Computer-Aided Detection of Breast Cancer: Has Promise Outstripped Performance?

Joann G. Elmore, Patricia A. Carney

The article by Gur et al. (1) in this issue of the Journal makes an important contribution to the literature on mammography screening in that it reports no difference in breast cancer detection and recall rates between mammograms read with computer-aided detection and those interpreted by a single radiologist without computer-aided detection. Gur et al. present the recall rates and cancer detection rates for 56,432 screening mammography examinations interpreted before the introduction of computer-aided detection and 59,139 screening mammography examinations interpreted after the introduction of computer-aided detection in their academic radiology practice. No statistically significant differences were observed in those rates. This finding is both remarkable and disappointing, given that computer-aided detection technology has been approved for use by the U.S. Food and Drug Administration (FDA) and is already widely used in clinical practice. Its widespread promotion was almost a promissory note to the public that it would outperform unaided radiologists.

How could the study conducted by Gur et al. result in findings so different from initial evaluations of this technology considered in the FDA approval process, which reported increased breast cancer detection rates with computer-aided detection? As Gur et al. noted, earlier studies of computer-aided detection relied on retrospective interpretations of selected breast cancer cases evaluated in a controlled laboratory testing situation rather than on prospective detection of breast cancer in a setting more similar to actual practice, as they have done. In addition, early studies on computer-aided detection were typically performed using a sample of mammography examinations that had a higher prevalence of breast cancer than that found in ordinary screening practices. Such differences in study design can result in vastly different findings. Indeed, there is reason for concern that a technology intended for a particular purpose (screening) was approved on the basis of results from studies whose designs did not reflect how that technology was intended for use in actual practice.

The evaluation of any new medical technology should be an ongoing process. The history of medical science has taught us that we should put our assumptions to the test in the real world. Unfortunately, evaluation of new technologies for cancer screening purposes is especially challenging for two reasons. First, a large number of individuals must often be screened before even one cancer is detected. Consequently, early studies (including those evaluating computer-aided detection) often assess new technology using laboratory or ‘testing’ situations that are augmented with cancer cases. Second, the most important clinical outcome in studies evaluating cancer screening technologies is overall survival, but this outcome requires many years of follow-up after the screening test is performed to observe possible effects on mortality. Because these data are not yet available for computer-aided detection, this technology has been evaluated using the more intermediary outcomes of cancer detection rates and cancer stage at time of diagnosis.

It is especially intriguing to consider how quickly computer-aided detection has been adopted in the United States. This rapid uptake by the medical community may be due, in part, to financial considerations. For example, the following appeared on the Web site of the manufacturer of one computer-aided detection system (2):

Computer-aided detection (CAD) can contribute to the earlier detection, by an average of over one year, of almost one out of four breast cancers. With the advent of federal and third party reimbursemens for computer-aided detection of breast cancer procedures, offering CAD can make both clinical and economic sense for the women’s health center. Using the example below, for a center with a caseload of fifty patients per day, every dollar invested in a CAD solution could return $7.54 in new revenues. On this basis, the payback period for a CAD solution could be under nine months.

Has the new revenue promised by manufacturers of computer-aided detection fostered the widespread use of such technology before its effectiveness and safety have been adequately tested? This is an important issue facing all new technologies spawned in our market-driven economy. The safeguard, of course, is evidence-based medicine, which purports to guide clinicians to clinically sound and cost-effective decisions based on valid and relevant research. We more closely approach this utopian goal with the data provided by Gur et al. (1) on computer-aided detection in actual screening practice. The standard to which we hold new, expensive technologies such as computer-aided detection should be high. However, questions remain about the efficacy of computer-aided detection, these new data notwithstanding.

Several limitations of the study by Gur et al. (1) merit discussion. First, Gur et al. could not adjust their analysis for characteristics of the women screened in their study, which may have affected their findings. They reported a stable recall rate after the introduction of computer-aided detection, but noted that more women in the later years of the study period returned for subsequent screening examinations. It is generally the case that the recall rate for subsequent examinations (prevalence screens)
is lower than the recall rate for initial examinations (incidence screens). This is likely due to the fact that radiologists often have prior films for comparison on prevalence screens, which helps inform their decisions, whereas abnormalities seen on incidence screens do not benefit from a baseline comparison. If their analysis was adjusted for the presence of women who had prior or comparison films and whether the screening mammogram was the woman’s first mammogram, the recall rate (or percentage of examinations with abnormalities noted) might have actually increased over the study period after the addition of computer-aided detection.

Second, it is not clear how complete the follow-up ascertainment was for breast cancer cases because the authors mention that a small number of women were lost to follow-up and their clinical practice is not linked to the Surveillance, Epidemiology, and End Results (SEER) Program or a health plan cancer registry. Thus, the number of breast cancer cases in their study may have been underestimated to an unknown extent. This lack of standardized and complete follow-up underscores how difficult it is to use what might be perceived as quality-assurance data from the real clinical world in a research study.

Despite these limitations, the study by Gur et al. raises some important questions. First, will computer-aided detection save lives? We cannot answer this question at the present time. Obviously, this is the hope, but it is still an untested hypothesis.

Second, will computer-aided detection affect medical malpractice claims in radiology? This consideration is important because the failure to detect breast cancer continues to be a leading cause of medical malpractice claims in the United States (3). However, we must consider the ways in which computer-aided detection might affect clinical practice and seek hard data on its efficacy in community practice rather than becoming mesmerized with new computer technology. For example, although computer-aided detection might seem like a helpful backup for the human interpreter, it is possible that radiologists might be falsely reassured by a negative computer-aided detection interpretation, or they may rely too heavily on computer-aided detection and risk becoming less vigilant and thorough in their own assessments.

Third, will computer-aided detection reduce the anxiety and uncertainty for the radiologist in the practice of mammography? This lowering of anxiety might improve the quality of the work environment for practicing radiologists, an important consideration. A future shortage of radiologists interpreting mammograms is possible because graduating radiology residents have little interest in entering the field of breast imaging (4).

Fourth, will women prefer to patronize facilities that specifically offer computer-aided detection? Americans love technology, and if women believe that computer-aided detection will prevent them from dying of breast cancer, they will likely seek it out. The U.S. population already overestimates the effectiveness of mammography [e.g., one study suggests that women in their 40s overestimate the relative risk reduction from screening by about sixfold and the absolute risk reduction from screening by more than 100-fold (5)]. Adding a highly technical feature, such as computer-aided detection, to mammography might give some women either an even more inflated estimate of how good mammography is or false reassurance that it is more effective than it actually is.

Finally, will computer-aided detection increase the cost of screening? One mammography facility offers women computer-aided detection for an additional fee of $28 (6). This potential influx of funds from computer-aided detection may make screening mammography less of a financial burden to radiology departments or screening facilities. For example, if we consider the 59,139 examinations performed with computer-aided detection in the Gur et al. study and a conservative reimbursement of $15 per computer-aided detection, the total additional revenue from computer-aided detection would be $887,085 (7). Assuming that the computer-aided detection unit cost $200,000, the practice would have generated more than $650,000 in direct income from computer-aided detection alone, with no discernible improvement in accuracy or improved outcomes for the women screened.

In summary, the results reported by Gur et al. highlight the importance of continued investigation of new screening technologies after they have received FDA approval. The study also shows the challenge of moving from the retrospective, enriched test situations used for FDA approval to cancer screening in actual practice. As more studies on computer-aided detection in the community setting are published and as computer-aided detection systems improve, we hope that women’s lives can ultimately be saved by this new technology.

REFERENCES


(6) Mammography-computer aided detection (CAD): is there an extra charge for CAD? Available at: http://www.washingtonradiology.com/exams/mam/cad.html. [Last accessed: 12/31/03.]


NOTE

Editors’ note: SEER is a set of geographically defined, population-based central cancer registries in the United States, operated by local nonprofit organizations under contract to the National Cancer Institute (NCI). Registry data are submitted electronically without personal identifiers to the NCI on a biannual basis, and the NCI makes the data available to the public for scientific research.