Measuring and Improving Radiologists’ Interpretative Performance on Screening Mammography

Karla Kerlikowske, MD
Diana Buist, PhD
Patricia Carney, PhD
Berta Geller, EdD
Diana Miglioretti, PhD
Robert Rosenberg, MD
Bonnie Yankaskas, PhD
Improving Breast Imaging Quality Standards

- Technical quality of mammography in the U.S. has improved since implementation of the Mammography Quality Standards Act, mammography interpretation remains quite variable.

We plan to study how best to reduce variability and improve interpretive performance among US radiologists.
Breast Cancer Surveillance Consortium

BCSC is an NCI-funded cooperative agreement designed to evaluate practice and performance of screening in community.
Size of the Pooled BCSC Data Resource

- Based on 1996-2004 examinations
  - Total mammograms = 5.4 million
  - Total women = 1.9 million

- Cancer data
  - Invasive cancers = 59,303
  - In situ cancers = 10,858

- Number of radiologists
  - 972
Mammography performance is highly variable across radiologists in BCSC

<table>
<thead>
<tr>
<th>Variability in Screening Mammography Performance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Measure</td>
</tr>
<tr>
<td>-----------------------</td>
</tr>
<tr>
<td>Recall rate</td>
</tr>
<tr>
<td>PPV2</td>
</tr>
<tr>
<td>Sensitivity</td>
</tr>
<tr>
<td>Specificity</td>
</tr>
</tbody>
</table>
EACH DOT REFLECTS ACCURACY OF A SINGLE PHYSICIAN

SENSEITIVITY (%)

FALSE POSITIVE RATE (%)
Recall rate 2-fold higher in U.S. versus U.K.

<table>
<thead>
<tr>
<th>Age</th>
<th>Recall per 100 UK</th>
<th>Recall per 100 BCSC</th>
<th>Recall per 100 CDC</th>
<th>Cancer per 1000 UK</th>
<th>Cancer per 1000 BCSC</th>
<th>Cancer per 1000 CDC</th>
</tr>
</thead>
<tbody>
<tr>
<td>50-54</td>
<td>3.9</td>
<td>8.7</td>
<td>8.0</td>
<td>3.8</td>
<td>2.6</td>
<td>2.8</td>
</tr>
<tr>
<td>55-59</td>
<td>3.6</td>
<td>8.3</td>
<td>7.0</td>
<td>4.9</td>
<td>3.6</td>
<td>3.5</td>
</tr>
<tr>
<td>60-64</td>
<td>3.4</td>
<td>7.9</td>
<td>6.7</td>
<td>5.9</td>
<td>3.9</td>
<td>3.7</td>
</tr>
</tbody>
</table>

Smith-Bindman, JAMA, 2003
<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Association with Performance</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Years of Experience</td>
<td>↓ FP, no Δ TP, ↓ FP, ↓ TP, ↓ FP</td>
<td>Smith-Bindman, 2005 Barlow, 2004 Elmore, 2002</td>
</tr>
<tr>
<td>Volume</td>
<td>↓ FP, no Δ TP, ↑ FP, ↑ TP, ↓ FP, no Δ CDR, ↓ FP, no Δ or ↑ CDR, ↑ PPV, no Δ CDR</td>
<td>Smith-Bindman, 2005 Barlow, 2004 Théberge (Quebec), 2005 Kan (BC), 2000 Coldman (Canada), 2006</td>
</tr>
<tr>
<td>Screening Focus</td>
<td>↓ FP, ↓ TP, no Δ FP or TP</td>
<td>Smith-Bindman, 2005 Barlow, 2004</td>
</tr>
<tr>
<td>Specialists</td>
<td>↓ FP, ↑ TP</td>
<td>Sickles, 2002</td>
</tr>
</tbody>
</table>
Goals of ACS-NCI Project

- Determine the effects of radiologists’ interpretive volume on clinical performance measures.
- Create and evaluate assessment test sets that consist of representative screening mammograms from community practice.
- Develop and pilot test innovative educational programs designed to improve radiologists’ mammography interpretation skills.
## Radiologists in BCSC by Average Annual Interpretive Volume -- 1998-2003

<table>
<thead>
<tr>
<th>Average annual volume</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;480</td>
<td>76</td>
</tr>
<tr>
<td>480-999</td>
<td>66</td>
</tr>
<tr>
<td>1000-1999</td>
<td>114</td>
</tr>
<tr>
<td>2000-2999</td>
<td>33</td>
</tr>
<tr>
<td>3000+</td>
<td>32</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>321</strong></td>
</tr>
</tbody>
</table>

- Survey Radiologists
- FAVOR survey
- Verify volume at non-BCSC facilities
- Verify prevalence of double reads
- Collect physician and facility characteristics
Examine if Performance on Test Set Reflects Performance in Clinical Practice

321 radiologists from GHC, NC, NH, NM, SF, VT

Radiologists detect <30 cancers from 1998-2003
N = 195

Radiologists detect 30+ cancers from 1998-2003
N = 126

High prevalence cancer
High difficulty

High prevalence cancer
Low difficulty

Low prevalence cancer
High difficulty

Low prevalence cancer
High difficulty
Randomized Controlled Trial

321 BCSC Radiologists

Intervention Group I

In person educational intervention

Control Group

Usual CME

Intervention Group II

Self-administered educational intervention

1. Improvement in performance on test set
2. Improvement in actual performance
Summary

- Variability exists in performance measures among radiologists in U.S.
- Understanding factors that explain variability in performance measures is important
- Need a means to assess interpretative skills that is associated with performance in clinical practice
- Need a means to improve interpretative skills
Acknowledgments

- National Cancer Institute
- American Cancer Society
- Women and Radiologists that have contributed information to BCSC
Thank you