

Developments In Business Simulation & Experiential Exercises, Volume 21, 1994

SIMULATION OF THE PREDICTIVE VALUE OF MAMMOGRAPHY

Susan A. Chesteen, University of Utah
Collin J. Watson, University of Utah

ABSTRACT

This article reviews data on incidence and mortality rates from breast cancer and provides the results from applications of Bayes' theorem, sensitivity and simulation analyses of the effects of changing numerical values in a probability model to analyze the predictive values of mammograms.

The results of our analyses indicate that the accuracy and effectiveness of mammographic screening for breast cancer are debatable. Consequently, we identify crucial unanswered questions for health systems about the heavy reliance on mammography. We recount the current recommendations, standards of practice, and utilization of mammography for breast cancer screening.

INTRODUCTION

Simulation is a useful method with substantial potential for business research and decision making. Although some types of simulations have been used for many years, the heightened interest in this method began with the arrival of the computer. In this paper we provide a brief discussion of simulation and an example of its use in health care management and consumer behavior.

Simulation has diverse definitions, various types of classifications, and many purposes and facets (Emshoff & Sisson, 1970; Teach, 1990). "In its most general form simulation may be described as the process of conducting experiments on a model of a system. As used in this context a 'model' is anything used to represent the system" (Emory, 1986). For example, Bayes' theorem is the symbolic or mathematics based model through which probabilities are simulated for random variables in a decision-making system for cancer screening. Computer simulation is particularly useful in this situation because of the presence of random variables, the relationships among the variables and the numerous time-consuming calculations.

The two essential elements of a simulation are a model of the process or system being studied and a sample of inputs which may be actual data or synthetic data created to reflect the general traits of real input data. When the model is operated a set of outputs is generated based on the sample of inputs (Emory, 1980). For example, in our simulation the inputs are sets of values for the probability of cancer, the probability of a positive test given cancer and the probability of a negative test given no cancer. The outputs are sets of probabilities for the predictive value of a positive test and for the predictive value of a negative test.

Using a computer-based model with simulation is especially useful in the circumstance in which it is not feasible to conduct an experiment or when there is no optimal solution. It is neither practical nor cost beneficial to conduct numerous experiments using randomized trials to estimate the probability distribution of the predictive value of mammographic tests. Furthermore, since the incidence rate of cancer, the probability of a positive mammogram given no cancer and the probability of a negative mammogram given cancer change over time and across populations we conduct a simulation analysis with randomly generated probabilities (Harris et al., 1992).

Simulations typically involve an application of the Monte Carlo technique to incorporate random variables into models. The value for a random variable is determined by a chance process described in the form of a

probability distribution (Emory, 1980). Whereas, parametric sensitivity analysis examines the effects of changing numerical values in a probability model (Samson, 1988, Irwin). In this study, we examine the stochastic properties of the predictive value of mammographic tests by using both sensitivity and simulation analyses.

Cancer is a primary cause of death of women ages 35 to 50 in the United States. Significantly, breast cancer accounts for the most frequent malignancy. In 1961 a woman had a five percent chance of developing breast cancer during her lifetime (Laurence, 1991); today it is estimated that one of every nine woman will develop breast cancer by age 85 (American Cancer Society, 1993). An estimated 182,000 new cases among women have been projected for 1993; in this same year about 46,300 will die of breast cancer. Even with the medical community's improved screening techniques, treatment has only slightly improved the chances of surviving (Marshall, 1991). As the incidence of breast cancer continues to rise at an alarming rate, health care consumers are asking hard questions such as: Just how effective is mammography as a screening procedure? Do health care consumers and providers truly understand the probabilities associated with the accuracy of mammography for screening for breast cancer? Why are there so few efficient methods, if any, for the detection, prevention and treatment for breast cancer? Why isn't more federal funding being allocated to support basic research of breast cancer? And the consumers are demanding satisfactory answers—not excuses. Without reasonable answers, how can the benefits of mammographic screening procedures be weighed against the drawbacks in any rational manner?

This article reviews data on incidence and mortality rates from breast cancer and provides the results from applications of Bayes' theorem and a simulation analysis of the effects of changing numerical values in a probability model to analyze the predictive value of a positive mammogram.

To reduce the morbidity and mortality due to breast cancer, many physicians propose that asymptomatic women without known risk factors be screened to detect cancer at a curable stage. The rationale for screening is based on the widely held view that screening is effective for detection and removal of breast cancer in its early stages as well as for abnormal growths suspected to be precancerous. However, the probabilities and statistics associated with breast cancer and the predictive value of mammography as a screening test vary considerably.

In this article we first describe the range of current recommendations and considerations for breast cancer screening of women with or without known risk factors. Then we briefly discuss the costs and utilization of mammography. Next, through the application of Bayes' Theorem and a test of homogeneity, we demonstrate reason for legitimate concerns regarding the heavy reliance on mammography. Our results indicate that the accuracy and effectiveness of the mammographic screening test for breast cancer are questionable. Our conclusions highlight professional and public sector issues related to management of health care systems and health policy towards women. In the last section of the paper we identify some of the important unanswered questions about mammographic screening (Watson & Chesteen, in press).

RECOMMENDATIONS AND CONSIDERATIONS IN SCREENING

Most physicians claim that breast screening including monthly self-

Developments In Business Simulation & Experiential Exercises, Volume 21, 1994

examination, annual professional breast examination, and appropriately timed use of mammography, is the most successful way to detect breast cancer at its earliest, most treatable stages. They emphasize that early detection through regular and thorough breast examination is still the single most important factor in determining a favorable outcome for breast disease. The American Cancer Society (ACS) of sets the following guidelines for women:

- 1 - Start monthly basic self examinations at age 20.
2. Have medical physical examination every three years between age 20 and 40, every year after age 40.
3. Have a baseline mammogram between age 35 and 40.
4. Have a mammogram every one to two years between age 40 and 49.
5. Have an annual mammogram after age 50.
6. High-risk women should consult their personal physician.

When making a decision about screening, one should consider at least three issues: (1) a woman's absolute risk of cancer without screening; (2) the extent to which screening can reduce the risk of cancer mortality; and (3) the costs of screening in terms of effort, morbidity, and dollars. Since the yearly incidence of breast cancer increases dramatically from the age of 50 to the age of 70, age is considered to be an important risk factor. Although age could be used to focus screening efforts on older women at high risk, the benefits of screening would become diluted at older ages by other causes of mortality and morbidity. A strong family history of breast cancer also indicates increased risk. Many breast cancers may some day be genetically determined by molecular markers. Other risk factors related to hormonal balances include early menarche, late childbearing, late menopause, and obesity or a high-fat diet. Another factor linked with breast cancer is exposure to radiation. A research study in Boston found that nearly eight percent of all breast cancer in the United States could result from the presence of a gene that occurs in more than a million American women and that radiation can trigger the gene to cause cancer. The study concluded that moderately strong X-rays significantly raise the risk of breast cancer in those women who carry the gene (Haney, 1992).

Knowledge about a woman's absolute risk of cancer could be used to determine the ideal intensity of screening. Even if efficacy is 100 percent, an individual with zero risk of cancer can gain nothing from screening. Thus, if genetic markers, for example, could be used to identify a large segment of the population with little or no risk of breast cancer, the decision not to screen would be an obvious one. At the other extreme, for high risk women a form of screening whose efficacy is supported only by indirect evidence may be appropriate.

Some argue that the costs of mammographic screening and extra interventional procedures are trivial when considering the benefits of early detection of cancer; however, examination of the financial costs associated with mammography for breast cancer screening for asymptomatic women is informative. The expenditures in Table 1 are impressively high and do not include costs for follow-up procedures. The direct cost of ACS recommendations for screening the approximately thirty million American women over the age of 35 who are now in the civilian labor force can be roughly estimated to be almost 2 billion dollars (\$1,876,814,400). Additional costs would be incurred by further diagnostic procedures, such as surgical biopsies and stereotactic automated large-core biopsies. Each year about 500,000 women in the United States of America undergo conventional surgical biopsies. The typical cost of the stereotactic needle biopsy is about \$750 to \$1000; this is about one-third to one-half the cost of a surgical biopsy (Bishop, 1992). Yet, where non-palpable lesions are usually first detected by mammography during routine checkups, an extremely large percentage of the lumps are benign.

Notwithstanding the anxiety arising from a suspicious mammogram and the subsequent pain of a biopsy, the aggregate cost of a policy decision to screen is huge, whereas the benefit is as yet uncertain. Ultimately, the efficacy of screening must be measured in terms of the resulting reductions in morbidity or mortality due to cancer. The screening program of Health Insurance Plan of New York (HIP) claims to have provided evidence that screening affects both survival and mortality (Moskowitz, 1983).

Whether a screening test can reduce mortality depends in part on the test's ability to detect the appropriate target lesion, such as an early curable cancer. The ability of a mammogram to detect a malignant tumor depends not only on the quality of the machine but also on the skill of the technician and the competence of the radiologist who reads the mammogram. Unfortunately, the results reported from mammographic screening tests are not always correct. In fact, not all radiologists will reach the same conclusion from the same mammogram, partially because "mammography is a very delicate interplay of technique and of the observer's behavior" (Moskowitz, 1983). Some forms of benign abnormalities, such as cysts, fibroadenomas and so forth, may be similar to cancer in appearance on film. Consequently, mammography can result in a positive test suggesting the possible presence of cancerous or abnormal cells. When the interpretations of mammographic screenings indicate the need for follow-up procedures, we call these positive mammograms. This scenario would likely lead to a recommendation for further diagnostic procedures when actually there is no cancer present.

Almost two thirds of women aged forty years or older have had at least one mammogram, yet only approximately one third of women in the same age group are following the screening guidelines for mammography. Women 50-59 years of age are the most prevalent users of mammography with frequency of use decreasing for younger women. Of the women having had a mammogram, almost three-fourths reported they did it because their doctors recommended it (Center for Disease Control, 1990).

Screening mammograms are now done by 96 percent of physicians, yet only 37 percent of physicians are following ACS guidelines with 72 percent stating that they agree with the guidelines. Many physicians who disagree with the guidelines do so because of the expense of following them. Many disagree with the recommended frequency for women after age 50, and many also think that mammography as a baseline screening procedure for asymptomatic women age 35-40 is too early and therefore not cost beneficial (American Cancer Society, 1989). The frequency of recommendations for mammographic screening may also be influenced by the fact that failure to diagnose breast cancer is the most expensive and the second most frequent reason for claims brought against physicians (Donegan, 1992). Furthermore, physicians' increasing use of self-referral imaging centers has led many critics to question whether physicians' advice to mammographically screen a substantial number of patients is perhaps based more on the entrepreneurial spirit than on the patients' best interests (Mason, 1992).

The determination of the efficacy of mammographic screening in reducing the morbidity and mortality of breast cancer is an extremely difficult problem. The rationale for mammographic screening is compelling: that detection and treatment of breast cancer in its early asymptomatic stages will increase survival and thereby decrease mortality from the disease. Yet, many providers within health care systems agree that there is a lack of a quick and convenient means for accurate diagnosis (Donagan, 1992). Still, breast cancer mortality in the U.S. has been growing despite increasing utilization of mammography. In the next sections of this paper, we use Bayes' Theorem to examine the predictive ability of mammographic breast cancer screening. We think it urgent that the probabilities and risks be understood by the public, policy makers, and health care practitioners.

Developments In Business Simulation & Experiential Exercises, Volume 21, 1994

BAYES' THEOREM APPLIED TO MAMMOGRAPHY

The annual incidence rate of breast cancer is somewhat rare and has been reported to be about 2 in 1000 women (Moskowitz, 1983). Thus, for this example, we take the probability that a patient selected randomly for screening has this type of cancer to be $P(\text{Cancer}) = .002$. However, the probabilities of acquiring cancer are not known exactly and they may change over time or across populations. The probability of a positive mammogram given that there is no cancer has been estimated (Moskowitz, 1983) to be about $P(\text{Positive test} | \text{No cancer}) = .04$. Furthermore, some cancers are not discovered by mammography due to poor quality in the radiographic image, questionable skills of the mammogram reader, minute size or subtle details of the lesions and so forth. Consequently, mammography can result in a negative test with no further diagnostic procedures recommended even though there actually is cancer present. The probability of a negative test, or mammogram that does not lead to further diagnosis or treatment, given that the individual has cancer has been estimated (Goldberg and Wines, 1981) to be about $P(\text{Negative test} | \text{Cancer}) = .36$. The probabilities can be summarized as shown in the probability tree diagram presented in Figure 1.

To statistically evaluate mammographic screening for cancer, one probability we need to know is the probability of having cancer given a positive test or mammogram (the predictive value of a positive test or mammogram). Using the information presented in Figure 1, the probability that a randomly selected patient has cancer given that the individual received a Positive test from mammography is determined by using Bayes' theorem to be .031.

Thus, the probability of cancer given a positive mammogram is .031. This probability (.031) is greater than the annual incidence rate of breast cancer (.002); nevertheless, it still seems somewhat low. Based on this probability, of the individuals who receive a positive mammogram, substantially more will not have the disease (about 97 out of 100) than will have the disease (about 3 in 100). The small magnitude of this probability is not simply due to the inaccuracy of mammography but also results because of the low annual incidence rate. Low predictive values of positive tests also occur with screening for other phenomena that have low incidence rates (Watson, 1993).

Next, we determine the probability of not having cancer given a negative test or mammogram (the predictive value of a negative test). Using the probability information presented in Figure 1, the probability that a randomly selected patient does not have cancer given that the individual received a negative test or mammogram is determined by using Bayes' theorem to be .99924.

Consequently, the probability that the individual does not have cancer given a negative mammogram increases just slightly from .998 to .99924. Thus, the increase in the probability that an individual does not have cancer is small. This also means that the probability of an individual actually having cancer given a negative mammogram is only .00076 (or $1 - .99924$). The hazard to these very few individuals is that they will not be diagnosed as actually having cancer as soon as they would be with a completely accurate screening test and the delay in treatment may decrease their chances of survival.

Since the incidence rate of cancer, the probability of a positive mammogram given no cancer and the probability of a negative test given cancer are estimated and they change over time and across populations, we conducted a sensitivity analysis using ranges of probability values to explore the effect on the predictive value of a negative mammogram. The results of the sensitivity analysis are reported in Table 4 and summarized in Figure 4. The highest Since the incidence rate of cancer, the probability of a positive

mammogram given no cancer and the probability of a negative test given cancer are estimated and they change over time and across populations, we also conducted a simulation analysis using ranges of probability values to explore the effect on the predictive value of a negative mammogram. The results of the simulation analysis are reported in Table 5 and Figure 5. The mean and standard deviation of the predictive value of a negative test are .99947 and .00035409. In this case, the histogram demonstrates in a visual fashion that the predictive value of a mammographic test for breast cancer is high.

Overall, the overwhelming majority of patients who receive positive mammograms actually do not have cancer. Statistical evaluation of mammographic screening for cancer suggests that it would be in the best interests of health care consumers if the accuracy of the mammographic technique could be improved or if other more accurate cancer screening methods could be developed.

DISCUSSION

Consumers have been shocked at the alarming statistics being disseminated about the frequency of breast cancer diagnoses. However, the annual incidence rate of this type of cancer, the accuracy of mammographic screening techniques, and the statistical evaluation of the effectiveness of mammography do not seem to be well known by either health care consumers or health care providers.

We have stated some legitimate reservations about the accuracy of mammography for screening for breast cancer. Our results underscore the need for all to be aware of the predictive value of mammography as a screening technique. Health care providers and consumers cannot afford to ignore the critical questions surrounding the issue of screening effectiveness.

According to the figures derived from our simulation and sensitivity analyses, screening improves one's chances of not having breast cancer at the time of screening from about 99.8 percent to a bit more than 99.9 percent given a negative mammogram. In a marketplace with numerous expensive medical procedures, there are difficult issues confronting consumers and corporations that require close scrutiny.

The results of our analysis of the predictive values, the costs, and the benefits of mammographic breast cancer screening have significant economic implications for corporate and public policy. Many corporations are embarking on wellness campaigns, which promote a proactive and prophylactic approach for employee health. At least two-thirds of American firms with 50 or more employees currently offer some type of screening or health awareness program for breast cancer. For example, Liz Claiborne has an extensive program involving education, free mammography screenings and physician counseling in addition to a mobile unit to the New Jersey and Pennsylvania sites. At the New York Claiborne conference rooms the Memorial Sloan-Kettering Cancer Center conducts breast cancer screenings (Fulman, 1991).

When weighing the evidence concerning the effectiveness and use of mammographic screening, one should recognize that mammography is limited by numerous factors such as the age and risk profile of the patient, the skill of the health care providers performing and interpreting the results of the tests, the condition of the equipment, and the inconsistency among subsequent diagnostic protocols for suspicious mammograms. Furthermore, the notion of weighing the value of screening in terms of costs per year of life saved is one that is frequently rejected for emotional reasons. Finally, it is widely held that only medical, not economic, considerations should provide the

Developments In Business Simulation & Experiential Exercises, Volume 21, 1994

decision to screen asymptomatic women for breast cancer.

But when primary care physicians influence some asymptomatic women to undergo mammographic screening, they are asking these patients to accept inconvenience, expense and the risk of side effects to gain benefits that may be elusive. Table 1 demonstrates that there are substantial costs associated with mammography for individual consumers, insurance providers, and businesses. Value judgments must be made about whether the benefits are worth the costs. Concerns similar to ours were recently raised by sixty-plus scientists in a conference in Washington where demands were made for a total overhaul of federal cancer policies. These concerns were further emphasized by a recent Canadian study that reported a 50 percent increased breast cancer mortality in woman over 40 given annual mammograms versus those given physical examinations only (Epstein, 1992).

We have provided information on the predictive value of mammography in tables and figures that may be useful for individual women and their health care providers for making a rational decision about screening. We believe that responsible clinical decision making should reflect all possibilities and limitations of mammography and that the outcomes, options, and probabilities associated with mammographic screening for breast cancer should be well understood by patients and their physicians. Only then is it possible for physicians to exercise their role as effective advisors for patients in their decision making. A well-informed and joint decision-making approach which also incorporates patients preferences encourages the rational utilization of screening and other diagnostic procedures and promotes the avoidance of litigation (Wennberg, 1991). A copy of the complete paper is available on request.

REFERENCES

- American Cancer Society 1989 Survey of physicians' attitudes and practices in early cancer detection CA 1990; 40:77-101
- American Cancer Society Cancer Facts & Figures 1993: 11 Bishop J. New biopsy technique for breast lumps is called quicker, cheaper, less painful. The Wall Street Journal. Feb. 5 1 1992; B 1.
- Brett A. Treating hypercholesterolemia: How should practicing physicians interpret the published data for patients? New England Journal of Medicine 19B9; 321: 676-79
- Centers for Disease Control. Use of mammography. Morbidity and Mortality Weekly Report 1990; 39: 621-30
- Cohen E. Keeping abreast of technology Lear's June 1992; 36.
- Donegan W Evaluation of a Palpable Breast Mass New England Journal of Medicine 1992; 327: 937-942
- Dodd G. Introduction to "Evaluation of periodic breast cancer screening with mammography: Methodology and early observations" by Shapiro, Strz, and Venet". CA-A Cancer Journal for Clinicians 1990; 40:109-10
- Emory, C.W. Business Research Methods Richard D. Irwin, Inc. Homewood, Illinois, 1980.
- Emshoff, J.R., and Sisson, R.L. Design and Use of Computer Simulation Models The Macmillan Company: London, 1970.
- Epstein S. The push for mammography may be off target. USA Today Feb. 4, 1991; 9A
- Fulman R. Leader in corporate fight against breast cancer The Salt Lake Tribune Oct. 21, 1991; A6
- Goldberg J, Wittes J. The evaluation of medical screening procedures The American Statistician 1981; 35:4-11.
- Goldstein M. Women emulate AIDS victims, mobilize against breast cancer. The Salt Lake Tribune Oct. 14, 1991; A9
- Haney O. Study: X-rays increase risk of breast cancer in some. The Salt Lake Tribune 1991; Dec. 26, A8
- Harris, J.R., Lippman, M.E., Veronesi, U., and Willett, W. Breast cancer New England Journal of Medicine: 327 (5); July, 1992, pp. 319-328.

- Kerlinger F. Foundations of Behavioral Research (2nd Ed.). Holt, Rinehart and Winston, Inc., New York, 1964, pp. 153-156.
- Laurence L The breast cancer epidemic: Women aren't just scared, were mad. McCall's Nov. 1991; 24-40.
- Lilford R, Johnson N. The alpha and beta errors in randomized trials New England Journal of Medicine 1990; 322:780-781
- Marshall E. Breast cancer: Stalemate in the war on cancer. Science 1991; 254: 171 9-1 720.
- Mason M. A little clinic on the side: Should doctors profit from their referrals? Newsweek. March 30, 1992; 71.
- McNiel B, Weichselbaum A. Pauker S. Fallacy of the five-year survival in lung cancer New England Journal of Medicine 1978; 299:1 397-1401.
- Moskowitz M. How can we decrease breast cancer mortality? CA-A Cancer Journal for Clinicians 1980; 30: 273-277
- Moskowitz M. Minimal breast cancer redux. Radiologic Clinics of North America 1983; 21: 93-113
- Reitman, J. Computer Simulation Applications Wiley-Interscience: NewYork, 1971.
- Samson d. 1988. Managerial Decision Analysis Irwin, Homewood, IL
- Teach R.D. Designing business simulations. Guide to Business Gaming and Experiential Learning. (J.W. Gentry, Editor) Nichols/GP Publishing: East Brunswick/Logan Page, London, 1990.
- U.S. Dept. of Labor, Bureau of Statistics Employment and earnings Dec. 1991; 15
- Watson C, Billingsley P. Huntsberger D, Croft D. Statistics for Management and Economics (5th Ed.). Allyn & Bacon, Inc., Newton MA, pp. 163-1 64.
- Watson, C. and S. Chesteen Statistical Evaluation of the Effectiveness of Mammography Journal of Health Systems In press
- Wennberg J. Outcomes research, cost containment, and the fear of health care rationing. New England Journal of Medicine 1991; 325: 66-67