
In their recent article, Gur et al. (1) concluded that “The introduction of computer-aided detection. . .was not associated with statistically significant changes in recall and breast cancer detection rates for the entire group of radiologists as well as for the subset of seven radiologists who interpreted high volumes of mammograms.” Three important points must be made regarding this conclusion.

First, the 95% confidence intervals (–11% to 19%) reported by Gur et al. allow for a wide range of detection rate changes. Even a 19% increase in detection rate would be consistent with their findings. The source of this large variation is not clear but may be a result of variability among readers.

Second, we show that the use of computer-aided detection was associated with a substantial, but not statistically significant, increase in cancer detection rates for the subgroup of 17 low-volume radiologists in the Gur et al. study (Table 1). Among the 24 radiologists in their study, the seven high-volume radiologists each read more than 8000 mammograms during this 3-year study and the 17 low-volume radiologists interpreted an average of 1967 mammograms during the same period (table 1 in Gur et al. (1)). The authors analyzed the combined performance of all 24 radiologists and separately analyzed the performance of the seven high-volume radiologists (table 2 in Gur et al. (1)). However, they did not similarly analyze the performance of the 17 low-volume radiologists.

As shown in Table 1, we used the data reported by Gur et al. (1) to address this issue. We subtracted their data for high-volume radiologists from that for all radiologists and found that use of computer-aided detection by the low-volume radiologists was associated with a 19.7% increase in their detection rate (from 3.05 cancers per 1000 mammograms without computer-aided detection to 3.65 cancers per 1000 mammograms with computer-aided detection). Thus, use of computer-aided detection allowed the low-volume radiologists to increase their collective detection rate to a rate equal to that of the high-volume radiologists. However, because of the small sample size, this improvement was not statistically significant ($P = .37$, chi-squared test). A sample size approximately five times larger than the one in the Gur et al. study would be needed to demonstrate that the nearly 20% increase in detection rate was statistically significant. Nevertheless, the increase we calculated using data reported by Gur et al. is similar to the 19.5% increase in cancer detection reported by Freer and Ulissey (2) in 2001 in their prospective study of computer-aided detection.

Our third point relates to the authors’ acknowledgement that during the study period, the percentage of women who were screened for the first time decreased from approximately 40% to 30%. This observation is clinically significant because a population undergoing periodic mammographic screening typically has lower cancer detection rates on second and third (i.e., inci-

Table 1. Mammography recall rates and breast cancer detection rates for high- and low-volume radiologists interpreting screening mammograms without and with computer-aided detection (CAD)

<table>
<thead>
<tr>
<th>Type of interpretation</th>
<th>No. of mammograms read</th>
<th>No. of recalls</th>
<th>No. of breast cancers detected</th>
<th>Recall rate, %</th>
<th>Breast cancer detection rate per 1000 mammograms read</th>
<th>Cancer detection rate, % change</th>
</tr>
</thead>
<tbody>
<tr>
<td>All radiologists (n = 24)</td>
<td>115 571</td>
<td>13 171</td>
<td>407</td>
<td>11.40</td>
<td>3.52</td>
<td>+1.72</td>
</tr>
<tr>
<td>Without CAD</td>
<td>56 432</td>
<td>6430</td>
<td>197</td>
<td>11.39</td>
<td>3.49</td>
<td></td>
</tr>
<tr>
<td>With CAD</td>
<td>59 139</td>
<td>6741</td>
<td>210</td>
<td>11.40</td>
<td>3.55</td>
<td></td>
</tr>
<tr>
<td>High-volume radiologists (n = 7)</td>
<td>82 129</td>
<td>9333</td>
<td>292</td>
<td>11.36</td>
<td>3.56</td>
<td>–3.32</td>
</tr>
<tr>
<td>Without CAD</td>
<td>44 629</td>
<td>5188</td>
<td>161</td>
<td>11.62</td>
<td>3.61</td>
<td></td>
</tr>
<tr>
<td>With CAD</td>
<td>37 500</td>
<td>4145</td>
<td>131</td>
<td>11.05</td>
<td>3.49</td>
<td></td>
</tr>
<tr>
<td>Low-volume radiologists (n = 17)</td>
<td>33 442</td>
<td>3838</td>
<td>115</td>
<td>11.48</td>
<td>3.44</td>
<td>+19.7</td>
</tr>
<tr>
<td>Without CAD</td>
<td>11 803</td>
<td>1242</td>
<td>56</td>
<td>10.52</td>
<td>3.05</td>
<td></td>
</tr>
<tr>
<td>With CAD</td>
<td>21 639</td>
<td>2596</td>
<td>79</td>
<td>12.00</td>
<td>3.65</td>
<td></td>
</tr>
</tbody>
</table>

Table 1. Mammosgraphy recall rates and breast cancer detection rates for high- and low-volume radiologists interpreting screening mammograms without and with computer-aided detection (CAD)
garding the confidence factor, which could obscure any evidence of benefit from computer-aided detection.

Results of previous studies suggest that computer-aided detection should improve breast cancer detection rates and that the value of computer-aided detection may vary among different readers (4,5). The article by Gur et al. and the editorial by Elmore and Carney (6) should have mentioned that the breast cancer detection rates of low-volume radiologists benefited from computer-aided detection. The article by Gur et al. indicates that additional studies will be helpful to provide more insight into the benefits of computer-aided detection in different clinical settings.

Stephen A. Feig
Edward A. Sickles
W. Phil Evans
Michael N. Linver

REFERENCES


NOTES

Affiliations of authors: Department of Radiology, Mount Sinai School of Medicine of New York University, New York, NY (SAF); Department of Radiology, University of California, San Francisco (EAS); Department of Radiology, University of Texas Southwestern Medical Center, Dallas (WPE); X-Ray Associates of New Mexico, Albuquerque (MNL).

Correspondence to: Stephen A. Feig, MD, Department of Radiology, The Mount Sinai Hospital, 1178 Fifth Ave., New York, NY 10029-6574 (e-mail: stephen.feig@mounsinai.org).

S. A. Feig, E. A. Sickles, and W. Phil Evans hold minor stock interests in R2 Technology, Inc., the manufacturer of the ImageChecker computer-aided detection system. S. A. Feig and E. A. Sickles are members of the Mammography Advisory Board of R2 Technology, Inc.

DOI: 10.1093/jnci/djh257

RESPONSE

We thank Feig et al. for their three comments. First, we clearly stated in our article (page 188) that because of the large 95% confidence intervals, “our findings are consistent with the range of improvement in detection rates estimated and reported by others” (1).

Namely, we did not rule out the possibility that a large improvement is indeed possible. Whether it is likely, however, is another matter that requires more data that are appropriately ascertained and, as we indicated in our correspondence (2), we welcome these additional data, whatever they may be.

Second, depending on the assumptions one makes, it is frequently possible to find subsets of participants that produce the desired result. As Feig et al. demonstrate, it is true that there was an increase (albeit not a statistically significant increase) in cancer detection rates among the low-volume radiologists, who interpreted fewer mammograms than the high-volume radiologists. However, during the same time period, the low-volume radiologists showed a statistically significant 14% increase in recall rates, from 10.5% (1242 women recalled per 11 803 mammograms read) to 12% (2596 women recalled per 21 639 mammograms read) (P<.001).

The persistent similar changes in recall and detection rates is the very reason we argued that, for radiologists who have had a certain level of training, recall rates may be highly correlated with detection rates, regardless of whether computer-aided detection is used (3). In other words, one could argue that in our group of low-volume radiologists, the effect of computer-aided detection may have been similar to that obtained if we had simply told the radiologists to be more conservative (or suspicious) in their reading of mammograms. Feig et al. assume that during the period in question, our low-volume radiologists were less trained or less experienced than the high-volume radiologists, which was not necessarily the case.

Third, we discussed the issues surrounding adjustment for confounding factors in this type of observational study in our own correspondence in response to the editorial by Elmore and Carney (2) who, unlike Feig et al., argued that we should have adjusted for recall rates to account for the gradual decrease in the fraction of women undergoing a first screen versus those undergoing a repeat screen. Elmore and Carney used our data to highlight the fact that recall rates (not cancer detection rates) have actually increased with computer-aided detection. Because we expected that the changes in recall rates and cancer detection rates among these two groups (first and repeat screens) would be similar in magnitude (4), we chose not to correct for this confounder.

We believe, in this observational study, this decision was appropriate (2).

David Gur
Jules H. Sumkin
Howard E. Rockette

REFERENCES


NOTES

Affiliations of authors: Department of Radiology, University of Pittsburgh, and Magee-Womens Hospital of the University of Pittsburgh Medical Center, Pittsburgh, PA (DG, JHS); Department of Biostatistics, University of Pittsburgh, Pittsburgh (HER).

Correspondence to: David Gur, ScD, University of Pittsburgh, Department of Radiology, Imaging Research, Ste. 4200 Magee-Womens Hospital, 300 Halket St., Pittsburgh, PA 15213-3180 (e-mail: gurd@upmc.edu).

DOI: 10.1093/jnci/djh258