Mass detection on mammograms: how do humans and models deal with signal uncertainty ?

C. Castella[1,2], K. Kinkel[3], M. P. Eckstein[4], C. K. Abbey[4], F. R. Verdun[1], R. S. Saunders[5], E. Samei[5], and F. O. Bochud[1]

[1] University Institute for Radiation Physics, DUMSC, CHUV and University of Lausanne, CH-1007 Lausanne, Switzerland;

[2] LPHE, Ecole Polytechnique Fédérale de Lausanne, CH-1015 Lausanne, Switzerland;

[3] Clinique des Grangettes, CH-1224 Chêne-Bougeries, Switzerland;

[4] Dept. of Psychology, University of California, Santa Barbara, CA, USA 93106-9660;

[5] Duke Advanced Imaging Laboratories, Duke University, Durham, NC, USA 27705

mail: cyril.castella@chuv.ch

Introduction

Mass detection in mammography is a task that is made very complex by the wide spectrum of possible signal locations, sizes, and shapes. These uncertainty sources are usually hard to reproduce and to analyze separately in clinical studies or with computer-generated images. In this study, we used hybrid images consisting of synthetic yet realistic masses embedded in mammographic backgrounds and computer generated clustered lumpy backgrounds (CLB) to investigate the influence of signal variability on human and model observer performances.

Material and Methods

We used realistic masses and backgrounds that have been validated by radiologists during previous studies, ensuring conditions close to the clinical situation. Four trained non-physician observers participated in two-alternative forced-choice (2-AFC) experiments. They were asked to detect synthetic masses superimposed on real mammographic backgrounds or CLB. Separate experiments were conducted with sets of benign and malignant masses.

Results under the signal-known-exactly (SKE) paradigm were compared with signal-known-statistically (SKS) experiments. In the latter case, the signal was chosen randomly for each of the 1,400 2-AFC trials among a set of 50 masses with similar dimensions, and the observers did not know which signal was present. Human observers' results were then compared with model observers in the same experimental conditions.

Results

The performance of the human observers does not differ significantly when benign masses are superimposed on real images or on CLB with locally matched gray level mean and standard deviation. For both benign and malignant masses, the performance does not differ significantly between SKE and SKS experiments, when the signals' dimensions do not vary throughout the experiment. However, there is a performance drop when the SKS signals' dimensions vary from 5.5 to 9.5 mm in the same experiment. Noise level in some model observers can be adjusted to reproduce human observers' proportion of correct answers in the 2-AFC task within 5% accuracy for most conditions.

Discussion and conclusion

By conducting detection experiments with realistic masses superimposed on real mammographic and synthetic backgrounds, we showed that human observers do not need to know (or do not use) the precise shape of the signals to be detected, as long as they have access to information about signal size. The excellent agreement obtained with some model observers provides a useful and objective tool for testing other conditions (mass type, position, or size) without involving time-consuming human observers studies.