Modeling the Effect of Computer-Aided Detection on the Sensitivity of Screening Mammography

Robert M. Nishikawa

Department of Radiology and Committee on Medical Physics, University of Chicago
5841 S Maryland Ave., MC-2026. Chicago, IL 60637, USA
r-nishikawa@uchicago.edu

Abstract. We have developed a Monte Carlo model to examine the cancer detection rate in screening mammography. We simulated the situation where screening was implemented for 9 years and then CADe was implemented for an additional 9 years. We investigated the effectiveness of two different methods for measuring changes in cancer detection rate. The first method was a sequential method in which the radiologist first reads without CADe and then immediately reads with CADe. The second method is temporal comparison where the cancer detection rates for two periods of time are compared: one without the use of CADe and one when CADe is in use. The model predictions have important implications for clinical studies of CADe. The temporal method is unlikely to measure a real affect, because the effect is small. A sequential method can measure an increase in the number of cancers detected because of CADe, but it cannot measure an overall increase in the cancer detection rate of the screening program.

1 Introduction

Computer-aided detection (CADe) has been proposed as a method for reducing the number of missed cancers. There have been six clinical studies of CADe published to date. The first by Freer and Ulisseys showed a 19.5% increase in the number of cancers detected with an increase in the recall rate from 6.5% to 7.7% when CADe was used [2]. Gur *et al.* reported that when CADe was used, the cancer detection rate increased from 3.49 to 3.55 with virtually no change in the recall rate [3]. Feig *et al.* performed a subanalysis of the Gur study and found that the low volume readers had a 19.7% increase in cancer detection rate, while the high volume readers had a 3.2% *decrease* [4]. Birdwell *et al.* measured a 7% increase in cancers detected due to CADe with 8% increase in recall rate [5]. Cupples *et al.* found a 16% increase in cancer detection rate with an 11% increase in recall rate [1]. Helvie *et al.* found a 10% increase in both number of cancers detected and recall rate, although it was a relatively small study [6]. Khoo *et al.* found a 1.7% increase in the number of cancers detected with a 6% increase in the recall rate [7]. This study was done in the context of double reading. None of the differences in any of the studies reach statistical significance.

On the surface, these studies seem to be contradictory, however, different methods were used to measure the effectiveness of CADe. The Freer, Helvie, Birdwell, and

Susan M. Astley et al. (Eds.): IWDM 2006, LNCS 4046, pp. 46 – 53, 2006.

Khoo studies used a sequential method. In this method, the radiologist first reads the mammograms without any computer assistance and he or she renders an interpretation. Immediately after, the radiologist reviews the computer analysis of the mammograms and renders another interpretation. By comparing the number of cancers detected in each of the two reading conditions, the impact of CADe was measured. The Gur and the Cupples studies used a temporal method based on historical comparisons. In this method, clinical data is collected retrospectively from two time periods. The first time period is from mammograms read without using CADe and the second time period is from mammograms read using CADe. A comparison of the cancer detection rates in the two time periods is a measure of the effectiveness of CADe.

The goal of the present study is to use a computer model of CADe in screening mammography to understand how these two methods can lead to different conclusions. We will show that the results of the clinical studies are not unexpected.

2 Method

An outline of the model is given in Figure 1. The model was implemented in Excel (Microsoft Corporation, Redmond, WA). Each decision outlined in the flowchart was implemented using a random number and comparing it to the probability of an event

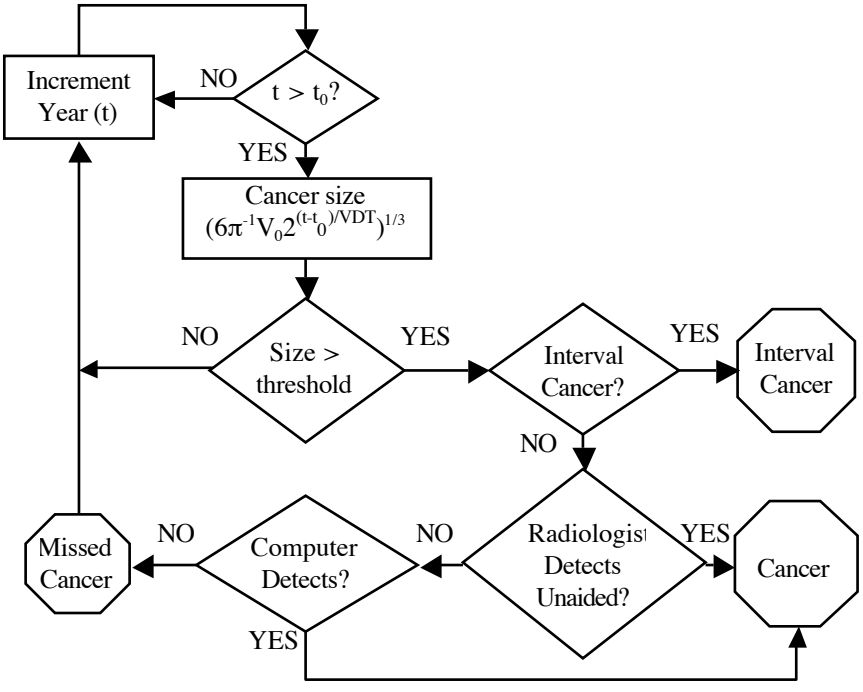

Fig. 1. Schematic of model. t_0 is the time the cancer starts to grow. This process is performed once for each cancer in the simulation.