetry and breast dose
• Choice of X-ray spectra
• Image quality assessment
  – Physical image quality measures
  – Quality measures using phantoms
  – Clinical measures of image quality
  – Do phantom measurements and clinical measures correlate?
• Conclusions
Clinical measures of image quality 1

Results from breast screening with different equipment

Differences in cancer detection using CR compared with DR in screening programs

<table>
<thead>
<tr>
<th>Study</th>
<th>DCIS</th>
<th>Invasive cancer</th>
<th>Combined lesions</th>
<th>CR/DR dose</th>
</tr>
</thead>
<tbody>
<tr>
<td>Belgium 2013</td>
<td>-27%</td>
<td>+2%</td>
<td>-3.5%</td>
<td>1.6</td>
</tr>
<tr>
<td>Canada 2013</td>
<td>-50%</td>
<td>-18%</td>
<td>-29%</td>
<td>1.17</td>
</tr>
<tr>
<td>France 2014</td>
<td>-53%</td>
<td>-15%</td>
<td>-23%</td>
<td>NA</td>
</tr>
</tbody>
</table>
Clinical measures of image quality 2

• Difficult to make direct comparison of systems for breast screening by imaging the same breast with each system
  – additional dose
  – low incidence of cancer
  – breast positioning may differ for each system

• Use virtual clinical trials
  – real (or simulated) images
  – each image modified mathematically to simulate different systems
  – can use real lesions or inserted simulated lesions

• Results of 3 virtual clinical trials from OPTIMAM
Our Experimental Approach

- Model the image formation process
- Real and simulated digital images of breasts
- Real and simulated features of cancers

Cancer Detection Observer Studies
Comparing different factors

Use results to optimise design, choice and use of new technology in cancer screening
Adjust image unsharpness using MTF measurements
Measure and understand noise power behaviour with dose and spatial frequency

Electronic noise

Quantum noise

Structure noise

A Mackenzie et al. Medical Physics 2012
Adapt images by adding noise

- Calculate extra noise
  - For dose change
  - For differences between detectors

- Create noise images to add to real image
  - On pixel-by-pixel basis to account for different signal levels

a-Se  
Csl/a-Si  
CR

A Mackenzie et al. Medical Physics 2012
Image quality modification (DR to CR)

Original DR image

Blur DR image to match CR MTF

‘Coloured’ noise to be added

CR like image
OBSERVER STUDY 1

Calcification detection at different image qualities
Observer study on calcification detection

- 81 normal cases
- 81 abnormal cases with 113 simulated subtle clusters
- 6 image quality levels
- 7 observers
- 6804 image readings
Simulated calcifications

Patches are 17.5x17.5 mm square
Each DR image used to create additional images at 6 different IQ levels

- Normal dose + calcs (2.3 mGy) DR
- Half dose (1.15 mGy) DR
- Quarter dose (0.67 mGy) DR
- Normal dose + calcs (2.3 mGy) CR
- Half dose (1.15 mGy) CR
- Processed using Agfa Musica-2
AFROC Curves for IQ levels (fitted to average of 7 observers)

Fraction of normal images with at least one NL

Lesion localisation fraction (LLF)

Normal dose DR
Half dose DR
Quarter dose DR

L Warren et al, Medical Physics 2012
AFROC Curves for IQ levels (fitted to average of 7 observers)

Fraction of normal images with at least one NL

Lesion localisation fraction (LLF)

L Warren et al, Medical Physics 2012
AFROC Curves for IQ levels (fitted to average of 7 observers)

- Normal dose DR (Hologic)
- Normal dose DR (Agfa)
- Half dose DR
- Normal dose CR
- Half dose CR
- Quarter dose DR

Fraction of normal images with at least one NL vs. Lesion localisation fraction (LLF)

L Warren et al, Medical Physics 2012
Lesion sensitivity at 0.1 FP per image

- Normal Dose DR (Hologic): 70%
- Normal Dose DR (Agfa): 72%
- Half dose DR: 47%
- Quarter dose DR: 27%
- Normal dose CR: 42%
- Half dose CR: 30%

L Warren et al, Medical Physics 2012
Does CDMAM test object predict calcification detection?

R² = 0.94
p = 0.04

Threshold gold thickness (µm) - 0.25mm disc diameter

Reader-averaged JAFROC FoM

Better

L Warren et al, Medical Physics 2012
Does CDMAM test object predict calcification detection?

Clear relationship between calcification detection and CDMAM results

Achievable minimum

Reader-averaged JAFROC FoM

Threshold gold thickness (µm) - 0.25mm disc diameter

\[ R^2 = 0.94 \]
\[ p = 0.04 \]
Conclusions from Observer Study 1

1. Threshold gold thickness using the CDMAM phantom is a good predictor of detection of clinical calcification
2. EU guidelines for IQ are clinically relevant
3. Traditional CR systems poorer than DR for calcification detection even at the relatively high dose level used (2.1 mGy, 50-60 mm breast).
4. Good IQ is important for calcification detection and systems should exceed achievable levels in EU guidelines. Acceptable level is too low
5. Image processing investigated had little effect on calcification detection
OBSERVER STUDY 2

Effect of image processing on detection of calcification and non-calcification lesions
Observer study on image processing

80 normal cases
80 cases with simulated subtle clusters
- avoids selection bias
80 cases with malignant non-calcification lesions (subtle or very subtle)
30 cases of biopsy proven benign lesions
3 types of image processing
7 observers

L Warren et al, AJR 2014
Image processing comparison

L Warren et al, AJR 2014
Image processing comparison

L Warren et al, AJR 2014
Image processing comparison

![Graphs showing comparison of standard, low contrast, and simulated screen-film versions for masses and calcifications.](image)

L Warren et al, AJR 2014
**Image processing comparison**

<table>
<thead>
<tr>
<th>Image processing</th>
<th>Figure of merit</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Masses</td>
</tr>
<tr>
<td>Standard</td>
<td>0.73</td>
</tr>
<tr>
<td>Low contrast</td>
<td>0.72</td>
</tr>
<tr>
<td>Simulated screen-film</td>
<td>0.72</td>
</tr>
</tbody>
</table>

L Warren et al, AJR 2014
1. For the detection of calcification clusters the standard image processing was significantly better than the low contrast or simulated screen-film processing

2. For the detection of non-calcification lesions there was no significant difference was found between any of the three image processing pairs

3. There were no significant differences in the number of false positive non-calcification marks for any of the image-processing pairs

4. The number of false positive calcification marks increased by 15% for film-screen processing compared to standard processing
OBSERVER STUDY 3

Effect of detector type on detection of calcification and non-calcification lesions
270 unprocessed image pairs – both breasts, one view (either CC or MLO)

- 80 with no cancer present
- 80 containing malignant subtle non-calcified cancers
- 80 with inserted malignant subtle calcification clusters
- 30 containing biopsy proven benign lesions
Image conversion

• Starting image set was converted to:
  – Arm 1: ‘a-Se’ detector
  – Arm 2: ‘CsI’ detector
  – Arm 3: ‘NIP CR’ detector
  – Arm 4: ‘Powder phosphor CR’ detector

All doses reduced by 20%
AGD = 1.08 mGy for 50 to 60 mm breast thickness

• Image processing
  – Agfa Musica²
  – Suitable for a wide range of image qualities
Calcification cluster detection and recall

Would you recall on the basis of this lesion?

No: Very confident  
No: Moderately confident  
No: Slightly confident  
Yes: Slightly confident  
Yes: Moderately confident  
Yes: Very confident
### Calcification cluster detection and recall

A Mackenzie et al, SPIE Medical Imaging 2014

<table>
<thead>
<tr>
<th>IQ pairs</th>
<th>$p$-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a$-Se v. CsI</td>
<td>0.27</td>
</tr>
<tr>
<td>$a$-Se v. NIP</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>$a$-Se v. CR</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>CsI v. NIP</td>
<td>0.0001</td>
</tr>
<tr>
<td>CsI v. CR</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>NIP v. CR</td>
<td>0.048</td>
</tr>
</tbody>
</table>

![Graph showing lesion localisation fraction vs. non-localisation fraction for different IQ pairs.](image)
Calcification cluster detection and recall

A Mackenzie et al, SPIE Medical Imaging 2014
Calcification cluster detection and recall

Compare with a-Se

<table>
<thead>
<tr>
<th></th>
<th>Change in recall</th>
</tr>
</thead>
<tbody>
<tr>
<td>CsI</td>
<td>n.s.</td>
</tr>
<tr>
<td>NIP</td>
<td>-28%</td>
</tr>
<tr>
<td>CR</td>
<td>-44%</td>
</tr>
</tbody>
</table>

A Mackenzie et al, SPIE Medical Imaging 2014
Non-Calcification detection and recall

<table>
<thead>
<tr>
<th>IQ pairs</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha$-Se v. CsI</td>
<td>0.93</td>
</tr>
<tr>
<td>$\alpha$-Se v. NIP</td>
<td>0.002</td>
</tr>
<tr>
<td>$\alpha$-Se v. CR</td>
<td>0.0007</td>
</tr>
<tr>
<td>CsI v. NIP</td>
<td>0.002</td>
</tr>
<tr>
<td>CsI v. CR</td>
<td>0.0009</td>
</tr>
<tr>
<td>NIP v. CR</td>
<td>0.77</td>
</tr>
</tbody>
</table>

A Mackenzie et al, SPIE Medical Imaging 2014
Non-Calcification detection and recall

A Mackenzie et al, SPIE Medical Imaging 2014
Non-Calcification detection and recall

A Mackenzie et al, SPIE Medical Imaging 2014
Non-Calcification detection and recall

<table>
<thead>
<tr>
<th>Arm 1</th>
<th>Arm 2</th>
<th>Arm 3</th>
<th>Arm 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>'a-Se'</td>
<td>'CsI'</td>
<td>'NIP CR'</td>
<td>'Powder CR'</td>
</tr>
</tbody>
</table>

- **Marked**
- **Recalled**

**Compare with a-Se**

<table>
<thead>
<tr>
<th></th>
<th>Change in recall</th>
</tr>
</thead>
<tbody>
<tr>
<td>CsI</td>
<td>n.s.</td>
</tr>
<tr>
<td>NIP</td>
<td>-6.5%</td>
</tr>
<tr>
<td>CR</td>
<td>-9.1%</td>
</tr>
</tbody>
</table>

A Mackenzie et al, SPIE Medical Imaging 2014
Relationship between physical image quality & cancer detection?
Calcification clusters

Clear relationship between calcification detection and CDMAM results

Achievable
Acceptable

R² = 0.98

Threshold gold threshold thickness (µm)
Non-calcification lesions

Relationship is not so clear

Larger detail
More complex task
Conclusions: 3 OPTIMAM observer studies

1. Threshold gold thickness using the CDMAM phantom is a good predictor of detection of clinical calcification
2. Traditional CR systems poorer than DR for calcification detection even at relatively high dose levels. NIP may be acceptable at high dose levels
3. EU guidelines for IQ are clinically relevant
4. Good IQ is important for calcification detection and to a lesser extent for masses.
Conclusions: 3 OPTIMAM observer studies

1. Systems should exceed achievable levels in EU guidelines.
2. Acceptable level is too low. Guidelines may need revision.
3. Image processing has an effect on calcification detection.
Thank you for your attention