Leveraging the Digital Mammography Image Screening Trial (DMIST) Data for the Evaluation of Computer-Aided Detection (CAD) Devices: A Proposal

Nicholas Petrick, Kyle J. Myers, Sophie Paquerault, Frank W. Samuelson, Brandon D. Gallas, and Robert F. Wagner

U.S. Food and Drug Administration, Center for Devices and Radiological Health, Rockville, MD 20852
{nicholas.petrick, kyle.myers, sophie.paquerault, frank.samuelson, brandon.gallas, robert.wagner}@fda.hhs.gov

Abstract. The availability of the large dataset of screen/film and full-field digital mammograms acquired through the Digital Mammography Imaging Screening Trial (DMIST) presents an extraordinary opportunity for the assessment of CAD devices. The National Cancer Institute and the National Institute of Biomedical Imaging and Bioengineering at the U.S. National Institutes of Health have engaged FDA scientists in the development of a plan to leverage this imaging resource to benchmark the performance of current CAD systems. In this talk, we will present an initial proposal for utilizing the DMIST data to quantitatively assess current CAD systems. It is our goal to engage the IWDM community and other interested groups in the development of a consensus on acceptable study designs for this purpose.

1 Background

A variety of computer-assist devices have been approved by the U.S. Food and Drug Administration (FDA) as aids to a mammographer in the detection of breast cancer. These devices were originally approved for screen/film mammography on the basis of a study design that demonstrated the potential for computer-aided detection (CAD) systems to identify missed cancers without the systems producing a substantial increase in the number of patients recalled for additional procedures [1]. Some of these initial screen/film approvals were later extended for application with select full-field digital mammography (FFDM) systems. The FDA approval studies performed by most of the CAD manufacturers were modest in their scope because of a limited patient population, their focus on screen/film mammography and the retrospective nature of the study design.

1.1 Recent Studies of Breast CAD

A recent prospective study by Gur, *et al.* on the benefits of mammographic CAD in an academic clinical radiology practice concluded that the introduction of CAD in their clinical practice was not associated with statistically significant changes in recall or breast cancer detection rates [2]. Similar results were reported for both the entire

Susan M. Astley et al. (Eds.): IWDM 2006, LNCS 4046, pp. 565 – 568, 2006.

group of 24 radiologists who participated in the study and the subset of radiologists who interpreted the highest volume of mammograms. While the conclusions of this study are limited to patients in the University of Pittsburgh system, screen/film mammography and one mammographic CAD device, they strongly suggest that additional studies to benchmark and evaluate the adjunctive benefit of CAD would be of practical value to the public at large.

Even fewer studies, and none with a large diverse patient population, have been conducted to benchmark and evaluate the adjunctive benefit of mammographic CAD with FFDM. The Blue Cross Blue Shield Technical Evaluation Center (TEC) recently conducted a MEDLINE literature search addressing the question as to whether the use of CAD can improve the sensitivity and specificity of FFDM [3]. A prior assessment of CAD in screen/film mammography conducted by the TEC concluded that evidence is available to support the conclusion that CAD improves the accuracy of screen/film mammography by increasing the true-positive rate without a disproportionate increase in the recalls compared with single-reader radiologist interpretation. For CAD as an adjunct to FFDM, the search yielded no high-quality articles in peer-reviewed journals assessing this combination. Therefore, the TEC concluded that "until results from better studies focusing on the use of CAD with FFDM become available, the benefits of CAD with FFDM cannot be determined." This conclusion again supports the need for additional studies to benchmark and evaluate the adjunctive benefit of CAD when combined with FFDM.

1.2 Digital Mammography Imaging Screening Trial

The American College of Radiology Imaging Network (ACRIN), under the direction of Etta Pisano, M.D., conducted the Digital Mammography Imaging Screening Trial (DMIST) [4, 5]. Funding for the trial was provided by the NIH National Cancer Institute and the total cost was on the order of $30 million. The primary goal of this large population-based trial was to compare the diagnostic accuracy of digital and screen/film mammography in a breast cancer screening population of asymptomatic women [5]. The trial was designed to measure small but potentially clinically important differences in diagnostic accuracy between digital and screen/film mammography in the overall population of asymptomatic women and in particular subgroups of denser breasted women where digital mammography might be expected to have an improved diagnostic ability [4].

The DMIST trial collected both digital and screen/film mammograms, in random order, from 49,528 women at 33 sites in the United States and Canada. Five digital mammography systems were utilized in the trial. These systems included the SenoScan (Fischer Imaging), the Computed Radiography for Mammography (Fuji), the Senographe 2000D (General Electric), the Lorad/Trex Digital Mammography System (Hologic) and the Selenia Full Field Digital Mammography System (Hologic) [5]. The screen/film and digital mammograms were each read independently by different radiologists with each reader rating patients using both a seven-point malignancy scale and a Breast Imaging Reporting and Data System (BI-RADS) [6] classification. All relevant information was available for 42,760 of these women, including 335 women subsequently identified as having breast cancer. Breast cancer