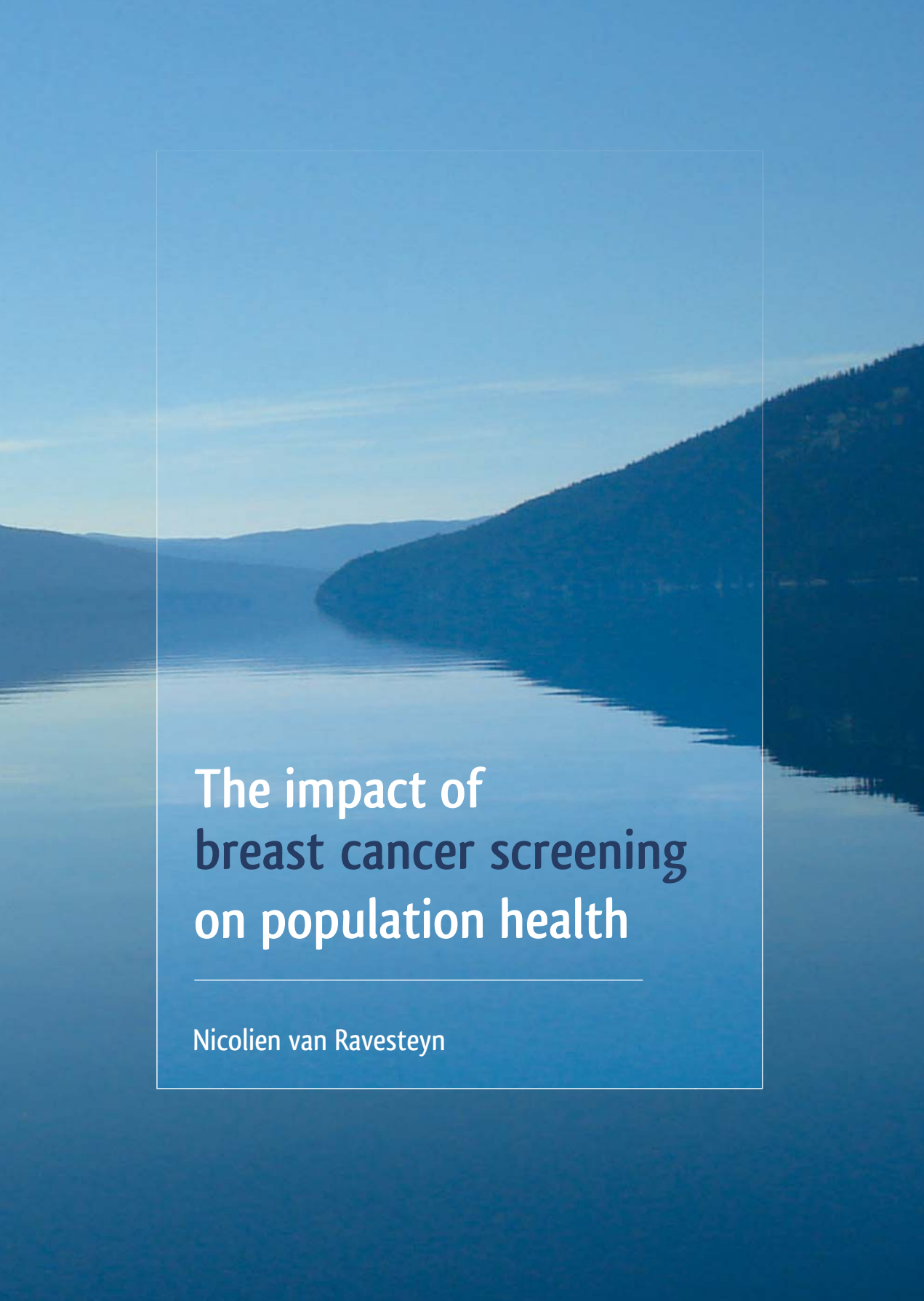


The impact of breast cancer screening on population health

Nicolien van Ravesteyn

THE IMPACT OF BREAST CANCER SCREENING ON POPULATION HEALTH

Nicolien van Ravesteyn

ISBN: 978-94-6169-367-9

Cover photo: Otto de Smeth
Layout and printing: Optima Grafische Communicatie, Rotterdam, The Netherlands

The studies reported in this thesis were primarily funded by the National Cancer Institute. For any grant pertaining to specific studies the reader is referred to the individual papers published in their respective journals.

This thesis was printed with financial support of the Department of Public Health Erasmus MC Rotterdam and the Erasmus University Rotterdam.

The Impact of Breast Cancer Screening on Population Health

De impact van borstkankerscreening op de volksgezondheid

Proefschrift

ter verkrijging van de graad van doctor aan de
Erasmus Universiteit Rotterdam
op gezag van de
rector magnificus

Prof.dr. H.G. Schmidt

en volgens besluit van het College voor Promoties.
De openbare verdediging zal plaatsvinden op

dinsdag 21 mei 2013 om 15.30 uur

door

Nicolien Thea van Ravesteijn
geboren te Leiden

PROMOTIECOMMISSIE:

Promotor:	Prof.dr. H.J. de Koning
Overige leden:	Prof.dr. M.G.M. Hunink
	Prof.dr. M.P.M.H. Rutten-van Mölken
	Prof.dr. G.J. den Heeten

CONTENTS

Chapter 1

General introduction

1.1 BREAST CANCER EPIDEMIOLOGY

Breast cancer is an important public health problem. It is the most common cancer among women both in developed and developing regions of the world.[1] In 2008 there were an estimated number of 1.38 million breast cancer cases and 458,000 deaths from the disease worldwide.[2] Breast cancer incidence and mortality increase with age (Figure 1.1) and with the projected growth and aging of many populations, the absolute number of cases and deaths is expected to increase in the future.[3]

Factors that strongly increase the risk of breast cancer include female sex and older age, as well as some genetic mutations, in particular BRCA1 and BRCA2 mutations.[4] Besides these well-established strong risk factors, two factors that increase risk by considerable amounts are breast density [5] and family history, especially when more than one family member has been diagnosed or when the relative was diagnosed at a young age.[6,7] In addition to these strong and moderate risk factors, there are numerous factors that increase risk by smaller amounts, with relative risks (RR's) in the range of 1.2 – 1.7, for example alcohol consumption,[8] physical inactivity [9] and many menstrual and reproductive factors such as parity, and ages at menarche, first birth, and menopause.[10] In addition, hormone replacement therapy (HRT), in particular use of estrogen plus progestin, has found to be associated with an increased risk for developing breast cancer.[11] Obesity decreases risk in premenopausal women, but

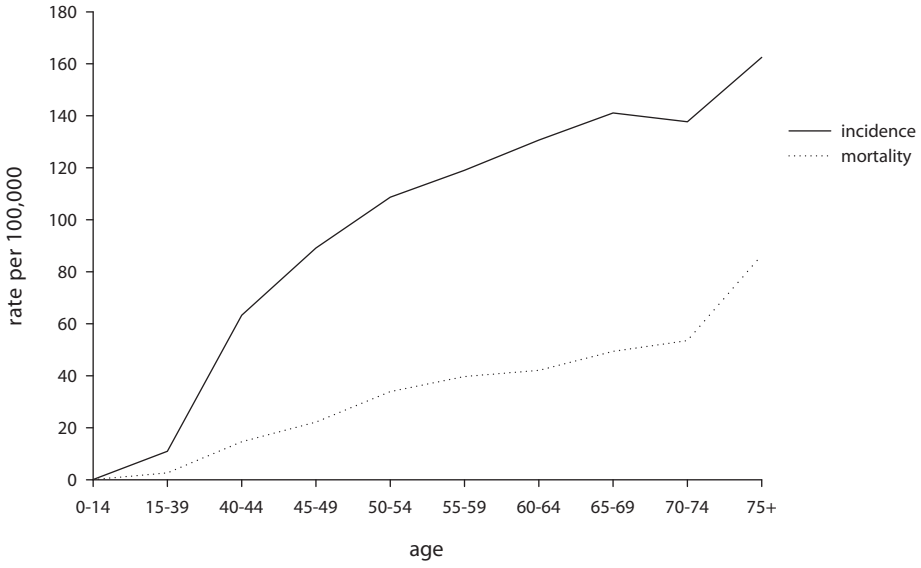


Figure 1.1. Worldwide breast cancer incidence and mortality by age in 2008. *Source: GLOBOCAN*

increases risk in postmenopausal women.[12] Risk of breast cancer has also been found to vary by race and ethnicity.[13] Non-Hispanic White women have higher risk of developing breast cancer than Black and Hispanic women,[13,14] although a part of the differences between race/ethnicity groups can be explained by differences in risk factor distribution.[15]

Partly due to changes in risk factor prevalence, for example, later age at first birth and increasing prevalence of obesity, the age-adjusted incidence rates have increased over time in many countries, for example in the United States (U.S.). At the same time, the age-standardized mortality rates have been found to decline since the early 1990's due to mammography screening and adjuvant therapy [16] (Figure 1.2). The decline was, however, not the same for all race groups. Among White women breast cancer mortality declined by 2.4% per year since 1990, while among Black women the decline was only 1.1% per year since 1991.[13]

The chance of surviving breast cancer has greatly improved over time. The chance of surviving invasive breast cancer is most strongly related to tumor size,[17] which is related to the extent of axillary lymph node involvement.[18] When tumors are diagnosed in their local stage (and are confined to the breast), in the U.S. the 5-year relative survival is 99%; the 5-year relative survival is 84% for regional disease (tumors that have spread to surrounding tissue or nearby lymph nodes) and only 23% for distant-stage disease (when the cancer has metastasized to distant organs).[19]

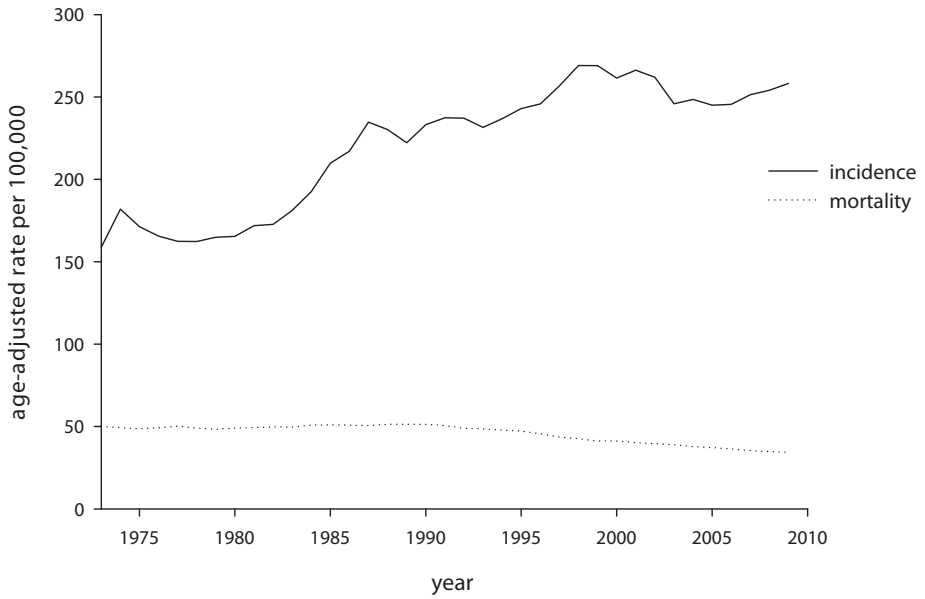


Figure 1.2. Age-adjusted (age 25+) breast cancer incidence and mortality in the U.S. from 1973-2009.
Source: Surveillance, Epidemiology and End Results (SEER) Program

1.2 INTERVENTIONS TO REDUCE THE NUMBER OF BREAST CANCER DEATHS

Primary prevention

There are few modifiable risk factors for breast cancer. Furthermore, the exceptions, such as obesity, alcohol consumption and physical inactivity, increase risk only modestly.[8,9,12] Tamoxifen has been found to decrease the incidence of breast cancer, and has been proposed as a breast cancer preventive agent, in particular for women at increased risk for the disease.[20] Its use has, however, serious side effects, which possibly leads to reluctance of physicians to prescribe it and a reluctance of women to take it.[21] The potential for primary prevention seems therefore unfortunately limited.

Screening

Secondary prevention (screening or early detection of the disease) has been first proposed in the 1950's.[22] Screening entails the examination of a group of asymptomatic individuals to detect disease at an earlier stage. The rationale behind screening is that if disease is detected earlier, the prognosis might be better and mortality might be decreased.

Screening for breast cancer can be performed using several tests, for example clinical and self breast exams, mammography, ultrasound, and magnetic resonance imaging. Mammography, which is an X-ray image of the breasts (Figure 1.3), is considered to be the most suitable test for screening, because of its relatively high sensitivity and specificity,[23] and relatively low cost.[24]

The first randomized controlled trial (RCT) of screening mammography was performed in New York in the 1960's, and has been referred to as the Health Insurance Plan (HIP) study.[25] Subsequently, several RCT's were performed in the 1970's and 1980's in Sweden [26], the United

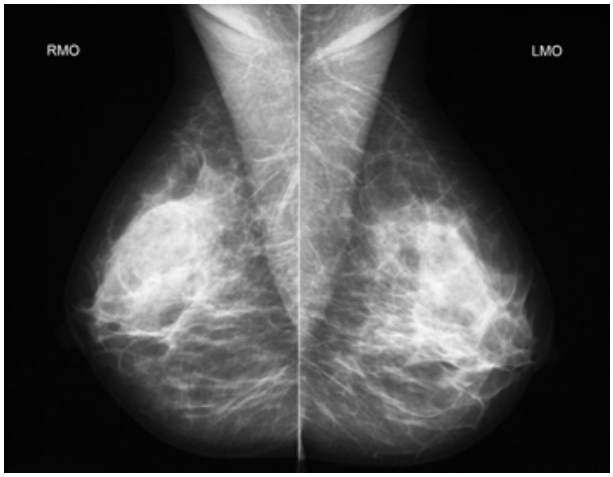


Figure 1.3. Mammogram (X-ray image of the breasts)

Kingdom [27] and Canada.[28,29] In those trials the breast cancer mortality in a group of women offered screening was compared with the breast cancer mortality in a group of women not invited to screening. Overall in a meta-analyses the breast cancer mortality reduction due to breast cancer screening was estimated to be 20% in women invited for screening.[30] Case-control studies have also shown a reduction in breast cancer mortality due to screening in nationwide programs, usually larger than the reduction found in randomized controlled trials. A recent meta-analysis found a 49% reduction in breast cancer mortality for women who are screened.[31] In addition to the lives that are saved due to screening there are also benefits for women who are diagnosed at an earlier stage, whose advanced disease is prevented and who might undergo less invasive treatment because of the earlier detection of their tumor.[32,33]

Besides the well-established benefits of mammography screening, there are unfortunately also harms. The harms vary widely in severity and occurrence. Harms include undergoing an uncomfortable and sometimes painful test,[34] and experiencing anxiety and undergoing biopsy from false-positive test results.[35,36] There is also a chance that the radiation from the mammogram causes a cancer to develop. This risk is, however, much smaller than the probability of having one's life saved because of screening.[37] In addition, it has been suggested that women attending screening might delay a visit to the doctor if they have symptoms in between screens, because of false reassurance. The risk has, however, been found to be small.[38] The most widely debated harm of breast cancer screening is overdiagnosis, i.e., the detection of tumors that would not have been detected in a woman's lifetime in the absence of screening. Although there is generally consensus that overdiagnosis occurs, the extent to which it happens is heavily debated and widely varying estimates of up to 54% have been published.[39-41] When only studies that adequately adjust for lead time and changes in breast cancer risk are included, the range of overdiagnosis estimates is considerably smaller (1% to 10%).[39]

Based on the findings of RCT's and a favorable cost-effectiveness ratio,[42,43] breast cancer screening has been implemented in many countries. In the Netherlands there has been a nationwide screening program since 1989. Since that year women between the age of 50 and 69 years old have been invited for mammography screening. Around the year 2000 the program has been extended to include 70-75 year-old women. In addition, the program began implementing digital mammography in 2004 in pilot regions, reaching nationwide coverage in 2010.[44]

In the U.S. there is no organised nationwide screening program. Breast cancer screening began to diffuse in the population in the early 1980's.[45] Some women are screened annually and start at an early age, while others get screened less frequently or start later.[45] Predictors of screening use have been found to be health insurance, income, age, family history of breast cancer, education, perceived risk, and receiving a physician recommendation.[46-48] The primary care physicians are advised by several groups about which screening scenario

to recommend. The recommended ages of starting and stopping breast cancer screening, as well as the recommended screening interval vary between the various groups that issue recommendations. For example, the U.S. Preventive Services Task Force recommends biennial screening between age 50 and 74,[49] while the American Cancer Society recommends annual screening starting at age 40 with no upper age limit as long as the woman is in good health.[50]

Treatment

In addition to these preventive measures, breast cancer mortality has been greatly reduced by advances in the treatment of breast cancer. Undergoing surgery, radiotherapy, and adjuvant treatment, i.e., chemotherapy and hormonal therapy, have all been found to decrease breast cancer mortality.[51-53]

Adjuvant therapy began diffusing in the population in the 1980's after the results of several randomized controlled trials that showed an improved survival for chemotherapy [54] and hormonal therapy.[55] The most recent meta-analysis showed that chemotherapy can reduce breast cancer mortality by about one third, depending on the type and duration of therapy.[56] Hormonal therapy (tamoxifen) has been found to reduce breast cancer mortality by approximately the same amount, but only in women with estrogen receptor (ER) positive disease.[57]

1.3 MICROSIMULATION OF BREAST CANCER SCREENING

It is not easy to estimate the effects of different mammography screening scenarios and compare them to different interventions to reduce the number of breast cancer deaths. Comparing multiple screening scenarios would require a very large number of women included in trials, which would become very expensive, while comparing the effects of screening vs. treatment would ideally require a control group without screening and treatment, which would be considered unethical. In these situations, models are considered especially valuable and can be used to make predictions. Models can be seen as a way to synthesize available data from different sources and project the findings of one study to another population, policy, or other conditions.[58]

In this thesis the MIcrosimulation SCreening ANalysis – Fatal Diameter (MISCAN-Fadia) model is used. The MISCAN model was developed in the 1980's to evaluate the effects of mass screening for cancer. It is a microsimulation model with continuous time and discrete events that simulates a dynamic population.[59] Several versions exist for different cancer types, for example for colorectal cancer [60], cervical cancer [61], lung cancer [62] and prostate cancer.[63]

MISCAN-Fadia was developed within the Cancer Intervention and Surveillance Modeling Network (CISNET).[64] CISNET is a consortium of investigators that use statistical modeling to improve our understanding of cancer control interventions in prevention, screening, and

treatment and their effects on population trends in incidence and mortality (www.cisnet.cancer.gov).

In MISCAN-Fadia, individual life histories are simulated, and the consequences of introducing a screening program on these life histories are assessed. A certain percentage of the modeled population develops pre-clinical disease. The natural history of breast cancer is modeled as a continuously growing tumor. Each tumor has a size (the fatal diameter, which differs between tumors) at which diagnosis and treatment will no longer result in cure given available treatment options. If the tumor is diagnosed and treated before it reaches the fatal diameter, the woman will be cured and will die of non-breast cancer causes. The times at which these events occur are determined by sampling from probability distributions (Figure 1.4).[64]

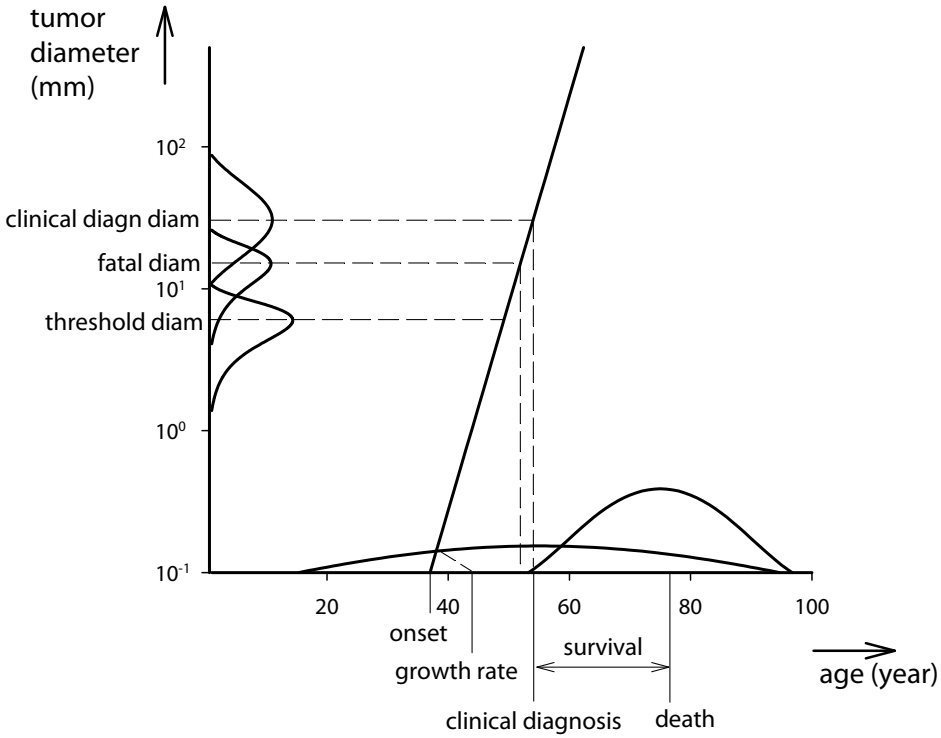


Figure 1.4. Schematic overview of the MISCAN-Fadia natural history model

When a screening program is applied, the pre-clinical tumor may be detected by screening. Each simulated tumor has a diameter at which it will be clinically diagnosed as well as a screen-detection threshold diameter. For the latter, screening test sensitivity is 0% below and 100% above this diameter. The threshold diameter is assumed to be age and period dependent. The model can incorporate the implementation of a nationwide organized screening program and its characteristics, such as screening ages, screening interval and attendance, as well as the dissemination of screening in a population.

MISCAN-Fadia includes a sub model for ductal carcinoma in situ (DCIS), with three different types of preclinical DCIS: regressive DCIS, DCIS that will be diagnosed clinically, and DCIS that will progress to invasive disease.

The model was developed using detailed data from the Two County Study.[65,66] Subsequently, several parameters were calibrated to U.S. data concerning the stage distribution and survival of breast cancer diagnosed in the period 1975-1979, before widespread use of adjuvant therapy and mammography screening.[67]

1.4 RESEARCH QUESTIONS AND OUTLINE OF THIS THESIS

The overarching goal of this thesis is to gain insights into the effects of mammography screening. The thesis consists of two parts. The first part describes the effects of screening in the population as well as in specific population subgroups and compares these effects to the effects of other interventions. In the second part the benefits and harms of different screening scenarios are evaluated.

In the following chapter, one of the main challenges in measuring the effects of breast cancer screening is described. Many recent observational studies might be looking for effects too early after implementation of the screening program. We highlight this issue by simulating the effects of a recent observational study, mimicking the study design and follow-up period and, subsequently, estimate the effect of a longer follow-up period on the estimated breast cancer mortality reduction (Chapter 2).

Next, the effects of several interventions aimed at reducing breast cancer incidence and/or mortality are assessed. The hypothetical effects of optimal adjuvant therapy use, optimal screening use (defined as annual screening between age 40-54 and biennial between age 55-99 years) and elimination of obesity are estimated for the U.S. population in future years up to 2025 (Chapter 3).

In Chapter 4 we estimate the impact of natural history, screening use, and adjuvant therapy use on the racial disparity in breast cancer mortality in the U.S. Subsequently, we extend this work by also including obesity and evaluating the impact of obesity on breast cancer incidence and mortality rates for U.S. White and Black women and assess its effect on the racial disparity in breast cancer mortality (Chapter 5).

The second part addresses questions around the optimal screening policy by predicting the effects of different screening strategies. In chapter 6, the effects of 20 screening strategies with varying starting and stopping ages applied annually or biennially were evaluated. Subsequently, the benefits and harms of breast cancer screening were evaluated in more detail for different starting ages (Chapter 7), stopping ages (Chapter 8 and Chapter 9), and screening intervals (Chapter 10). In addition, we investigated the costs and effects of screening using film vs. digital mammography and the implications of replacing film by digital mammography for a program for low-income, uninsured women in the U.S. (Chapter 11).

The thesis ends with a discussion section in which the research questions are answered and discussed. In addition, implications and directions for future research are presented.

REFERENCES

1. Bray F, Ren JS, Masuyer E, Ferlay J. Global estimates of cancer prevalence for 27 sites in the adult population in 2008. *Int J Cancer* 2013;132:1133-45.
2. Ferlay J, Shin HR, Bray F, Forman D, Mathers C, Parkin DM. Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. *Int J Cancer* 2010;127:2893-917.
3. Bray F, Moller B. Predicting the future burden of cancer. *Nat Rev Cancer* 2006;6:63-74.
4. Antoniou A, Pharoah PD, Narod S, Risch HA, Eyfjord JE, Hopper JL, et al. Average risks of breast and ovarian cancer associated with BRCA1 or BRCA2 mutations detected in case Series unselected for family history: a combined analysis of 22 studies. *Am J Hum Genet* 2003;72:1117-30.
5. Boyd NF, Guo H, Martin LJ, Sun L, Stone J, Fishell E, et al. Mammographic density and the risk and detection of breast cancer. *N Engl J Med* 2007;356:227-36.
6. Collaborative Group on Hormonal Factors in Breast Cancer. Familial breast cancer: collaborative reanalysis of individual data from 52 epidemiological studies including 58,209 women with breast cancer and 101,986 women without the disease. *Lancet* 2001;358:1389-99.
7. Egan KM, Stampfer MJ, Rosner BA, Trichopoulos D, Newcomb PA, Trentham-Dietz A, et al. Risk factors for breast cancer in women with a breast cancer family history. *Cancer Epidemiol Biomarkers Prev* 1998;7:359-64.
8. Smith-Warner SA, Spiegelman D, Yaun SS, van den Brandt PA, Folsom AR, Goldbohm RA, et al. Alcohol and breast cancer in women: a pooled analysis of cohort studies. *JAMA* 1998;279:535-40.
9. Sprague BL, Trentham-Dietz A, Newcomb PA, Titus-Ernstoff L, Hampton JM, Egan KM. Lifetime recreational and occupational physical activity and risk of in situ and invasive breast cancer. *Cancer Epidemiol Biomarkers Prev* 2007;16:236-43.
10. Reeves GK, Pirie K, Green J, Bull D, Beral V, Million Women Study C. Reproductive factors and specific histological types of breast cancer: prospective study and meta-analysis. *Br J Cancer* 2009;100:538-44.
11. Chlebowski RT, Anderson GL. Changing concepts: Menopausal hormone therapy and breast cancer. *J Natl Cancer Inst* 2012;104:517-27.
12. van den Brandt PA, Spiegelman D, Yaun SS, Adami HO, Beeson L, Folsom AR, et al. Pooled analysis of prospective cohort studies on height, weight, and breast cancer risk. *Am J Epidemiol* 2000;152:514-27.
13. Smigal C, Jemal A, Ward E, Cokkinides V, Smith R, Howe HL, et al. Trends in breast cancer by race and ethnicity: update 2006. *CA Cancer J Clin* 2006;56:168-83.
14. Siegel R, Naishadham D, Jemal A. Cancer statistics for Hispanics/Latinos, 2012. *CA Cancer J Clin* 2012;62;283-98.
15. Chlebowski RT, Chen Z, Anderson GL, Rohan T, Aragaki A, Lane D, et al. Ethnicity and breast cancer: factors influencing differences in incidence and outcome. *J Natl Cancer Inst* 2005;97:439-48.
16. Berry DA, Cronin KA, Plevritis SK, Fryback DG, Clarke L, Zelen M, et al. Effect of screening and adjuvant therapy on mortality from breast cancer. *N Engl J Med* 2005;353:1784-92.
17. Michaelson JS, Silverstein M, Wyatt J, Weber G, Moore R, Halpern E, et al. Predicting the survival of patients with breast carcinoma using tumor size. *Cancer* 2002;95:713-23.
18. Carter CL, Allen C, Henson DE. Relation of tumor size, lymph node status, and survival in 24,740 breast cancer cases. *Cancer* 1989;63:181-7.
19. Siegel R, Naishadham D, Jemal A. Cancer statistics, 2012. *CA Cancer J Clin* 2012;62:10-29.
20. Fisher B, Costantino JP, Wickerham DL, Redmond CK, Kavanah M, Cronin WM, et al. Tamoxifen for prevention of breast cancer: report of the National Surgical Adjuvant Breast and Bowel Project P-1 Study. *J Natl Cancer Inst* 1998;90:1371-88.
21. Waters EA, Cronin KA, Graubard BI, Han PK, Freedman AN. Prevalence of tamoxifen use for breast cancer chemoprevention among U.S. women. *Cancer Epidemiol Biomarkers Prev* 2010;19:443-6.
22. Gershon-Cohen J, Ingleby H, Moore L. Can mass x-ray surveys be used in detection of early cancer of the breast? *JAMA* 1956;161:1069-71.
23. Humphrey LL, Helfand M, Chan BK, Woolf SH. Breast cancer screening: a summary of the evidence for the U.S. Preventive Services Task Force. *Ann Intern Med* 2002;137:347-60.
24. Elmore JG, Armstrong K, Lehman CD, Fletcher SW. Screening for breast cancer. *JAMA* 2005;293:1245-56.
25. Shapiro S. Evidence on screening for breast cancer from a randomized trial. *Cancer* 1977;39:2772-82.
26. Nystrom L, Andersson I, Bjurstam N, Frisell J, Nordenskjold B, Rutqvist LE. Long-term effects of mammography screening: updated overview of the Swedish randomised trials. *Lancet* 2002;359:909-19.
27. Alexander FE, Anderson TJ, Brown HK, Forrest AP, Hepburn W, Kirkpatrick AE, et al. 14 years of follow-up from the Edinburgh randomised trial of breast-cancer screening. *Lancet* 1999;353:1903-8.

28. Miller AB, To T, Baines CJ, Wall C. Canadian National Breast Screening Study-2: 13-year results of a randomized trial in women aged 50-59 years. *J Natl Cancer Inst* 2000;92:1490-9.
29. Miller AB, To T, Baines CJ, Wall C. The Canadian National Breast Screening Study-1: breast cancer mortality after 11 to 16 years of follow-up. A randomized screening trial of mammography in women age 40 to 49 years. *Ann Intern Med* 2002;137:305-12.
30. Independent U. K. Panel on Breast Cancer Screening. The benefits and harms of breast cancer screening: an independent review. *Lancet* 2012;380:1778-86.
31. Nickson C, Mason KE, English DR, Kavanagh AM. Mammographic Screening and Breast Cancer Mortality: A Case-Control Study and Meta-analysis. *Cancer Epidemiol Biomarkers Prev* 2012;21:1479-88.
32. Fracheboud J, Otto SJ, van Dijck JA, Broeders MJ, Verbeek AL, de Koning HJ, et al. Decreased rates of advanced breast cancer due to mammography screening in The Netherlands. *Br J Cancer* 2004;91:861-7.
33. de Koning HJ, van Ineveld BM, de Haes JC, van Oortmarssen GJ, Klijn JG, van der Maas PJ. Advanced breast cancer and its prevention by screening. *Br J Cancer* 1992;65:950-5.
34. Aro AR, Absetz-Ylostalo P, Eerola T, Pamilo M, Lonnqvist J. Pain and discomfort during mammography. *Eur J Cancer* 1996;32A:1674-9.
35. Brewer NT, Salz T, Lillie SE. Systematic review: the long-term effects of false-positive mammograms. *Ann Intern Med* 2007;146:502-10.
36. Espasa R, Murta-Nascimento C, Bayes R, Sala M, Casamitjana M, Macia F, et al. The Psychological Impact of a False-Positive Screening Mammogram in Barcelona. *J Cancer Educ* 2012;27:780-5.
37. de Gelder R, Draisma G, Heijnsdijk EA, de Koning HJ. Population-based mammography screening below age 50: balancing radiation-induced vs prevented breast cancer deaths. *Br J Cancer* 2011;104:1214-20.
38. de Gelder R, van As E, Tilanus-Linthorst MM, Bartels CC, Boer R, Draisma G, et al. Breast cancer screening: evidence for false reassurance? *Int J Cancer* 2008;123:680-6.
39. Puliti D, Duffy SW, Miccinesi G, de Koning H, Lynge E, Zappa M, et al. Overdiagnosis in mammographic screening for breast cancer in Europe: a literature review. *J Med Screen* 2012;19 Suppl 1:42-56.
40. Jorgensen KJ, Keen JD, Gotzsche PC. Is mammographic screening justifiable considering its substantial overdiagnosis rate and minor effect on mortality? *Radiology* 2011;260:621-7.
41. Kopans DB, Smith RA, Duffy SW. Mammographic screening and "overdiagnosis". *Radiology* 2011;260:616-20.
42. de Koning HJ, van Ineveld BM, van Oortmarssen GJ, de Haes JC, Collette HJ, Hendriks JH, et al. Breast cancer screening and cost-effectiveness; policy alternatives, quality of life considerations and the possible impact of uncertain factors. *Int J Cancer* 1991;49:531-7.
43. van der Maas PJ, de Koning HJ, van Ineveld BM, van Oortmarssen GJ, Habbema JD, Lubbe KT, et al. The cost-effectiveness of breast cancer screening. *Int J Cancer* 1989;43:1055-60.
44. National Evaluation Team for Breast cancer screening (NETB). Main results 2010 breast cancer screening programme in the Netherlands. Erasmus MC, University Medical Center Rotterdam, Radboud University Medical Centre, 2012.
45. Cronin KA, Yu B, Krapcho M, Miglioretti DL, Fay MP, Izmirlian G, et al. Modeling the dissemination of mammography in the United States. *Cancer Causes Control* 2005;16:701-12.
46. Meissner HI, Breen N, Taubman ML, Vernon SW, Graubard BI. Which women aren't getting mammograms and why? (United States). *Cancer Causes Control* 2007;18:61-70.
47. Selvin E, Brett KM. Breast and cervical cancer screening: sociodemographic predictors among White, Black, and Hispanic women. *Am J Public Health* 2003;93:618-23.
48. Katapodi MC, Lee KA, Facione NC, Dodd MJ. Predictors of perceived breast cancer risk and the relation between perceived risk and breast cancer screening: a meta-analytic review. *Prev Med* 2004;38:388-402.
49. U. S. Preventive Services Task Force. Screening for breast cancer: U.S. Preventive Services Task Force recommendation statement. *Ann Intern Med* 2009;151:716-26, W-236.
50. Smith RA, Cokkinides V, Brawley OW. Cancer screening in the United States, 2012: A review of current American Cancer Society guidelines and current issues in cancer screening. *CA Cancer J Clin* 2012;62:129-142.
51. Verkooijen HM, Fioretta GM, Rapiti E, Bonnefoi H, Vlastos G, Kurtz J, et al. Patients' refusal of surgery strongly impairs breast cancer survival. *Ann Surg* 2005;242:276-80.
52. Early Breast Cancer Trialists' Collaborative Group. Effects of chemotherapy and hormonal therapy for early breast cancer on recurrence and 15-year survival: an overview of the randomised trials. *Lancet* 2005;365:1687-717.

53. Early Breast Cancer Trialists' Collaborative G, Darby S, McGale P, Correa C, Taylor C, Arriagada R, et al. Effect of radiotherapy after breast-conserving surgery on 10-year recurrence and 15-year breast cancer death: meta-analysis of individual patient data for 10,801 women in 17 randomised trials. *Lancet* 2011;378:1707-16.
54. Early Breast Cancer Trialists' Collaborative Group. Polychemotherapy for early breast cancer: an overview of the randomised trials. *Lancet* 1998;352:930-42.
55. Early Breast Cancer Trialists' Collaborative Group. Tamoxifen for early breast cancer: an overview of the randomised trials. *Lancet* 1998;351:1451-67.
56. Early Breast Cancer Trialists' Collaborative Group, Peto R, Davies C, Godwin J, Gray R, Pan HC, et al. Comparisons between different polychemotherapy regimens for early breast cancer: meta-analyses of long-term outcome among 100,000 women in 123 randomised trials. *Lancet* 2012;379:432-44.
57. Early Breast Cancer Trialists' Collaborative Group, Davies C, Godwin J, Gray R, Clarke M, Cutter D, et al. Relevance of breast cancer hormone receptors and other factors to the efficacy of adjuvant tamoxifen: patient-level meta-analysis of randomised trials. *Lancet* 2011;378:771-84.
58. Morrissey JP, Lich KH, Price RA, Mandelblatt J. Computational modeling and multilevel cancer control interventions. *J Natl Cancer Inst Monogr* 2012;2012:56-66.
59. Habbema JD, van Oortmarssen GJ, Lubbe JT, van der Maas PJ. The MISCAN simulation program for the evaluation of screening for disease. *Comput Methods Programs Biomed* 1985;20:79-93.
60. Loeve F, Boer R, van Oortmarssen GJ, van Ballegooijen M, Habbema JD. The MISCAN-COLON simulation model for the evaluation of colorectal cancer screening. *Comput Biomed Res* 1999;32:13-33.
61. Habbema JD, van Oortmarssen GJ, Lubbe JT, van der Maas PJ. Model building on the basis of Dutch cervical cancer screening data. *Maturitas* 1985;7:11-20.
62. Schultz FW, Boer R, de Koning HJ. Chapter 7: Description of MISCAN-lung, the Erasmus MC Lung Cancer microsimulation model for evaluating cancer control interventions. *Risk Anal* 2012;32 Suppl 1:S85-98.
63. Draisma G, Boer R, Otto SJ, van der Cruijsen IW, Damhuis RA, Schroder FH, et al. Lead times and overdetection due to prostate-specific antigen screening: estimates from the European Randomized Study of Screening for Prostate Cancer. *J Natl Cancer Inst* 2003;95:868-78.
64. Tan SY, van Oortmarssen GJ, de Koning HJ, Boer R, Habbema JD. The MISCAN-Fadia continuous tumor growth model for breast cancer. *J Natl Cancer Inst Monogr* 2006:56-65.
65. Tabar L, Fagerberg G, Duffy SW, Day NE, Gad A, Grontoft O. Update of the Swedish two-county program of mammographic screening for breast cancer. *Radiol Clin North Am* 1992;30:187-210.
66. Tabar L, Vitak B, Chen HH, Duffy SW, Yen MF, Chiang CF, et al. The Swedish Two-County Trial twenty years later. Updated mortality results and new insights from long-term follow-up. *Radiol Clin North Am* 2000;38:625-51.
67. Cronin KA, Mariotto AB, Clarke LD, Feuer EJ. Additional common inputs for analyzing impact of adjuvant therapy and mammography on U.S. mortality. *J Natl Cancer Inst Monogr* 2006:26-9.

PART 1

Modeling the impact of different interventions on breast cancer incidence and mortality

Chapter 2

More on screening mammography: larger effect of screening on breast cancer mortality when a longer follow-up time is used

Nicolien T. van Ravesteyn, Eveline A.M. Heijnsdijk, Harry J. de Koning

N Engl J Med. 2011 Jan 20;364(3):282-3

Kalager et al. find a smaller effect of breast-cancer screening on breast-cancer mortality than previously reported estimates.[1-2] We think that their approach, although well executed, underestimates the true effect of breast-cancer screening considerably.

If we use U.S. data on incidence, survival, and mortality, and we use the exact same method, with a 77% attendance and a follow-up time of an average 2.2 years, we also find an identical 10% effect of screening on incidence-based mortality among who were between the ages of 50 and 69 years at diagnosis. The microsimulation model known as "microsimulation of screening analysis–fatal diameter" (MISCAN-Fadia),[3] which was developed to estimate screening and treatment effects in the United States,[1,3,4] enables us to simulate a longer follow-up time. When the follow-up time is prolonged by 5 years, the effect of screening is predicted to increase to 16% (Table 2.1). Furthermore, an additional benefit is expected for women 70 years of age or older, since screening reduces the incidence and thus the incidence-based mortality in this age group.

The reported modest reduction in mortality is associated with the very short follow-up time. The actual effect of the Norwegian breast program is predicted to be larger.

Table 2.1. Rate ratios for death from breast cancer among women between the ages of 50 and 69 years at diagnosis*

Study and Input Assumption		Rate Ratio†		Screening Effect‡
		No Screening	Screening	
				percent
Kalager et al.	77% attendance and follow-up time (average, 2.2 yr)	0.82	0.72	9.9%
MISCAN-Fadia	77% attendance and short follow-up time	0.88	0.78	9.9%
	100% attendance and short follow-up time	0.88	0.75	13.4%
	77% attendance and 5 yr longer follow-up time	0.88	0.72	16.2%
	100% attendance and 5 yr longer follow-up time	0.88	0.68	20.3%

* MISCAN-Fadia denotes microsimulation of screening analysis–fatal diameter.

† The rate ratios are calculated by dividing the rate of death in the current group by the rate of death in the historical group.

‡ For the screening effect, the value shown is the difference between the rate ratio for death among women in the screening group and the rate ratio for death among women in the nonscreening group.

REFERENCES

1. Berry DA, Cronin KA, Plevritis SK, et al. Effect of screening and adjuvant therapy on mortality from breast cancer. *N Engl J Med* 2005;353:1784-92.
2. Otto SJ, Fracheboud J, Looman CW, et al. Initiation of population-based mammography screening in Dutch municipalities and effect on breast-cancer mortality: a systematic review. *Lancet* 2003;361:1411-7.
3. Tan SY, van Oortmarssen GJ, de Koning HJ, Boer R, Habbema JD. The MISCAN-Fadia continuous tumor growth model for breast cancer. *J Natl Cancer Inst Monogr* 2006:56-65.
4. Mandelblatt JS, Cronin KA, Bailey S, et al. Effects of mammography screening under different screening schedules: model estimates of potential benefits and harms. *Ann Intern Med* 2009;151:738-47.

Chapter 3

Which strategies would reduce breast cancer mortality most? Collaborative modeling of optimal screening, treatment, and obesity prevention

Jeanne S. Mandelblatt, Nicolien T. van Ravesteyn, Clyde B. Schechter, Yaojen Chang, An-Tsun Huang, Aimee M. Near, Harry J. de Koning, Ahemdin Jemal

Accepted for publication in Cancer

ABSTRACT

Background: US breast cancer mortality is declining but thousands of women still die each year.

Methods: Two established simulation models examine 6 strategies that include increased screening and/or treatment or elimination of obesity vs. continuation of current patterns. The models use common national data on incidence and obesity prevalence, competing causes of death, mammography characteristics, treatment effects and survival/cure. Parameters are modified based on obesity (defined as BMI > 30kg/m2). Outcomes are presented for the year 2025 among women age 25+ and include numbers of cases, deaths, mammograms and false positives; age-adjusted incidence and mortality; breast cancer mortality reduction and deaths averted; and probability of dying of breast cancer.

Results: If current patterns continue, the models project that there would be about 50,100-57,400 (range across models) annual breast cancer deaths in 2025. If 90% of women were screened annually from ages 40-54 and biennially from ages 55-99 (or death) then 5,100-6,100 fewer deaths would occur vs. current patterns, but incidence, mammograms and false-positives increase. If all women receive indicated systemic treatment (with no screening change) then 11,400-14,500 more deaths would be averted vs. current patterns, but increased toxicity could occur. If 100% receive screening plus indicated therapy, there would be 18,100-20,400 fewer deaths. Eliminating obesity yields 3,300-5,700 fewer breast cancer deaths vs. continuation of current obesity levels.

Conclusions: Maximal reductions in breast cancer deaths could be achieved through optimizing treatment use, followed by increasing screening use and obesity prevention.

INTRODUCTION

Breast cancer mortality continues to decrease in the United States largely due to improved treatment and screening,[1] but it remains the most commonly diagnosed non-skin cancer and the second leading female cause of cancer death, with about 40,000 dying each year.[2] Reasons for the continuing burden of breast cancer are multi-factorial and include the "graying of America", high rates of obesity that affect incidence and complicate treatment, sub-optimal access to screening and timely diagnostic follow-up, non-standard and/or delayed treatment, limits in existing screening and therapeutic paradigms, socio-economic factors that diminish survival, and unknown aspects of this disease.[3-9] Evaluating the impact of these factors on breast cancer mortality across the population separately and jointly is not feasible through empirical research.

Modeling can be used as a population laboratory to estimate the impact of changing a number of these contributing factors alone or in combination.[10-12] In this paper, we use two well-established models [1,11-15] to evaluate combinations of screening and treatment strategies[16] or elimination of obesity to decrease breast cancer deaths beyond what would be expected if current patterns persist.

METHODS

The two models (MISCAN-Fadia and SPECTRUM) were developed independently within the Cancer Intervention and Surveillance Modeling Network (CISNET) [11,12,17] and were exempt from institutional review board approval. The models estimate the impact of applying six strategies in the US female population from 2012 to 2025 vs. maintaining current patterns: 1) 90% of women screen annually from ages 40-54 and biennially from ages 55-99 (or death) and the remaining 10% do not screen at all; women receive treatment based on current patterns; 2) current screening, but 100% receive treatment indicated by age, stage and ER/HER2 status;[18] 3) 90% screening and 100% receipt of indicated treatment; 4) 100% screening and current patterns of treatment; 5) 100% screening and 100% indicated treatment; and 6) eliminate obesity but maintain current screening and treatment. While we will never achieve 100% compliance or eliminate obesity, these strategies demonstrate upper bounds of possible known approaches. We examined a hybrid strategy of more frequent screening intervals at younger ages than at older ages since there are shorter age-dependent sojourn times prior vs. after menopause. We did not impose an upper age limit to provide an estimate of the impact of screening and treatment over the entire life-course.

Model overview

Both models begin with estimates of incidence and mortality trends without screening and systemic treatment and then overlay screening use and improvements in survival associated with systemic therapy.[19] We overlay actual dissemination of screening and systemic treatments as our "base case", carrying these rates forward into the future. We superimpose the six strategies beginning in 2012 through 2025. Women are followed until death, even if that date is after 2025.

Breast cancer is depicted as having a preclinical screening-detectable period and a clinical detection point. On the basis of mammography sensitivity (or thresholds of detection), screening identifies disease in the preclinical period and results in the identification of earlier-stage or smaller tumors than occur via clinical detection. In MISCAN-Fadia, treatment results in cure for some women and in SPECTRUM results in reductions in the hazards of death. Obesity (body mass index (BMI) of ≥30kg/m^2) affects outcomes based on its age- and cohort-specific prevalence [20-23] through its impact on multiple model parameters.[15] Current obesity prevalence is projected forward.[24]

Model parameters

Both models use a common set of age-specific variables along with model-specific inputs to represent disease history (e.g., incidence, stage shifts or tumor growth).[1,10-12] For instance, based on the varying model structures, SPECTRUM uses an age–period–cohort (APC) model [25] to represent incidence from 1975-2000 without screening, while MISCAN-Fadia only uses it to estimate tumor onset. Consequently, the models have slightly different incidence rates beginning in 1975, but comparable results for trends over time.[11,12,14] Both extrapolate age-specific incidence rates forward based on rates in 2000. Obesity increases the risk for breast cancer in postmenopausal women (relative risk [RR] =1.25) but decreases the risk in premenopausal women (RR=0.60).[7,8]

The current dissemination of mammography is depicted based on the age of receipt of the first mammography and the interval between subsequent mammograms using data from the National Health Interview Survey and the Breast Cancer Surveillance Consortium (BCSC), respectively; current rates are carried forward to 2025 for each cohort and age-group.[26,27] Because mammography use does not vary by BMI except at extremes values,[28] we assume that obesity has no effect on mammography use. We use age- and BMI-specific mammography sensitivity and specificity observed in the BCSC (unpublished data) to develop model inputs[15,29] and with other data[11] to define thresholds of detection. The impact of digital mammography is evaluated in sensitivity analyses.

The American Joint Committee on Cancer (AJCC) stage distribution in the absence of screening is estimated from Surveillance, Epidemiology, and End Results (SEER) data in 1975-79.[30] After 1979 we phase in distributions among clinically detected women using BCSC data. Stage distributions among screened women are estimated using unpublished BCSC data

from 1996-2007. Since obesity is associated with more advanced tumors at diagnosis,[31,32] we use BCSC data on stage by BMI- and age-group for unscreened and screened women.

The joint distribution of ER and HER2 status by age, year, and stage is estimated from women diagnosed from 1997-2005.[14,33] Since obesity affects the rate of ER-positive tumors differentially by menopausal status,[34] we apply a RR of 0.86 and 1.78 to the probability of ER-positive cancer among obese pre- and post-menopausal women, respectively. We assume that obesity has no impact on HER2 status.

Dissemination of systemic chemo- and hormonal-therapy from 1975 to 2000 is estimated using NCI Patterns of Care data by age, year, stage, and ER status [35,36] and updated through 2010 (including trastuzumab for HER2-positive cases) using unpublished data from the National Comprehensive Cancer Network (NCCN) Outcomes Database; these data are carried forward. Strategies that include 100% indicated treatment assume: 1) ER-positive invasive cases receive chemotherapy and hormonal treatment based on age and year (tamoxifen between 1980 and 1999; tamoxifen <50 years and anastrozole if ≥50 years from 2000-2010); DCIS cases only receive hormonal therapy; 2) ER-negative invasive cases receive chemotherapy; and 3) HER2-positive tumors diagnosed in 2005 or later receive trastuzamab.[18,35,36] We assume treatment patterns do not vary by obesity. Treatment effectiveness is based on a synthesis of clinical trials.[37-40] Chemotherapy effectiveness is reduced in ~30% of obese ER-negative women based on dose reductions.[3,5,6] We assume obesity has no impact on hormonal or trastuzumab effectiveness.[8] SEER data from 1975-1979 are used to estimate breast cancer survival before screening and adjuvant treatment were available.[30] Non-breast cancer mortality is calculated by subtracting breast cancer from all-cause mortality.[41-43] The impact of obesity was incorporated using NHANES-mortality linked data.[21]

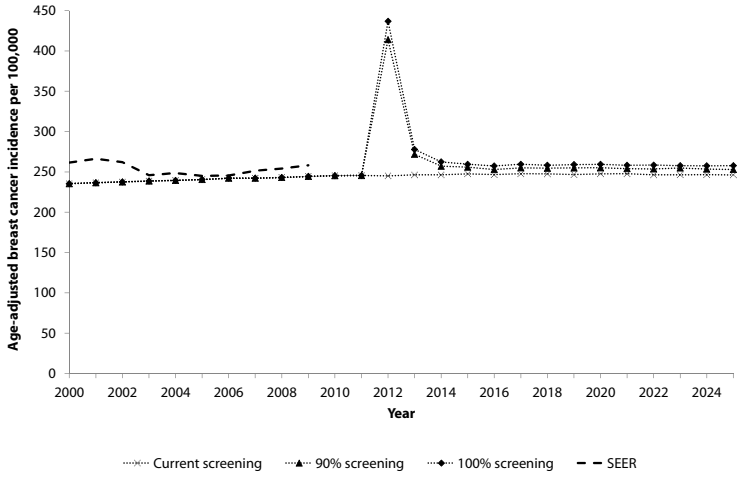
Benefits and burden

The models project the probability of breast cancer death and age-adjusted mortality rates from age 25 years to death based on continuing current patterns of screening and treatment and each of the six alternative strategies in 2025. The number of breast cancer deaths averted is calculated by applying model projections of age-specific mortality rates to the age-specific US population projections for 2025.[44] False positive mammograms are a proxy for burden and are defined as the number of mammograms read as needing further follow-up in women without cancer. Results are presented in absolute terms and incremental numbers compared to current patterns.

RESULTS

Observed incidence and mortality rates (Figures 3.1 and 3.2) and stage distributions (not shown) from 2000-2009 are accurately reproduced by both models. If current patterns continue to 2025, the models project that breast cancer mortality rates would be 31.8-36.3 (range across the models) per 100,000 women, or 50,000-57,400 breast cancer deaths in women age 25 or older (Table 3.1).

A



B

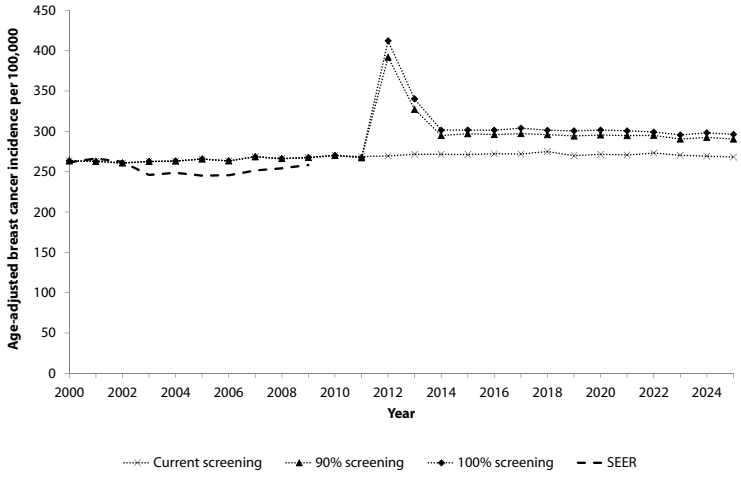


Figure 3.1. Age-adjusted breast cancer incidence rates from 2000-2025 predicted by the models for alternative screening strategies[1] vs. reported to SEER[2] for women 25 years and older (panel A SPECTRUM; panel B MISCAN-Fadia)

[1] Strategies that include treatment are not included since they do not affect incidence.

[2] Breast cancer incidence reported to SEER from 2000 to 2009.

Table 3.1. Predicted absolute outcomes in 2025 by model and alternative screening and treatment strategies vs. continuation of current patterns for women 25 years and older

Strategy	Number of Mammograms/1,000		False Positives/1,000		Age-adjusted Mortality Rate/100,000		Number of Breast Cancer Deaths [4]		Probability of Dying of Breast Cancer [5]	
	Spectrum	MISCAN-Fadia	Spectrum	MISCAN-Fadia	Spectrum	MISCAN-Fadia	Spectrum	MISCAN-Fadia	Spectrum	MISCAN-Fadia
Current screening and Rx patterns [1]	261.6	278.1	30.9	28.4	36.3	31.8	57,400	50,100	2.9%	2.9%
Current screening and 100% Rx [2]	261.4	277.9	30.9	28.4	29.1	23.0	46,000	35,600	2.3%	2.1%
90% screening and current Rx [1,3]	417.9	419.7	51.8	48.2	32.2	28.4	51,300	45,000	2.5%	2.6%
90% screening and 100% Rx [2,3]	417.7	419.4	51.9	48.1	26.1	19.7	41,700	30,700	2.1%	1.8%
100% screening and current Rx [1,3]	464.2	465.9	57.6	53.5	30.4	27.5	48,400	43,600	2.4%	2.5%
100% screening and 100% Rx [2,3]	463.9	465.6	57.5	53.5	24.7	19.1	39,300	29,700	2.0%	1.7%

[1] Current refers to screening and/or treatment as actually disseminated in the US population.

[2] All women receive indicated treatment based on age, stage and ER/HER2 status.

[3] 90% or 100% schedules are annual screening from age 40-54 and biennially from 55-99 (or death). In the 90% strategy, the remaining 10% are assumed to not have any screening.

[4] Rounded to the nearest hundred.

[5] Calculated using the Probability of Developing or Dying of Cancer Software, Version 6.6.1. Surveillance Research Program, Statistical Methodology and Applications Branch, National Cancer Institute, 2012. http://surveillance.cancer.gov/devcan.

Rx= treatment

Increased screening or indicated treatment vs. continuation of current patterns

If screening rates increase in 2012 from current patterns to 90% of women screened annually from age 40 to 54 and then biennially from age 55 onward (with no change in treatment patterns), incidence rates would have a transient increase at the start, followed by a leveling off at a higher rate than seen presently (Figure 3.1). The higher incidence would be accompanied by mortality reductions of 10.7% -11.5% in 2025 compared to continuing current screening, or about 5,100- 6,100 deaths averted among women >25 years (Table 3.2 and Figure 3.2). These benefits would require more than 140 additional mammograms per 1,000 women in 2025, including about 20 more false positive tests per 1,000 women than would occur if current patterns continue (Table 3.2).

If 100% of women are screened, incidence increases further (Figure 3.1), but mortality could be reduced (Figure 3.2) by 13.4%-16.3% in 2025 vs. continuation of current screening use. This translates into almost 6,500-8,900 more deaths averted than continuation of current patterns, but with an even greater increase in mammograms and false positives (Table 3.2). However, if screening continues at current levels, but all women receive indicated therapy, then mortality rates could be decreased by 19.8%-27.5% vs. continuation of current treatment patterns and 11,400-14,500 deaths could be avoided (Table 3.2; Figure 3.2).

Table 3.2. Predicted incremental outcomes in 2025 by model and alternative screening and treatment scenario vs. continuation of current patterns for women 25 years and older[1]

Strategy (each compared incrementally to current)	Mammograms/1,000		False Positives/1,000		Percent Mortality Reduction		Breast Cancer Deaths Averted [4]	
	Spectrum	MISCAN-Fadia	Spectrum	MISCAN-Fadia	Spectrum	MISCAN-Fadia	Spectrum	MISCAN-Fadia
Current screening and Rx patterns [1]	-	-	-	-	-	-	-	-
Current screening and 100% Rx [2]	NA	NA	NA	NA	19.8	27.5	11,400	14,500
90% screening and current Rx [1,3]	156.3	141.6	20.9	19.8	11.5	10.7	6,100	5,100
90% screening and 100% Rx [2,3]	156.1	141.3	21.0	19.7	28.1	37.9	15,700	19,400
100% screening and current Rx [1,3]	202.6	187.8	26.7	25.1	16.3	13.4	8,900	6,500
100% screening and 100% Rx [2,3]	202.3	187.5	26.6	25.1	32.1	39.9	18,100	20,400

[1] Current refers to screening and/or treatment as actually disseminated in the US population.

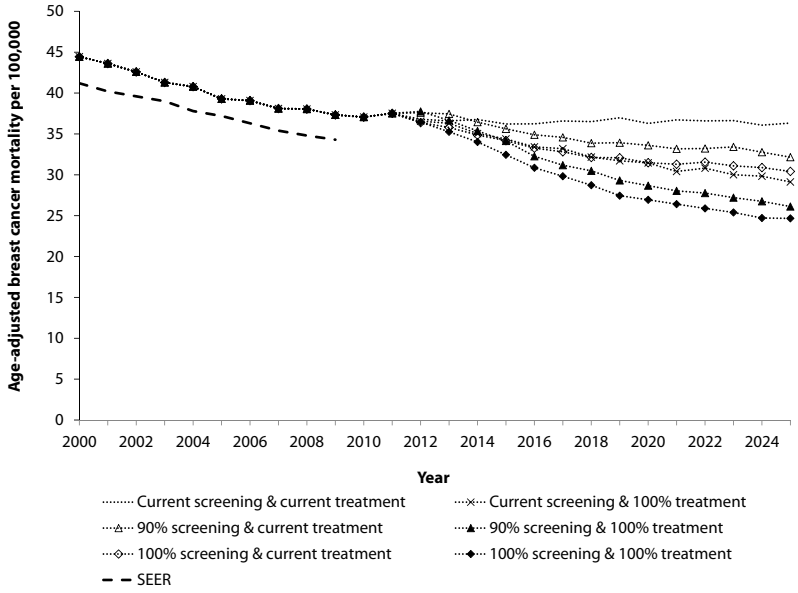
[2] All women receive indicated treatment based on age, stage and ER/HER2 status.

[3] 90% or 100% schedules are annual screening from age 40-54 and biennially from 55-99 (or death). In the 90% strategy, the remaining 10% are assumed to not have any screening.

[4] Rounded to the nearest hundred.

NA=not applicable, since no change in screening

A



B

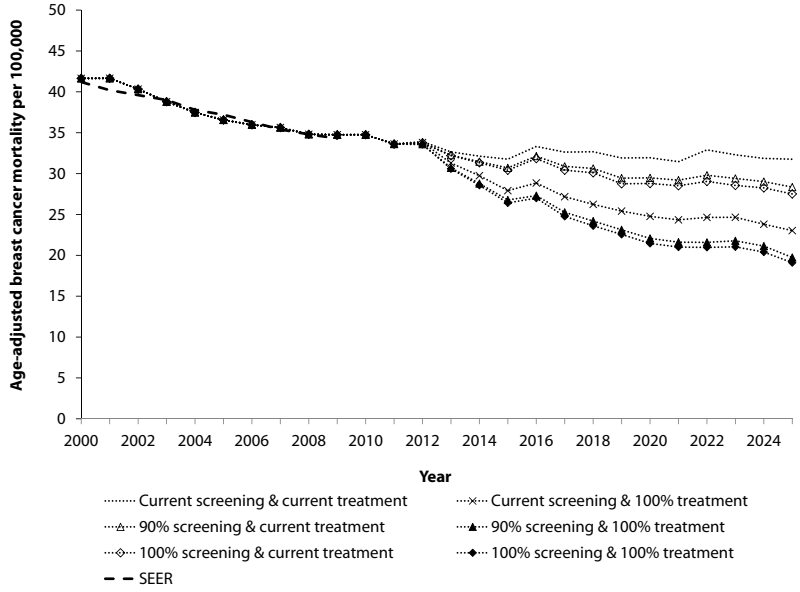


Figure 3.2. Predicted age-adjusted breast cancer mortality from 2000-2025 by alternative screening and treatment strategies vs. reported to SEER[1] for women 25 years and older (panel A SPECTRUM; panel B MISCAN-Fadia)

[1] Breast cancer mortality reported in SEER from 2000 to 2009

100% Screening and 100% indicated treatment vs. continuation of current patterns

Optimizing screening and treatment could reduce mortality by less than the sum of each approach because they interact (e.g., the better treatment is, the less screening contributes to mortality reduction). Thus, the maximum aggregate reductions that could be achieved under optimal conditions are about 18,100- 20,400 more deaths averted in 2025 vs. maintaining current patterns (Table 3.2; Figure 3.2). This corresponds to reducing a woman's lifetime probability of dying of breast cancer after age 25 years from 2.9% to 1.7-2.0% (range across models) (Table 3.1). However, even under these idealized circumstances, there would still be 29,700-39,300 breast cancer deaths.

Elimination of obesity

Obesity increases the incidence of breast cancer; 5.4%-5.6% of cases expected to occur in 2025 would be attributable to obesity if current rates are maintained (Table 3.3). If we could eradicate obesity, there could be about 3,300-5,700 fewer breast cancer deaths in 2025 in women ≥ 25 years.

Table 3.3. Projected impact of obesity on breast cancer outcomes for US women ages 25 and older in 2025 assuming current patterns of care are maintained

	Obese		Non-obese		All Women	
Incidence	Spectrum	MISCAN-Fadia	Spectrum	MISCAN-Fadia	Spectrum	MISCAN-Fadia
Age adjusted incidence rate per 100,000	270.7	287.0	241.9	252.5	251.5	263.4
# of breast cancer cases (invasive and in-situ) [1]	157,200	167,100	228,700	245,600	385,900	412,700
Attributable fraction of breast cancer cases due to obesity [2]	--	--	--	--	5.4%	5.6%
# of cases that could be avoided if obesity were eliminated [1]	--	--	--	--	20,700	23,000
Mortality						
Age adjusted mortality rate per 100,000	40.3	35.2	30.8	29.8	33.8	31.6
% Mortality Reduction [3]					9.1%	6.4%
# of breast cancer deaths [1]	23,100	20,100	32,700	30,000	55,900	50,100
Attributable fraction of breast cancer deaths due to obesity [4]	--	--	--	--	10.2%	6.6%
# of deaths that could be averted if obesity were eliminated [1]	--	--	--	--	5,700	3,300

[1] Rounded to the nearest hundred.

[2] Attributable fraction of incident cases based on formula: p*(i_O – i_N) / (p*i_O+(1-p)*i_N), where p=prevalence of obesity, i_O = incidence in obese, and i_N= incidence in non-obese

[3] Percent mortality reduction is calculated as the difference in the age-adjusted breast cancer mortality in 2025 between the current pattern and non-obese scenario divided by the age-adjusted mortality in the current pattern scenario.

[4] Attributable fraction of deaths based on formula: p*(m_O – m_N) / (p*m_O+(1-p)*m_N), where p=prevalence of obesity, m_O = mortality in obese, and m_N= mortality in non-obese

Sensitivity analysis

Improving test sensitivity or switching to digital mammography does not change the results substantially (not shown), since most lesions are slow growing and if missed on one exam are detected on the subsequent screen without much impact on mortality.

DISCUSSION

Maximal reductions in US deaths from breast cancer might be achieved through ensuring that all women have clinically indicated systemic therapy, followed by increasing screening, then obesity prevention after age 50, although greater screening could exacerbate false positives and increase incidence. Even if optimal deployment of these currently available breast cancer control strategies were achievable, the number of projected future breast cancer deaths remains high.

Optimizing use of currently available systemic therapies results in nearly double the number of deaths averted than enhancing screening levels compared to current patterns. But greater uptake of systemic therapy could lead to more therapy-related toxicity and there are barriers to use at the system, provider, and patient level. For instance, sub-optimal compliance with the full course of hormonal therapy has been noted in other research, so that modeled mortality reductions may not be realized.[45,46]

In past research, we examined differences between lifetime screening annually or biennially starting at age 40 or 50.[13] In the current study, we extend those results by examining a hybrid approach of screening annually starting at age 40 and changing to a biennial schedule at age 55 and including different rates of use. The results suggest that the majority of added benefit of increased screening use is from increasing screening levels to 90% of women using regular screening compared to current patterns, even though 100% compliance could avoid additional deaths. However, recommendations for and compliance with 90-100% regular screening at all ages will be difficult to achieve,[47,48] so that fewer deaths will be averted than projected by the models. Additional screening also imposes a burden of added false positive results and increased incidence (and over-diagnosis).[13,49,50]

Greater program efficiencies might be achieved by using a community-based approach in populations where screening and treatment services are suboptimal, as well as "personalized" risk-based approaches to target screening and treatment. The latter approach could result in more intensive screening of women with the highest risk of developing disease and deployment of therapies by women most likely to benefit and decreased use by women unlikely to benefit, minimizing harm and toxicity. However, to date, there is only a limited empiric database to support personalized approaches.[51,52] Future modeling should consider the impact of individual risk-based cancer control strategies as well as targeting geographic areas and communities with the highest burden of cancer and the least resources.

The models estimate that obesity, which presently occurs in about one-third of the female population over age 50 [20] accounts for only a modest number of breast cancer cases and deaths. Moreover, these estimates are an upper bound on what is achievable with intensive campaigns to lower obesity rates. As more data become available, it will be interesting to re-examine how strategies to reduce obesity will affect breast cancer outcomes via influences on the immune and metabolic systems that are implicated in breast cancer risk or probability of recurrence.[53]

Overall, the collaboration of two groups with different modeling approaches and structures to estimate the same endpoints by using common data provides a reasonable range of expected results. Despite these strengths and our consistent results, our study has limitations. We do not capture decrements in quality of life associated with false positive results, living with earlier knowledge of a cancer diagnosis and possible side-effects of treatment, or over-diagnosis.[50,54] We extrapolate current data forward and patterns may not continue as projected. We include mammography resources and do not include resources associated with increased use of therapy. The models also do not consider other primary prevention approaches beyond obesity reduction (e.g., tamoxifen use by high-risk women) or improvements in multimodality local therapy over time.

In summary, our results suggest that substantial improvements in US breast cancer control can be made by ensuring that all women receive indicated systemic therapy, use regular screening and avoid obesity after age 50. Multiple leverage points will be required to realize these improvements, but increasing the use of indicated systemic therapy is a necessary component of strategies for women diagnosed with breast cancer. Combinations of other approaches and new paradigms, guided by evidence from modeling, novel trials, and new scientific discovery will be needed for further reductions in the future burden of breast cancer.

Acknowledgements

We thank the Breast Cancer Surveillance Consortium (BCSC) for data provided for this study. BCSC data collection was supported by NCI cooperative agreements U01CA63740, U01CA86076, U01CA86082, U01CA63736, U01CA70013, U01CA69976, U01CA63731, U01CA70040, and HHSN261201100031C. The collection of cancer data was also supported by public health departments and cancer registries. For a full description of these sources, please see: http://www.breastscreening.cancer.gov/work/acknowledgement.html. A list of the BCSC investigators and procedures for requesting data are provided at: http://breastscreening.cancer.gov/.

REFERENCES

1. Berry DA, Cronin KA, Plevritis SK et al. Effect of screening and adjuvant therapy on mortality from breast cancer. *New Engl J Med* 2005;353:1784-92.
2. Siegel R, Naishadham D, Jemal A. Cancer statistics, 2012. *CA Cancer J.Clin.* 2012;62:10-29.
3. Rosner GL, Hargis JB, Hollis DR et al. Relationship between toxicity and obesity in women receiving adjuvant chemotherapy for breast cancer: results from cancer and leukemia group B study 8541. *J Clin Oncol* 1996;14:3000-8.
4. Gorin SS, Heck JE, Cheng B, Smith SJ. Delays in breast cancer diagnosis and treatment by racial/ethnic group. *Arch Intern Med* 2006;166:2244-52.
5. Griggs JJ, Sorbero ME, Lyman GH. Undertreatment of obese women receiving breast cancer chemotherapy. *Arch Intern Med* 2005;165:1267-73.
6. Colleoni M, Li S, Gelber RD et al. Relation between chemotherapy dose, oestrogen receptor expression, and body-mass index. *Lancet* 2005;366:1108-10.
7. van den Brandt PA, Spiegelman D, Yaun SS, et al. Pooled analysis of prospective cohort studies on height, weight, and breast cancer risk. *Am J Epidemiol* 2000;152:514-27.
8. Carmichael AR, Bates T. Obesity and breast cancer: a review of the literature. *Breast* 2004;13:85-92.
9. Vona-Davis L, Rose DP. The influence of socioeconomic disparities on breast cancer tumor biology and prognosis: a review. *J.Womens Health (Larchmt.)* 2009;18:883-93.
10. Clarke LD, Plevritis SK, Boer R, Cronin KA, Feuer EJ. A comparative review of CISNET breast models used to analyze U.S. breast cancer incidence and mortality trends. *J Natl Cancer Inst Monogr* 2006;96-105.
11. Tan SY, van Oortmarssen GJ, de Koning HJ, Boer R, Habbema JD. The MISCAN-Fadia continuous tumor growth model for breast cancer. *J Natl Cancer Inst Monogr* 2006;56-65.
12. Mandelblatt J, Schechter CB, Lawrence W, Yi B, Cullen J. The SPECTRUM population model of the impact of screening and treatment on U.S. breast cancer trends from 1975 to 2000/ *J Natl Cancer Inst Monogr* 2006;47-55.
13. Mandelblatt JS, Cronin KA, Bailey S et al. Effects of mammography screening under different screening schedules: model estimates of potential benefits and harms. *Ann Intern Med* 2009;151:738-47.
14. van Ravesteyn NT, Schechter CB, Near AM et al. Race-specific impact of natural history, mammography screening, and adjuvant treatment on breast cancer mortality rates in the United States. *Cancer Epidemiol Biomarkers Prev* 2011;20:112-22.
15. Chang Y, Schechter CB, van Ravesteyn NT, et al. Collaborative Modeling of the Impact of Obesity on Race-specific Breast Cancer Incidence and Mortality. *Breast Cancer Res Treat* 2012;136(3):823-35.
16. American Cancer Society. Cancer Prevention & Early Detection Facts & Figures 2012. Atlanta: American Cancer Society, Inc. Available from: http://www.cancer.org/acs/groups/content/@epidemiologysurveilance/documents/document/acspc-033423.pdf [Accessed October 2012].
17. CISNET Cancer Intervention and Surveillance Modeling Network. Available from: http://cisnet.cancer.gov [Accessed May 2012].
18. National Comprehensive Cancer Network. NCCN Clinical Practice Guidelines in Oncology v.2.2008. Available from: www.nccn.org/professional/physician_gls/f_guidelines.asp [Accessed September 2009].
19. Carter SB, Gartner SS, Haines MR, et al. Historical Statistics of the United States, Volume One: Population. New York: Cambridge University Press, 2006.
20. Centers for Disease Control and Prevention (CDC). National Center for Health Statistics (NCHS). National Health and Nutrition Examination Survey Data. Hyattsville, MD: U.S. Department of Health and Human Services, CDC. Available from: www.cdc.gov/nchs/nhanes.htm [Accessed May 2012].
21. Wang YC, Graubard BI, Rosenberg MA et al. Derivation of background mortality by smoking and obesity in cancer stimulation models. *Med Decis Making* 2012 Nov 6. [Epub ahead of print].
22. Wang YC, Colditz GA, Kuntz KM. Forecasting the obesity epidemic in the aging U.S. population. *Obesity* 2007;15:2855-65.
23. Wang Y, Beydoun MA. The obesity epidemic in the United States--gender, age, socioeconomic, racial/ethnic, and geographic characteristics: a systematic review and meta-regression analysis. *Epidemiol Rev* 2007;29:6-28.
24. Flegal K, Carroll M, Ogden C, Johnson C. Prevalence of obesity and trends in the distribution of body mass index among US adults, 1999-2010. *JAMA* 2012;307:491-7.
25. Holford TR, Cronin KA, Mariotto AB, Feuer EJ. Changing patterns in breast cancer incidence trends. *J Natl Cancer Inst Monogr* 2006;19-25.

26. Cronin KA, Mariotto AB, Clarke LD, Feuer EJ. Additional common inputs for analyzing impact of adjuvant therapy and mammography on U.S. mortality. *J Natl Cancer Inst Monogr* 2006;26-9.
27. Cronin KA, Yu B, Krapcho M et al. Modeling the dissemination of mammography in the United States. *Cancer Causes Control* 2005;16:701-12.
28. Zhu K, Wu H, Jatoi I, Potter J, Shriver C. Body mass index and use of mammography screening in the United States. *Preventive Medicine* 2006;381-5.
29. Breast Cancer Surveillance Consortium. Performance Measures for 3,884,059 Screening Mammography Examinations from 1996 to 2007 by Age --- based on BCSC data as of 2008. Available from: http://breastscreening.cancer.gov/data/performance/screening/2008/perf _age.html [Accessed April 2008].
30. Altekruse SF, Kosary CL, Krapcho M, Neyman N, Aminou R, Waldron W. SEER Cancer Statistics Review, 1975-2007, National Cancer Institute. Bethesda, MD. Available from: http://seer.cancer.gov/csr/1975-2007/ [Accessed January 2012].
31. Cui Y, Whiteman M, Flaws J et al. Body mass and stage of breast cancer at diagnosis. *Int J Cancer* 2002;98:279-83.
32. Hahn KM, Bondy ML, Selvan M, et al. Factors associated with advanced disease stage at diagnosis in a population-based study of patients with newly diagnosed breast cancer. *Am J Epidemiol* 2007;166:1035-44.
33. Bekele BN, Nieto-Barajas LE, Munsell MF. Analysis of partially incomplete tables of breast cancer patients' characteristics with an ordinal variable. [in press]. *Journal of Statistical Theory and Practice* 2012.
34. Suzuki R, Orsini N, Saji S, Key TJ, Wolk A. Body weight and incidence of breast cancer defined by estrogen and progesterone receptor status--a meta-analysis. *Int J Cancer* 2009;124:698-712.
35. Mariotto AB, Feuer EJ, Harlan LC, Abrams J. Dissemination of adjuvant multiagent chemotherapy and tamoxifen for breast cancer in the United States using estrogen receptor information: 1975-1999. *J Natl Cancer Inst Monogr* 2006;7-15.
36. Mariotto A, Feuer EJ, Harlan LC, Wun LM, Johnson KA, Abrams J. Trends in use of adjuvant multi-agent chemotherapy and tamoxifen for breast cancer in the United States: 1975-1999. *J Natl Cancer Inst* 2002;94:1626-34.
37. Early Breast Cancer Trialists' Collaborative Group (EBCTCG). Effects of chemotherapy and hormonal therapy for early breast cancer on recurrence and 15-year survival: an overview of the randomised trials. *Lancet* 2005; 365:1687-717.
38. Early Breast Cancer Trialists' Collaborative Group (EBCTCG). Polychemotherapy for early breast cancer: an overview of the randomised trials. *Lancet* 1998;352:930-42.
39. Early Breast Cancer Trialists' Collaborative Group (EBCTCG) Tamoxifen for early breast cancer: an overview of the randomised trials. *Lancet* 1998;351:1451-67.
40. Clarke M, Coates AS, Darby SC et al. Adjuvant chemotherapy in oestrogen-receptor-poor breast cancer: patient-level meta-analysis of randomised trials. *Lancet* 2008;371:29-40.
41. National Center for Health Statistics. United States, 2000. Hyattsville, Maryland: Public Health Service 2000. Available from: http://www.cdc.gov/nchs/products/life_tables.htm [Accessed May 2012].
42. Human Mortality Database, Centers for Disease Control and Prevention, Publications and Information Products, Life Tables 2011. Available from: http://www.mortality.org [Accessed May 2012].
43. Rosenberg MA. Competing risks to breast cancer mortality. *J Natl Cancer Inst Monogr* 2006;15-19.
44. Projected population by single year of age, sex, race, and hispanic origin for the United States: July 1, 2000 to July 1, 2050. Available from: http://www.census.gov/population/projections/data/ [Accessed May 2012].
45. Hershman DL, Shao T, Kushi LH et al. Early discontinuation and non-adherence to adjuvant hormonal therapy are associated with increased mortality in women with breast cancer. *Breast Cancer Res. Treat.* 2011;126:529-37.
46. Wigertz A, Ahlgren J, Holmqvist M et al. Adherence and discontinuation of adjuvant hormonal therapy in breast cancer patients: a population-based study. *Breast Cancer Res.Treat.* 2012;133:367-73.
47. Trivedi AN, Rakowski W, Ayanian JZ. Effect of cost sharing on screening mammography in Medicare health plans. *New Engl J Med* 2008;358:375-83.
48. Vernon SW, McQueen A, Tiro JA, del Junco DJ. Interventions to promote repeat breast cancer screening with mammography: a systematic review and meta-analysis. *J Natl Cancer Inst* 2010;102:1023-39.
49. Bleyer A, Welch HG. Effect of three decades of screening mammogrpahy on breast cancer incidence. *New Engl J Med* 2012;367:1998-2005.
50. de Gelder R, Heijnsdijk EA, van Ravesteyn NT, Fracheboud J, Draisma G, de Koning HJ. Interpreting overdiagnosis estimates in population-based mammography screening. *Epidemiol.Rev*. 2011;33:111-21.
51. Mandelblatt JS, Stout N, Trentham-Dietz A. To screen or not to screen women in their 40s for breast cancer: is personalized risk-based screening the answer? *Annals of Internal Med* 2011;155:58-60.

52. Elmore JG, Reisch LM, Barton MB et al. Efficacy of breast cancer screening in the community according to risk level. *J Natl Cancer Inst* 2005;97:1035-43.
53. Goodwin P. Energy Balance and Cancer Prognosis. In: McTiernan A. Mechanisms Associating Physical Activity With Cancer Incidence: Exercise and Immune Function. Boca Raton, FL: Taylor & Francis Group, LLC, 2006:405-36.
54. Bonomi AE, Boudreau DM, Fishman PA et al. Quality of life valuations of mammography screening. *Qual Life Res* 2008;17:801-14.

Chapter 4

Race-specific impact of natural history, mammography screening, and adjuvant treatment on breast cancer mortality rates in the United States

Nicolien T. van Ravesteyn, Clyde B. Schechter, Aimee M. Near, Eveline A.M. Heijnsdijk, Michael A. Stoto, Gerrit Draisma, Harry J. de Koning, Jeanne S. Mandelblatt

Cancer Epidemiol Biomarkers Prev. 2011 Jan;20(1):112-22

ABSTRACT

Background: U.S. Black women have higher breast cancer mortality rates than White women despite lower incidence. The aim of this study is to investigate how much of the mortality disparity can be attributed to racial differences in natural history, uptake of mammography screening, and use of adjuvant therapy.

Methods: Two simulation models use common national race-, and age-specific data for incidence, screening and treatment dissemination, stage distributions, survival, and competing mortality from 1975 to 2010. Treatment effectiveness and mammography sensitivity are assumed to be the same for both races. We sequentially substituted Black parameters into the White model to identify parameters that drive the higher mortality for Black women in the current time period.

Results: Both models accurately reproduced observed breast cancer incidence, stage and tumor size distributions, and breast cancer mortality for White women. The higher mortality for Black women could be attributed to differences in natural history parameters (26-44%), use of adjuvant therapy (11-19%) and uptake of mammography screening (7-8%), leaving 38 to 46% unexplained.

Conclusion: Black women appear to have benefited less from cancer control advances than White women, with a greater race-related gap in the use of adjuvant therapy than screening. However, a greater portion of the disparity in mortality appears to be due to differences in natural history and undetermined factors.

Impact: Breast cancer mortality may be reduced substantially by ensuring that Black women receive equal adjuvant treatment and screening as White women. More research on racial variation in breast cancer biology and treatment utilization is needed.

INTRODUCTION

In 2009, an estimated 192,370 women in the United States were diagnosed with invasive breast cancer and approximately 40,170 women were expected to die of this disease.[1] After remaining relatively constant for many years, breast cancer mortality in the United States decreased by 24% from 1990 to 2000 because of diffusion of mammography screening and improved adjuvant breast cancer treatment.[2] However, trends show a growing disparity in breast cancer mortality between Black and White women. While the breast cancer mortality rates for White women steadily decreased from 1990 onward at an average annual rate of 2.4%, the rates in Black women have only decreased by 1.1% per year during this same period.[3] The higher mortality rate for Black women (i.e., in 2006, 49 per 100,000 vs. 35 per 100,000 for White women 25 years and older) is particularly striking since breast cancer incidence is lower for Black than White women.[3]

Several factors are thought to contribute to the observed race disparity in breast cancer mortality. Black women are more likely to present with breast cancer at a later stage than White women.[4-6] This difference has been hypothesized to be due to low or irregular rates of use of mammography screening,[7] delays in follow-up after an abnormal mammogram,[8] and/or cultural beliefs and attitudes that may lead to delayed presentation of clinically diagnosed cases.[9] Even within stage categories, Black women have significantly worse survival than White women after controlling for age and tumor markers.[10] This racial difference in stage-specific survival has been hypothesized to be due to underuse of appropriate adjuvant therapy [11] and delays in treatment initiation.[12,13] Also, higher rates of comorbidities, including cardiovascular disease and diabetes may affect Black women's ability to tolerate chemotherapy and lead to dose reductions that diminish treatment effectiveness.[14] In addition, differences in tumor biology, such as higher rates of poor-prognosis triple-negative tumors in Blacks have been hypothesized to contribute to the Black-White disparities in breast cancer mortality.[15,16]

In the present study, the impact of natural history, screening use and adjuvant therapy use on the disparity in breast cancer mortality between U.S. Black and White women is estimated using 2 established, independent population simulation models.[17,18] Modeling provides an excellent "laboratory" for the evaluation of the separate contribution of these factors, because hypothetical scenarios can be simulated (e.g., changing 1 factor at a time). Our results are intended to inform health policy debates about the most effective strategies to reduce the disparity in breast cancer mortality between Black and White women and ultimately reduce the burden of breast cancer for all Americans.

METHODS

Model overviews

MISCAN-Fadia (MIcrosimulation of SCreening ANalysis-Fatal diameter) and SPECTRUM (Simulating Population Effects of Cancer Control inTerventions – Race and Understanding Mortality) are 2 simulation models developed within the Cancer Intervention and Surveillance Modeling Network (CISNET). CISNET is an international collaborative modeling effort funded by the National Cancer Institute (NCI). Collaborative modeling provides an opportunity to evaluate how model differences affect results.

The models have been described in detail elsewhere [17,18] and information about the models can be found online.[19] Briefly, both models simulate breast cancer trends in the U.S. population in the absence of screening or adjuvant treatment and then overlay screening and adjuvant treatment diffusion over time. MISCAN-Fadia models tumor growth, where tumors can be detected once they are beyond a detection threshold and cured if the tumor diameter is below a fatal diameter. In SPECTRUM, tumors progress through stages, with screening effects due to age and stage shifts and adjuvant treatment reducing the hazard of death. In both models ductal carcinoma *in situ* (DCIS) is represented as a state that can regress, remain, and be diagnosed or progress to invasive cancer.

Model parameters

Race-specific common data inputs

MISCAN-Fadia and SPECTRUM use a common race-specific set of data inputs to model breast cancer mortality by race. The demographic characteristics of multiple birth cohorts of Black and White women born between 1890 and 1985 were based on historical data for number of births and deaths from the U.S. Census and the National Center for Health Statistics (NCHS).[20]

The background incidence of breast cancer in the absence of screening was estimated from the Connecticut Tumor Registry and Surveillance, Epidemiology and End Results (SEER) data with the use of an age-period-cohort (APC) model.[21] The original APC model was used for White women and adapted for Black women using an age-specific relative risk of Black versus White incidence.

SEER data for stage distribution and breast cancer-specific survival from the period 1975 to 1979 were used to model the natural history of breast cancer in the absence of mammography screening and adjuvant therapy as these cancer control interventions did not begin to disseminate into the population in a substantial manner until after 1980.

The dissemination of mammography in the population was estimated using a 2-part model described elsewhere.[22,23] The first component of the model involves estimating the distribution of age at first mammography and the second component estimates the interval between successive screenings. For both components a race-specific variant has been used

resulting in somewhat lower screening rates for Black women.[24] For example, the screening rates were approximately 13% lower in Black than that in White women ageing 50 to 74 years in the period 1995 to 2005.

Age-, year-, AJCC (American Joint Committee on Cancer) stage-, estrogen receptor (ER), and human epidermal growth factor receptor 2 (HER2)-specific use of adjuvant therapy among Black and White women from 1975 to 2000 was estimated from data from the NCI's Patterns of Care (POC) studies [25,26] and updated through 2010 based on data from patients presenting at National Comprehensive Cancer Network (NCCN) sites. Overall, Black women were 22% and 15% less likely to receive multi-agent chemotherapy and hormonal therapy, respectively, than White women. These Black-White differences were applied to the adjuvant treatment dissemination curves from 1975 to 2010.

Non-race-specific inputs

Treatment effectiveness estimates are based on meta-analyses of randomized trial results from the Early Breast Cancer Trialists' Collaborative Group (EBCTCG).[27-29] We assume that hormonal and chemotherapy regimens are equally effective in Black and White women.[30]

The sensitivity of mammography screening is based on data from screening trials and Breast Cancer Surveillance Consortium (BCSC), and is assumed to be equal for both race groups (D. Miglioretti, personal communication, January 2008).

Model validation

SPECTRUM and MISCAN-Fadia have used several approaches to assess the internal reliability of the models and the validity of the results against external data for the U.S. population.[17,18] For the present study, we compared model predictions for incidence rates by race over time (1975-2006) with SEER data.[31] Breast cancer incidence by race for women 25 years and older was directly age standardized to 2000 U.S. standard population. We also compared model predictions of the stage (SPECTRUM) and tumor size (MISCAN-Fadia) distribution by race (assuming observed race-specific dissemination of screening) with observed SEER data in the period 2004 to 2006 (the last year of publically available SEER data at the time of analysis).

Impact of screening and adjuvant therapy on breast cancer mortality

The models were used to estimate age-adjusted breast cancer mortality rates between 1975 and 2010 for Black and White women in the United States. We calculated percent mortality reductions by comparing the mortality in scenarios with screening, adjuvant treatment, and both with the background mortality predicted in the absence of screening and adjuvant treatment. Breast cancer mortality by race for women 25 years and older was directly age standardized to 2000 U.S. standard population. The predicted breast cancer mortality rates were compared to the observed rates by race.[32]

Factors contributing to the observed mortality difference

We investigated the effect of the following factors on the difference between White and Black women in age-adjusted breast cancer mortality in a current period (the years 2004-2006): demography and breast cancer incidence, natural history (defined as the stage distribution and survival in the absence of screening and adjuvant treatment, and ER/HER2 distribution), screening use, and adjuvant treatment use. To this end, we sequentially substituted parameter values relating to these factors in the White version of each of the 2 models by corresponding values from the Black version and computed the fraction of the mortality difference between White and Black women explained by each factor.

RESULTS

Model validation

From 1975 to 2006, the observed age-adjusted breast cancer incidence rates steadily rose from 173 to 249 per 100,000 in White women and from 144 to 227 per 100,000 in Black women. These trends were accurately reproduced by both models for both races (Figure 4.1). The difference between the observed and predicted incidence was not more than 10% in either model in any year.

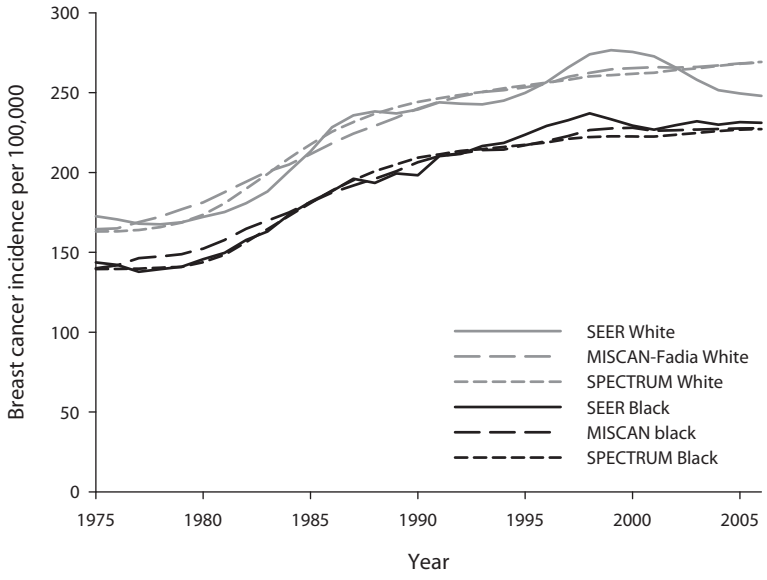


Figure 4.1. Age-adjusted incidence rates (3-year moving average) over time as observed (SEER) and predicted by MISCAN-Fadia and SPECTRUM for White and Black U.S. women 25 years and older

The observed stage distribution at diagnosis for the period 2004 to 2006 was more favorable in White than in Black women (Figure 4.2). This observation was reproduced by both models, with a more favorable tumor size distribution (MISCAN-Fadia) and stage distribution (SPECTRUM) for White than for Black women (Figure 4.2A and 4.2B). However, for Black women, both models predicted a slightly more favorable stage or tumor size than actually observed.

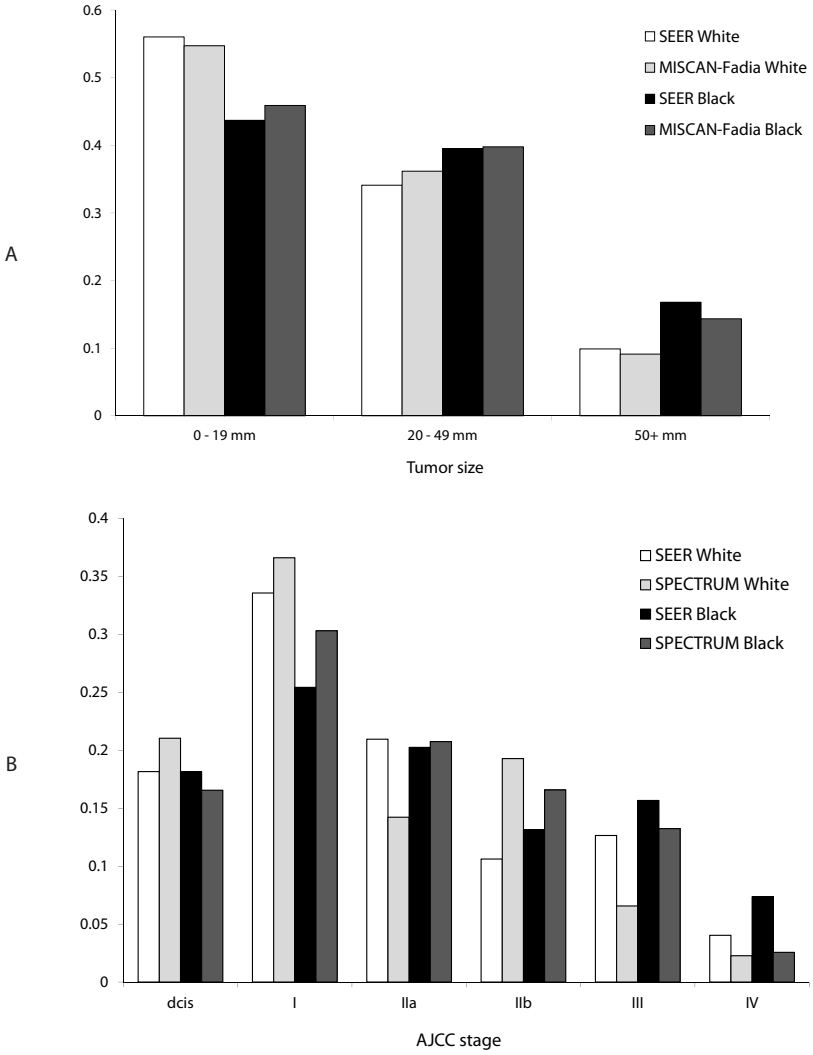


Figure 4.2 A. Age-adjusted tumor size distribution of invasive breast cancers for White and Black U.S. women 25 years and older as observed and predicted by MISCAN-Fadia in 2004 to 2006. **B**. Age-adjusted stage distribution for White and Black U.S. women 25 years and older as observed and predicted by SPECTRUM in 2004 to 2006

Impact of screening and adjuvant therapy on breast cancer mortality

There have been different trends of age-adjusted breast cancer mortality observed over time (1975-2006) by race [Figure 4.3A and 4.3B, i.e., (MISCAN-Fadia) and (SPECTRUM) for White and Figure 4.4A and 4.4B, i.e., (MISCAN-Fadia) and (SPECTRUM) for Black women].

For White women, the model-predicted breast cancer mortality rates with screening and adjuvant treatment as disseminated in the population were similar to the observed rates. The difference between the observed and predicted rates was less than 8% for all years between 1975 and 2006 in both models. Both mammography screening (19-22% mortality reduction

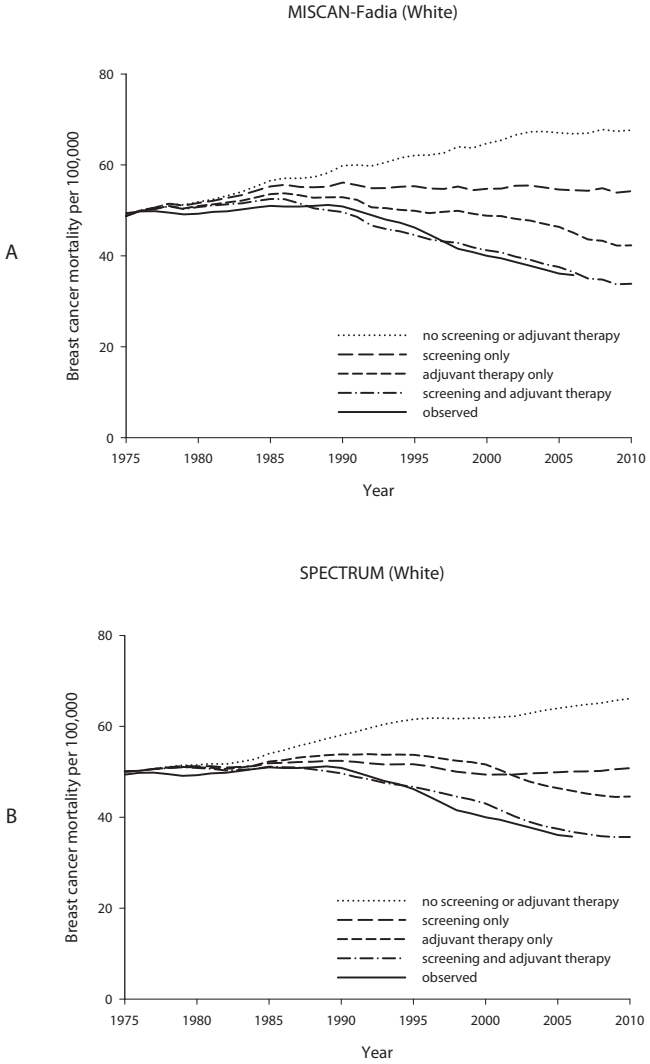


Figure 4.3. Age-adjusted breast cancer mortality rates (3-year moving averages) over time as observed and predicted in 4 scenarios for White women 25 years and older

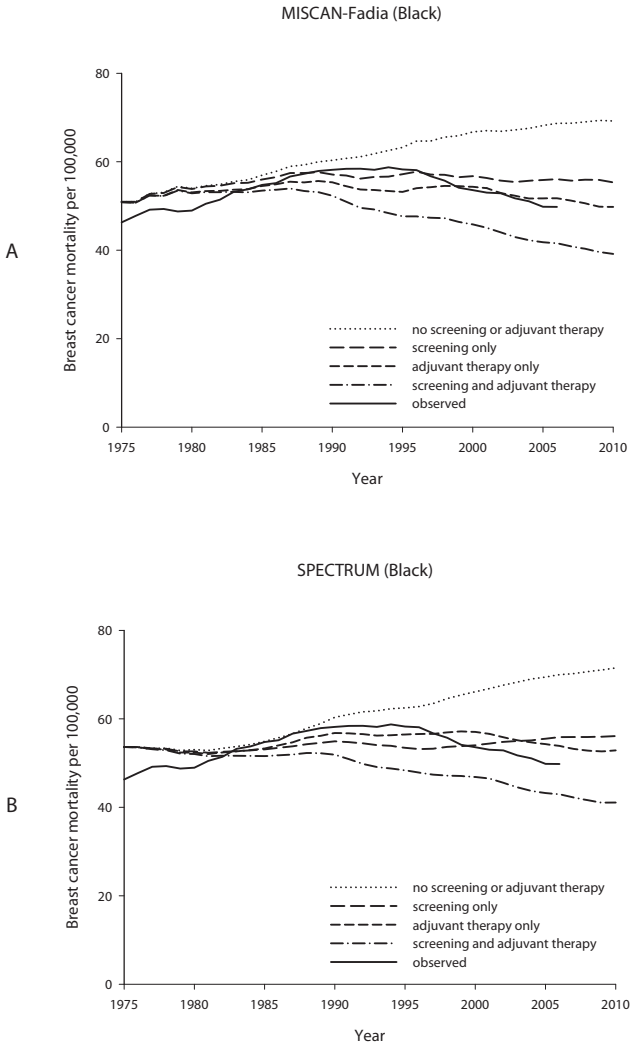


Figure 4.4. Age-adjusted breast cancer mortality rates (3-year moving averages) over time as observed and predicted in 4 scenarios for Black women 25 years and older

for MISCAN-Fadia and SPECTRUM, respectively) and adjuvant treatment (27-31% mortality reduction) contributed substantially to the observed reduction in breast cancer mortality among White women in both models (Table 4.1). The combination of mammography and adjuvant therapy is estimated to have resulted in substantially lower breast cancer mortality among White women in 2004 to 2006 (41-44% reduction) compared with a hypothetical situation without screening and adjuvant treatment.

For Black women, the model-predicted breast cancer mortality rates with screening and adjuvant treatment as disseminated in the population diverge from the observed rate. The observed breast cancer mortality decreases less and later than the predicted rates. The

predicted mortality reductions in both models were somewhat lower than for White women: mammography screening (18-20% mortality reduction), adjuvant treatment (22-24% mortality reduction), and the combination of screening and treatment (38-39% mortality reduction; Table 4.1).

Factors contributing to the observed mortality difference

Table 4.2 compares observed age-adjusted breast cancer mortality in 2004 to 2006 among White women (36.1 per 100,000 women-years) and Black women (49.8 per 100,000) to predictions from a series of models with White parameter values sequentially replaced by Black values. The models for the White population predict mortality correctly (37.4 and 37.5 per 100,000 respectively, in MISCAN-Fadia and SPECTRUM). First, replacing demographic characteristics and breast cancer incidence lowered mortality predictions to 32.5 and 32.2 per 100,000, as a result of the lower incidence for Black women. Next, changing natural history parameters responsible for a less favorable stage distribution and survival in Black women raised predicted mortality to 36.9 and 40.1 per 100,000. The lower rate of screening among Black women raised mortality to 38.4 and 41.3 and the lower use of adjuvant therapy raised

Table 4.1. Model predicted age-adjusted breast cancer mortality rates in 2004 to 2006 per 100,000 U.S. women 25 years and older

Scenarios	White				Black			
	MISCAN-Fadia		SPECTRUM		MISCAN-Fadia		SPECTRUM	
	Mortality rate per 100,000	Mortality reduction[a] (%)	Mortality rate per 100,000	Mortality reduction[a] (%)	Mortality rate per 100,000	Mortality reduction[a] (%)	Mortality rate per 100,000	Mortality reduction[a] (%)
No screening or adjuvant therapy	67.0	-	64.0	-	68.2	-	69.5	-
Screening only (as disseminated in the population)	54.6	18.6%	49.9	22.0%	55.9	18.1%	55.5	20.1%
Adjuvant therapy only (as disseminated in the population)	46.3	30.9%	46.4	27.5%	51.7	24.2%	54.3	21.9%
Screening and adjuvant therapy (as disseminated in the population)	37.5	44.0%	37.4	41.4%	41.9	38.6%	43.2	37.8%
Observed mortality rate	36.1				49.8			

[a] Mortality reductions (%) are calculated by comparing the predicted mortality to the background mortality in the scenario without screening and adjuvant therapy.

Table 4.2. The effect of sequential replacement of parameters for Black women in the White model on the predicted breast cancer mortality rate for Black women 25 years and older for the period 2004 to 2006

	White value replaced with Black value (in **bold**)							
	Observed (White)	None (White model)	**Demography and incidence**	Demography, incidence**, and natural history**	Demography, incidence, natural history, **and screening**	Demography, incidence, natural history, **and treatment**	**All (Black model)**	Observed (Black)
MISCAN-Fadia								
Mortality per 100,000	36.1	37.5	32.5	36.9	38.4	40.3	41.9	49.8
Difference, (obs-pred)			*17.4*	12.9	11.5	9.6	8.0	
% explained by replaced value [a]				26%	8%	19%	54%	
SPECTRUM								
Mortality per 100,000	36.1	37.4	32.2	40.1	41.3	42.0	43.2	49.8
Difference, (obs-pred)			*17.6*	9.8	8.5	7.8	6.6	
% explained by replaced value [a]				44%	7%	11%	62%	

[a] Calculated as the ratio of reduction of the difference between observed and predicted mortality rate and the *difference between observed and predicted mortality,* taken into account the lower incidence among Black women. So, in MISCAN-Fadia substituting Black natural history parameters into the White model explains 26% of the Black-White differences based on a reduction in the difference from 17.4 to 12.9 per 100,000, or 4.5 of the 17.4 per 100,000, i.e. 26%. obs, observed; pred, predicted

mortality to 40.3 and 42.0 per 100,000. Changing all parameters to Black values resulted in mortality predictions of 41.9 and 43.2 per 100,000 in MISCAN-Fadia and SPECTRUM, respectively. Of the difference between observed mortality and predicted mortality after taking into account the lower incidence among Blacks, natural history explained 26% (44%), screening use 8% (7%), and use of adjuvant therapy 19% (11%), leaving 46% (38%) unexplained in MISCAN-Fadia (SPECTRUM).

DISCUSSION

To our knowledge, this is the first study using collaborative population modeling to evaluate the separate and combined impact of natural history, screening use, and adjuvant therapy use on race disparities in breast cancer mortality in the United States. Both models find that the majority of the Black-White disparities in mortality outcomes is attributable to variations in natural history and yet unknown factors, and to a lesser extent to differences in use of cancer screening or treatment services. In addition, the results suggest that racial differences in adjuvant treatment dissemination contribute to the racial disparity in breast cancer mortality to a greater extent than differences in screening uptake.

Our results indicate that breast cancer natural history parameters were a major driver of race-specific differences in mortality. Also, reduced screening and treatment use in Black women, which might be related to the higher proportion of un(der)insured Black women,[33] contributed to the mortality disparity. However, the models also agree that a substantial part (38-46%) of the mortality difference by race remains unexplained, which is in line with previous work showing that several predictor variables contribute to, but do not fully explain, race differences in breast cancer survival.[34]

Several factors might account for the unexplained part of the mortality difference. First, our assumptions about some inputs being equal for Blacks and Whites might be too optimistic for Black women (e.g., equal sensitivity of screening by race). Although the predicted incidence and stage distribution for Black women fit the observed data reasonably well, both models predict a slightly more favorable stage or tumor size distribution than observed for the period 2004 to 2006. This might indicate a somewhat reduced sensitivity of mammography screening for Black women, perhaps due to lower quality imaging or interpretation. In addition, the time interval between mammogram and follow-up might differ by race. For example, women who experienced a delay between the time of mammogram and diagnosis or last diagnostic test ruling out cancer were found to be more likely to be Black than White (odds ratio 1.45; 95% confidence interval = 1.13, 1.85).[35]

Also, as observed in several randomized clinical trials, treatment efficacy was assumed to be equal for Blacks and Whites in our models.[30] However, the higher prevalence of comorbidities for Black women might lead to dose reductions outside clinical trials, resulting in

somewhat reduced treatment effectiveness in community practice. Also, Black women have been found to be less likely than white women to be treated at high-quality hospitals [36] and experience more delays between diagnosis and the beginning of treatment.[37] In addition, Black women have been found to be more likely than White women to have no surgery,[34] to discontinue treatment before completion of all courses (11% vs. 7%, respectively; $P = .07$) [38] and more likely to miss appointments (19% vs. 9%, respectively; $P=.0002$).[38] Those factors are not captured in our models, because high-quality data on the frequency of occurrence and effect on breast-cancer survival by race, age, stage, and calendar year were not available in this level of detail.

While we modeled racial differences in the distribution of known tumor prognostic markers (ER and HER2), an alternative explanation for our inability to explain the full mortality disparity is that Black women have experienced an increasing amount of aggressive tumor types over time based on less clearly defined prognostic markers. This might, for example, be related to racial differences in the prevalence of obesity in the United States, which have been increasing over the past 3 decades, with the most pronounced increases among Black women.[39] Obesity affects breast cancer mortality rates in several ways.[40] First, obesity may decrease treatment efficacy, because lower doses are delivered relative to what is recommended based on body surface area.[41] In addition, obesity may influence breast cancer survival,[42] mammography use,[43] screening performance,[44] and mammography follow-up [e.g., a higher frequency of obese women delayed return for mammography resolution compared with non-obese women (64.7% vs. 35.3%)].[45] Including obesity directly in our models would help to partition the effect of race and obesity on the disparity in breast cancer mortality. More research on the race-specific types of tumor diagnosed over time will be critical to developing the knowledge base needed to refine the natural history components of our, and other, population surveillance models.

Both models indicate that both mammography screening and adjuvant treatment contributed substantially to the observed reduction in breast cancer mortality over the past several decades for both Black and White women. This result is consistent with conclusions from past modeling work for the overall U.S. female population.[2] The predicted mortality reductions from the present study are somewhat larger than reported in past studies, probably due to greater penetration of screening in recent years and our inclusion of newer treatments (e.g., trastuzumab and aromatase inhibitors). Also, the percent mortality reductions depend somewhat on what age range is evaluated. For example, the percentages due to screening will be somewhat larger when a smaller age range excluding women unlikely to benefit from screening (age 25-40 year) is evaluated. For Black women, the predicted percent mortality reductions were somewhat lower than for White women, in particular the mortality reduction attributed to adjuvant treatment.

Our finding that treatment variations accounted for a greater amount of race variation in mortality than screening is consistent with previous research. For instance, an earlier

modeling study showed that efforts to ensure that Black women receive the same treatment as White women was a more cost-effective approach to reducing their disproportionate mortality than investing in increased screening use.[46] The finding that the effect of reduced screening use was relatively small (7-8%) is also consistent with previous work showing that the difference in screening rates between Black and White women is not very large.[47] Previous work showed that differences in mammography use can explain 10-12% of excess late-stage breast cancer among Black women compared with White women.[48,49]

The collaboration of 2 groups with different model assumptions and structure provides an excellent opportunity to cross-replicate modeling results, quantify uncertainty, and indicate which results are consistent across modeling approaches and therefore less dependent on unverifiable model assumptions. The resulting conclusions about race-specific differences in the impact of natural history, screening and adjuvant treatment on breast cancer mortality rates were similar across the 2 models and should provide greater credibility than inferences based on 1 model alone**.**

The most important limitation of the current study is the relative paucity of data on Black women, especially for the use of adjuvant treatment. Several studies that assessed the use of treatment by Black women in comprehensive cancer centers found no difference in treatment between races.[50] However, data on treatment use in the population over time are sparse for Black women. In addition, the data that are available for Black women might suffer from selection bias, with Black women who participate in trials potentially not being representative of the overall Black population. In addition, although we used the best quality data available for Black and White women as input parameters for the models, this approach led to the use of several different data sources for different variables, with the potential problem of one (or more) of these data sources not being representative of the total Black (female) population. Next, while we portrayed known differences in biology by race and age (e.g., distribution of ER- and HER2-positive tumors), some aspects of the race-specific natural history of disease are not known and/or cannot be fully captured. Even with these acknowledged limitations, the 2 models demonstrate meaningful, qualitatively similar outcomes despite variations in structure and assumptions.

The findings of the current study have important policy implications. Our results indicate that breast cancer mortality may be reduced substantially by ensuring that Black women receive adjuvant treatment and mammography screening equal in quantity and quality to that which White women receive. However, a considerable portion of the observed race differences in mortality remains unexplained. More research on racial variation in breast cancer biology, racial differences in actual treatment utilization, and responses to treatment is needed to refine optimal strategies for eliminating disparities and ensuring that all women benefit equally from medical advances and public health efforts to reduce the burden of breast cancer.

Acknowledgements

We thank the NCCN investigators for use of data on treatment dissemination.

We thank the BCSC investigators for the data they have provided for this study. A list of the BCSC investigators and procedures for requesting BCSC data for research purposes are provided at: http://breastscreening.cancer.gov/.

REFERENCES

1. American Cancer Society. Cancer Facts & Figures 2009. Atlanta, GA: American Cancer Society; 2009.
2. Berry DA, Cronin KA, Plevritis SK, Fryback DG, Clarke L, Zelen M, et al. Effect of screening and adjuvant therapy on mortality from breast cancer. *N Engl J Med* 2005;353:1784-92.
3. Smigal C, Jemal A, Ward E, Cokkinides V, Smith R, Howe HL, et al. Trends in breast cancer by race and ethnicity: update 2006. *CA Cancer J Clin* 2006;56:168-83.
4. Lantz PM, Mujahid M, Schwartz K, Janz NK, Fagerlin A, Salem B, et al. The influence of race, ethnicity, and individual socioeconomic factors on breast cancer stage at diagnosis. *Am J Public Health* 2006;96:2173-8.
5. Li CI, Malone KE, Daling JR. Differences in breast cancer stage, treatment, and survival by race and ethnicity. *Arch Intern Med* 2003;163:49-56.
6. Mandelblatt J, Andrews H, Kao R, Wallace R, Kerner J. Impact of access and social context on breast cancer stage at diagnosis. *J Health Care Poor Underserved* 1995;6:342-51.
7. Smith-Bindman R, Miglioretti DL, Lurie N, Abraham L, Barbash RB, Strzelczyk J, et al. Does utilization of screening mammography explain racial and ethnic differences in breast cancer? *Ann Intern Med* 2006;144:541-53.
8. Press R, Carrasquillo O, Sciacca RR, Giardina EG. Racial/ethnic disparities in time to follow-up after an abnormal mammogram. *J Womens Health (Larchmt)* 2008;17:923-30.
9. Lannin DR, Mathews HF, Mitchell J, Swanson MS, Swanson FH, Edwards MS. Influence of socioeconomic and cultural factors on racial differences in late-stage presentation of breast cancer. *JAMA* 1998;279:1801-7.
10. Joslyn SA, West MM. Racial differences in breast carcinoma survival. *Cancer* 2000;88:114-23.
11. Bickell NA, Wang JJ, Oluwole S, Schrag D, Godfrey H, Hiotis K, et al. Missed opportunities: racial disparities in adjuvant breast cancer treatment. *J Clin Oncol* 2006;24:1357-62.
12. Lund MJ, Brawley OP, Ward KC, Young JL, Gabram SS, Eley JW. Parity and disparity in first course treatment of invasive breast cancer. *Breast Cancer Res Treat* 2008;109:545-57.
13. Gwyn K, Bondy ML, Cohen DS, Lund MJ, Liff JM, Flagg EW, et al. Racial differences in diagnosis, treatment, and clinical delays in a population-based study of patients with newly diagnosed breast carcinoma. *Cancer* 2004;100:1595-604.
14. Tammemagi CM, Nerenz D, Neslund-Dudas C, Feldkamp C, Nathanson D. Comorbidity and survival disparities among black and white patients with breast cancer. *JAMA* 2005;294:1765-72.
15. Carey LA, Perou CM, Livasy CA, Dressler LG, Cowan D, Conway K, et al. Race, breast cancer subtypes, and survival in the Carolina Breast Cancer Study. *JAMA* 2006;295:2492-502.
16. Vona-Davis L, Rose DP. The influence of socioeconomic disparities on breast cancer tumor biology and prognosis: a review. *J Womens Health (Larchmt)* 2009;18:883-93.
17. Mandelblatt J, Schechter CB, Lawrence W, Yi B, Cullen J. The SPECTRUM population model of the impact of screening and treatment on U.S. breast cancer trends from 1975 to 2000: principles and practice of the model methods. *J Natl Cancer Inst Monogr* 2006(36):47-55.
18. Tan SY, van Oortmarssen GJ, de Koning HJ, Boer R, Habbema JD. The MISCAN-Fadia continuous tumor growth model for breast cancer. *J Natl Cancer Inst Monogr* 2006(36):56-65.
19. Available from: http://cisnet.cancer.gov/.
20. Carter SB, Gartner SS, Haines MR, Olmstead AL, Sutch R, Wright G. Historical Statistics of the United States, Volume One: Population. New York: Cambridge University Press; 2006.
21. Holford TR, Cronin KA, Mariotto AB, Feuer EJ. Changing patterns in breast cancer incidence trends. *J Natl Cancer Inst Monogr* 2006(36):19-25.
22. Cronin KA, Mariotto AB, Clarke LD, Feuer EJ. Additional common inputs for analyzing impact of adjuvant therapy and mammography on U.S. mortality. *J Natl Cancer Inst Monogr* 2006(36):26-9.
23. Cronin KA, Yu B, Krapcho M, Miglioretti DL, Fay MP, Izmirlian G, et al. Modeling the dissemination of mammography in the United States. *Cancer Causes Control* 2005;16:701-12.
24. Cronin KA, Miglioretti DL, Krapcho M, Yu B, Geller BM, Carney PA, et al. Bias associated with self-report of prior screening mammography. *Cancer Epidemiol Biomarkers Prev* 2009;18:1699-705.
25. Mariotto A, Feuer EJ, Harlan LC, Wun LM, Johnson KA, Abrams J. Trends in use of adjuvant multi-agent chemotherapy and tamoxifen for breast cancer in the United States: 1975-1999. *J Natl Cancer Inst* 2002;94:1626-34.
26. Mariotto AB, Feuer EJ, Harlan LC, Abrams J. Dissemination of adjuvant multiagent chemotherapy and tamoxifen for breast cancer in the United States using estrogen receptor information: 1975-1999. *J Natl Cancer Inst Monogr* 2006(36):7-15.

27. Early Breast Cancer Trialists' Collaborative Group. Polychemotherapy for early breast cancer: an overview of the randomised trials. *Lancet* 1998;352:930-42.
28. Early Breast Cancer Trialists' Collaborative Group. Tamoxifen for early breast cancer: an overview of the randomised trials. *Lancet* 1998;351:1451-67.
29. Early Breast Cancer Trialists' Collaborative Group. Effects of chemotherapy and hormonal therapy for early breast cancer on recurrence and 15-year survival: an overview of the randomised trials. *Lancet* 2005;365:1687-717.
30. Dignam JJ. Efficacy of systemic adjuvant therapy for breast cancer in African-American and Caucasian women. *J Natl Cancer Inst Monogr* 2001(30):36-43.
31. SEER*Stat Database: Incidence – SEER 9 Regs Research Data. Nov 2008 Sub (1973–2006) <Katrina/Rita Population Adjustment> – Linked ToCounty Attributes – TotalU.S., 1969–2006 Counties, National Cancer Institute, DCCPS, Surveillance Research Program, Cancer Statistics Branch, released April 2009, based on the November 2008 submission.
32. SEER*Stat Database: Mortality – All COD. Aggregated With State, Total U.S. (1969–2006) <Katrina/Rita Population Adjustment>, National Cancer Institute, DCCPS, Surveillance Research Program, Cancer Statistics Branch, released May 2009.
33. Thamer M, Richard C, Casebeer AW, Ray NF. Health insurance coverage among foreign-born US residents: the impact of race, ethnicity, and length of residence. *Am J Public Health* 1997;87:96-102.
34. Curtis E, Quale C, Haggstrom D, Smith-Bindman R. Racial and ethnic differences in breast cancer survival: how much is explained by screening, tumor severity, biology, treatment, comorbidities, and demographics? *Cancer* 2008;112:171-80.
35. Wujcik D, Shyr Y, Li M, Clayton MF, Ellington L, Menon U, et al. Delay in diagnostic testing after abnormal mammography in low-income women. *Oncol Nurs Forum* 2009;36:709-15.
36. Keating NL, Kouri E, He Y, Weeks JC, Winer EP. Racial differences in definitive breast cancer therapy in older women: are they explained by the hospitals where patients undergo surgery? *Med Care* 2009;47:765-73.
37. Gorin SS, Heck JE, Cheng B, Smith SJ. Delays in breast cancer diagnosis and treatment by racial/ethnic group. *Arch Intern Med* 2006;166:2244-52.
38. Hershman DL, Unger JM, Barlow WE, Hutchins LF, Martino S, Osborne CK, et al. Treatment quality and outcomes of African American versus white breast cancer patients: retrospective analysis of Southwest Oncology studies S8814/S8897. *J Clin Oncol* 2009;27:2157-62.
39. Flegal KM, Carroll MD, Ogden CL, Johnson CL. Prevalence and trends in obesity among US adults, 1999-2000. *JAMA* 2002;288(14):1723-7.
40. Carmichael AR. Obesity as a risk factor for development and poor prognosis of breast cancer. *BJOG* 2006;113:1160-6.
41. Griggs JJ, Sorbero ME, Lyman GH. Undertreatment of obese women receiving breast cancer chemotherapy. *Arch Intern Med* 2005;165:1267-73.
42. Daling JR, Malone KE, Doody DR, Johnson LG, Gralow JR, Porter PL. Relation of body mass index to tumor markers and survival among young women with invasive ductal breast carcinoma. *Cancer* 2001;92(4):720-9.
43. Cohen SS, Palmieri RT, Nyante SJ, Koralek DO, Kim S, Bradshaw P, et al. Obesity and screening for breast, cervical, and colorectal cancer in women: a review. *Cancer* 2008;112:1892-904.
44. Elmore JG, Carney PA, Abraham LA, Barlow WE, Egger JR, Fosse JS, et al. The association between obesity and screening mammography accuracy. *Arch Intern Med* 2004;164:1140-7.
45. Fair AM, Wujcik D, Lin JM, Grau A, Wilson V, Champion V, et al. Obesity, gynecological factors, and abnormal mammography follow-up in minority and medically underserved women. *J Womens Health (Larchmt)* 2009;18:1033-9.
46. Mandelblatt JS, Schechter CB, Yabroff KR, Lawrence W, Dignam J, Muennig P, et al. Benefits and costs of interventions to improve breast cancer outcomes in African American women. *J Clin Oncol* 2004;22:2554-66.
47. Breen N, Wagener DK, Brown ML, Davis WW, Ballard-Barbash R. Progress in cancer screening over a decade: results of cancer screening from the 1987, 1992, and 1998 National Health Interview Surveys. *J Natl Cancer Inst* 2001;93:1704-13.
48. Jones BA, Kasl SV, Curnen MG, Owens PH, Dubrow R. Can mammography screening explain the race difference in stage at diagnosis of breast cancer? *Cancer* 1995;75:2103-13.
49. McCarthy EP, Burns RB, Coughlin SS, Freund KM, Rice J, Marwill SL, et al. Mammography use helps to explain differences in breast cancer stage at diagnosis between older black and white women. *Ann Intern Med* 1998;128:729-36.
50. Du W, Simon MS. Racial disparities in treatment and survival of women with stage I-III breast cancer at a large academic medical center in metropolitan Detroit. *Breast Cancer Res Treat* 2005;91:243-8.

Chapter 5

Collaborative modeling of the impact of obesity on race-specific breast cancer incidence and mortality

Yaojen Chang, Clyde B. Schechter, Nicolien T. van Ravesteyn, Aimee M. Near, Eveline A.M. Heijnsdijk, Lucile Adams-Campbell, David Levy, Harry J. de Koning, Jeanne S. Mandelblatt

Breast Cancer Res Treat. 2012 Dec;136(3):823-35

ABSTRACT

Obesity affects multiple points along the breast cancer control continuum from prevention to screening and treatment, often in opposing directions. Obesity is also more prevalent in Blacks than Whites at most ages so it might contribute to observed racial disparities in mortality. We use two established simulation models from the Cancer Intervention and Surveillance Modeling Network (CISNET) to evaluate the impact of obesity on race-specific breast cancer outcomes. The models use common national data to inform parameters for the multiple US birth cohorts of Black and White women, including age- and race-specific incidence, competing mortality, mammography characteristics, and treatment effectiveness. Parameters are modified by obesity (BMI of ≥30 kg/m^2) in conjunction with its age-, race-, cohort- and time-period-specific prevalence. We measure age-standardized breast cancer incidence and mortality and cases and deaths attributable to obesity. Obesity is more prevalent among Blacks than Whites until age 74; after age 74 it is more prevalent in Whites. The models estimate that the fraction of the US breast cancer cases attributable to obesity is 3.9–4.5 % (range across models) for Whites and 2.5–3.6 % for Blacks. Given the protective effects of obesity on risk among women <50 years, elimination of obesity in this age group could increase cases for both the races, but decrease cases for women ≥50 years. Overall, obesity accounts for 4.4–9.2 % and 3.1–8.4 % of the total number of breast cancer deaths in Whites and Blacks, respectively, across models. However, variations in obesity prevalence have no net effect on race disparities in breast cancer mortality because of the opposing effects of age on risk and patterns of age- and race-specific prevalence. Despite its modest impact on breast cancer control and race disparities, obesity remains one of the few known modifiable risks for cancer and other diseases, underlining its relevance as a public health target.

INTRODUCTION

The burden of breast cancer has been decreasing over time,[1,2] but mortality gaps between Black and White women have been persistent.[3] The higher mortality rate for Black women is particularly striking now, given virtually equivalent screening rates and lower incidence than White women.[4,5] Racial differences in breast cancer outcome are cast on a backdrop of an obesity epidemic that disproportionately affects Black women. Currently more than 50% of Black women are obese (defined as a BMI of ≥30kgm^2), compared to 32.0% of White women[6,7] but there are exceptions to this overall trend, with Whites having higher rates of obesity after age 74.[6-8]

Obesity exerts numerous, often opposing effects on the chain of events leading to possible death from breast cancer. It increases breast cancer incidence in post-menopausal women, but reduces risk in pre-menopausal women.[9-13] Obesity also leads to more favorable tumor types,[13,14] greater sensitivity of detection,[15] but more advanced stage,[16-18] lower treatment effectiveness,[19,20] and greater competing mortality.[8,21,22] The Institute of Medicine recently noted that simulation modeling is particularly useful for evaluating the net impact of a factor such as obesity that affects multiple points in a disease process differentially.[23] Modeling is also helpful in evaluating the role obesity plays in racial disparities by providing a "virtual laboratory" to evaluate the impact of varying conditions that cannot be readily tested in the population, such as the net impact of reductions in obesity prevalence on breast cancer rates.[24,25]

In this article, we use two established, independent simulation models to evaluate how obesity affects breast cancer incidence and mortality in US Black and White women. We also investigate how much of the disparity in breast cancer mortality is due to the differential prevalence of obesity. Our results are intended to inform debates about effective strategies to reduce racial gaps in breast cancer mortality and reduce the burden of breast cancer for all women.

METHODS

The two models, called MISCAN and SPECTRUM, were developed within the Cancer Intervention and Surveillance Modeling Network (CISNET) [26] and were exempt from institutional review board approval. The models have been described in detail elsewhere.[2,24,27,28] Briefly, both are discrete event-driven, continuous-time state transition models that project US breast cancer population trends in the absence of screening or treatment and then overlay screening and adjuvant treatment diffusion over time. Breast cancer is depicted as having a preclinical screen-detectable period (sojourn time) and a clinical detection point. On the basis of mammography sensitivity (or thresholds of tumor size detection), screening identi-

fies disease in the preclinical screen-detectable period and results in the identification of earlier stage or smaller, more curable tumors than might occur via clinical detection, resulting in reduced breast cancer mortality. Age at diagnosis, estrogen-receptor (ER) and human epidermal growth factor-2 (HER2) status and tumor size- or stage-specific treatment have independent effects on probability of a cure (or survival). Women can die of breast cancer or of other causes.

MISCAN portrays tumor growth, where tumors can be detected once they are larger than a detection threshold and cured if the diameter at detection is below a fatal threshold. In SPECTRUM, tumors progress through stages, with screening effects due to age- and stage-shifts and treatment reducing the hazard of death. In both, ductal carcinoma in situ (DCIS) can regress, remain and be diagnosed, or progress to invasive cancer. In previous collaborations the models generated similar estimates of incidence and mortality.[2,24,29]

Model parameters

Using data from clinical trials and epidemiological studies, the models employ a common set of parameters to portray race-specific effects and then superimpose the impact of obesity on each parameter (Table 5.1). Each includes model-specific parameters to represent sojourn, lead and dwell time within stages or tumor diameter growth times and stage distribution or tumor size.[2,24,27,28]

Demographic data and obesity

The population consists of cohorts of US Black and White women born between 1890 and 2010.[30] We start in 1890 to project prevalent cancers in the 1970s.Women are assigned to being obese or non-obese based on rates observed for their age, birth cohort, race and the calendar year using prevalence data from the National Health and Nutrition Examination Survey I (1970-1975),II (1976-1980), III (1988-1994), and 1999-2004.[8] These data are extrapolated to 2009-2010 using the most recent NHANES data .[7] We do not allow transitions from obese to non-obese (or back again to obese) because there are insufficient data on how such transitions would affect all the model parameters.

Breast cancer incidence

Breast cancer incidence in the absence of screening is based on an age-period-cohort (APC) model.[31] We extrapolate forward based on rates in 2000 so do not capture the more recent decrease in incidence.[32] The APC model is adapted for Black women using an age-specific relative risk (RR) based on Black vs. White incidence.[24] Based on a synthesis of studies, obesity is modeled as increasing breast cancer incidence in post-menopausal women by a RR of 1.25 and decreasing rates in pre-menopausal women by a RR of 0.60.[9-11,33]

Table 5.1. Common model input parameters

Parameter	Race-Specific	Source	Obesity-Specific	Source
Births	Birth cohorts born from 1890 to 2000 by race	[30]	---	
Obesity prevalence	Race, cohort and year-specific	[74,75]	---	[74,75]
Incidence	Age-period cohort model with age-specific relative risk of Black versus White incidence	[24,31]	For obese (BMI ≥ 30) vs. non-obese: < 50: RR 0.6 (95% CI 0.4-1.0) 50+: RR 1.25 (95% CI 1.1-2.0) Assume obesity effect equal by race	[9-11,33]
Mammography use	Dissemination based on age- and race-specific rates for first and subsequent exams and intervals between screenings.	[34-37]	Assume obesity does not affect rate of screening	[38]
Mammography sensitivity	Age-specific rates for first and subsequent screening exams; equal by race based on unpublished BCSC data.	[39]	BMI-specific	Unpub. BCSC data
Sojourn time	2 years if age ≤ 40 2 + 0.2(age-40) if age 40-49 4 if age ≥ 50. Assume equal by race	[76]	Assume sojourn time is equal across BMI categories.	---
ER/HER2	Regression model using NCCN data from 2,646 women	[24]	Risk of ER+ breast cancer, obese vs. non-obese: < 50: RR 0.86 (95% CI 0.77-0.95) 50+: RR 0.78 (95% CI 1.50-2.11) Assume no effect HER2	[14]
Mean stage dwell time	DCIS 5 years Stage 1 2.60 yrs Stage 2a 1.26 yrs Stage 2b 1.27 yrs Stage 3 4.08 yrs Stage 4 N/A Assume equal by race	[27,28]	Assume no effect of obesity	---
Stage distribution				
Unscreened	Varies by age, race and year	[40]	BMI-specific stage	Unpub. BCSC data
Screened	Varies by age and race	Unpub. BCSC data		
Survival without Rx	Survival by race from SEER in 1975-1979	[40]	Assume no effect of obesity on breast cancer-specific	---
Treatment dissemination	Blacks 22% less likely to receive chemo; 10% (< age 50) to 15% (age 50+) less likely to get hormonal Rx than Whites	[43,44]	Obesity has no effect on treatment dissemination	[52]
Treatment effectiveness	Meta-analyses of randomized trial results; assume treatment effectiveness is equal by race	[47-51]	Reduce hazard ratios by 0.55 for obese ER-negative women who dose reductions; 30% of obese women have a dose-reduction	[19,20]
Other cause mortality	Age-race, and cohort-specific all-cause mortality rates by year	[54,55]	NHANES-linked mortality database	[8]

Mammography

The dissemination of mammography is depicted based on the age of receipt of the first mammography and the interval between subsequent mammograms using data from the Breast Cancer Surveillance Consortium (BCSC).[34,35] This parameter was extended using BCSC data to include different screening rates and intervals for Blacks and Whites.[36,37] Mammography use does not generally vary by BMI (except for the extremes for underweight and extremely obese),[38] so we assume obesity has no effect on mammography dissemination.

Sensitivity and specificity of mammographic screening for DCIS and invasive cancer were estimated by age group (under and over 50), screening round (first or subsequent), and obesity group using unpublished BCSC data.[39] There was no difference in test characteristics by race.

Stage distribution

The tumor stage distribution in absence of screening for Black and White women was estimated from the SEER data in 1975-79 before widespread use of mammography [40] and updated over time using race-specific BCSC data for unscreened (clinically detected) women. Stage distributions among screened women were estimated using race-specific BCSC data from 1996 to 2007 by screening intervals and first versus subsequent screen detection (unpublished data).

Obesity is associated with more advanced tumors at diagnosis overall [16,41] and in Blacks and Whites,[17,18] even after accounting for mammography use.[15] Therefore, we used BCSC data stratified by BMI and age group to represent the impact of obesity on stage for unscreened and screened women of both races.

Tumor biomarkers

We estimated the joint distribution of ER and HER2 status by age, year, stage, and race using data from 1997 to 2005.[24,42] As obesity affects the rate of ER+ tumors differentially by menopausal status [14], we applied RRs of 0.86 and 1.78 to the probability of having ER+ cancer among obese women <50 and 50+, respectively. We assumed that obesity had no direct impact on ER- tumors. There were insufficient data on obesity and HER2, so we assumed that obesity had no impact on HER2 distribution.

Treatment

Age-, year-, AJCC stage- (or tumor size), and ER/HER2-specific use of adjuvant hormonal and chemotherapy as disseminated from 1975 to 2000 was estimated from NCI's Patterns of Care studies [43,44] and updated through 2010 using unpublished data from the National Comprehensive Cancer Network Outcomes Database. Compared to White women, Black women were 22% less likely to receive multi-agent chemotherapy and 10% (< age 50) to 15% (age 50+) less likely to receive hormonal therapy. These Black-White differences were applied

to the treatment dissemination curves. Obese and non-obese women had similar treatment patterns and obesity did not modify treatment in Blacks.[45]

Women with ER+ invasive tumors receive hormonal treatment (tamoxifen from 1980 to 1999; tamoxifen if <50 years and anastrozole if ≥50 years from 2000 to present) and non-hormonal treatment (CMF or anthracycline-based regimen from 1975 to 1999; anthracycline-based plus taxanes from 2000 to present). Women with ER- invasive tumors receive non-hormonal therapy only. Women with DCIS and ER+ tumors receive hormonal therapy. Women with HER2+ tumors received trastuzumab beginning in 2005.[46]

Treatment effectiveness is based on RCTs.[47-50] Hormonal and chemotherapy regimens are equally effective in Black and White women.[51] We adjusted survival to reflect the fact that ~ 30% of obese patients experience dose reductions and that ER- cases having dose-reductions experience decrements in survival of 55%.[52,53] No adjustment was applied for ER+ patients. We assumed that the impact of dose reductions was the same across all race groups and that obesity had no effects on treatment effectiveness of hormonal or trastuzumab therapy.[9]

Mortality

SEER data from 1975 to 1979 were used to estimate breast cancer survival before screening and adjuvant treatment was available. Age-, race- and cohort-specific non-breast cancer mortality were calculated by subtracting breast cancer from all-cause mortality.[54,55] The impact of obesity on non-breast cancer mortality was derived from NHANES-mortality linked data.[8]

Analysis

The models simulate 1975-2020 age-adjusted breast cancer incidence and mortality rates for Black and White obese and non-obese women; adjusted is based on the standard US million population. This common referent population allows comparisons of results, as the age distributions of the population differ by race and obesity status. Overall, the US rates were estimated using a prevalence-weighted sum of the age-adjusted rates for the obese and non-obese women. Age-specific rates by race are used with the projected 2010 age-specific distribution of the respective populations [56] weighted by obesity prevalence to calculate numbers of cases and deaths. Additionally, we calculate the fraction of cases and deaths attributable to obesity. Finally, we investigate the effect of obesity on the difference in mortality between White and Black women, after considering the effect of race-differences in demography, incidence, natural history, screening use alone, adjuvant treatment use alone, and both screening and treatment. To this end, we sequentially substitute parameter values relating to these factors in the White version of the models with corresponding values from the Black version. In the final step, we add the prevalence of obesity among Blacks to the models. At each step we compute the fraction of the mortality difference between White and

Black women explained by each factor. The predicted breast cancer mortality rates at each step are also compared to the 3-year average of observed rates for Black women.[3]

Model validation

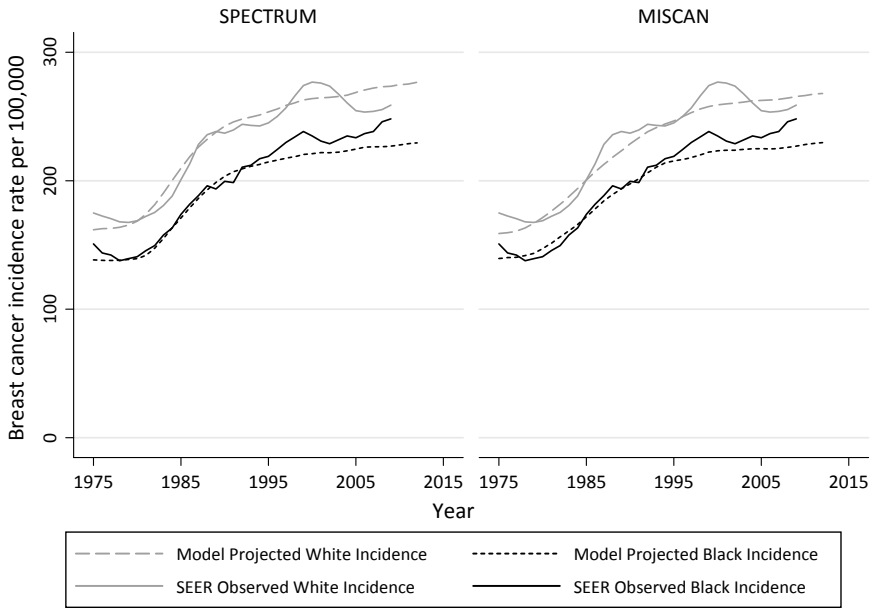
The model validation has been described in previous publications.[27,28] Results from two models provides implicit cross-validation, a range of plausible impacts of obesity, and a measure of uncertainty. Internal reliability was evaluated by combining incidence and mortality rates for obese and non-obese women in proportion to their age- and race-specific prevalence of obesity to verify that we reproduced overall population rates. Reliability of the model adaptations for obesity was also evaluated by comparing model outputs to observed SEER data.

RESULTS

Observed incidence trends were accurately reproduced by both models. Mortality rates are reproduced for White women, and the shape of the curve is similar for Blacks, but lower than SEER (Figure 5.1). The models mirror observed data showing that stage distribution (or tumor size) is more favorable in White than in Black women, but the models predict a slightly more distribution for Blacks than actually observed (Figure 5.2).

Impact of obesity on incidence

Obesity increases the incidence of breast cancer for both races and the fraction of cases attributable to obesity is similar for Whites (3.9-4.5%, across models) and Blacks (2.5-3.6%) (Table 5.2; Figure 5.3). The net impact of obesity on incidence is the result of opposing risk by age. The predicted incidence in obese women < 50 years is 37-47/100,000 for Whites and 32-44/100,000 for Blacks. Among non-obese White and Black women < 50 years, the corresponding rates are 47-60/100,000 and 43-60/100,000, respectively. Thus, elimination of obesity would actually increase the number of cases among women <50 years. For women aged 50+, obesity accounts for 5.5-6.4% and 5.3-8.1% of cases for White and Black women, respectively (data not shown).



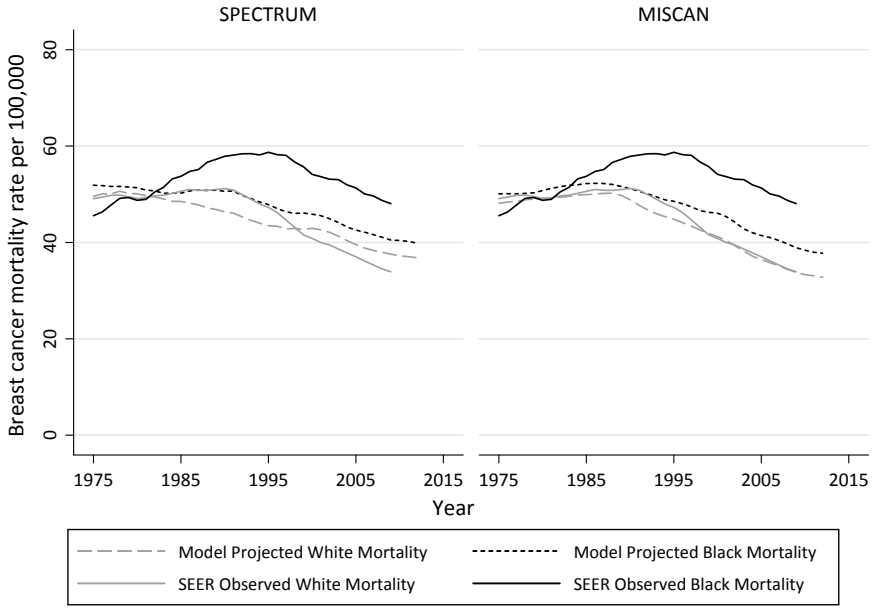


Figure 5.1. Model predicted age-adjusted breast cancer incidence and mortality by model, race and calendar year versus observed SEER Rates for US women age 25+

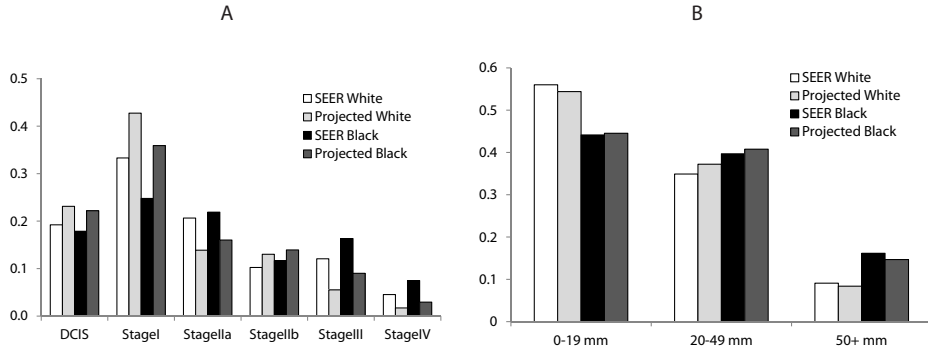


Figure 5.2. Model predicted age-adjusted breast cancer stage distributions and tumor sizes by model and race vs. observed SEER Rates in 2007-2009 for US women age 25+ (panel A SPECTRUM; panel B MISCAN)

Table 5.2. Projected breast cancer incidence in 2012 among women 25+ by race, obesity and model

	White			Black		
	Obese	Non-obese	All (weighted sum of obese + non-obese)	Obese	Non-obese	All (weighted sum of obese + non-obese)
SPECTRUM model						
Age- adjusted incidence rate per 100,000[a]	300.3	267.4	277.4	247.6	223.7	229.8
# of cases[b]	91,688	175,643	267,331	11,619	17,461	29,080
Proportion of cases in the US population attributable to obesity[c]			4.5%			3.6%
MISCAN model						
Age- adjusted incidence rate per 100,000[a]	290.7	258.2	267.5	252.7	223.0	229.7
# of cases[b]	88,989	172,768	261,757	11,463	17,717	29,180
Proportion of cases in the US population attributable to obesity[c]			3.9%			2.5%

[a] For comparability the model outputs for both race groups are age adjusted using the standard US million-population

[b] The number of cases is calculated from the model projected age- and race-specific rates, the age- and race-specific population distribution projected for 2010, and the age and race-specific prevalence of obesity

[c] The attributable fraction of cases in the overall US population that are due to obesity is estimated by the prevalence of obesity*(incidence in obese- incidence in non-obese) / (prevalence of obesity*incidence in obese+ (1-prevalence of obesity)*incidence in non-obese)

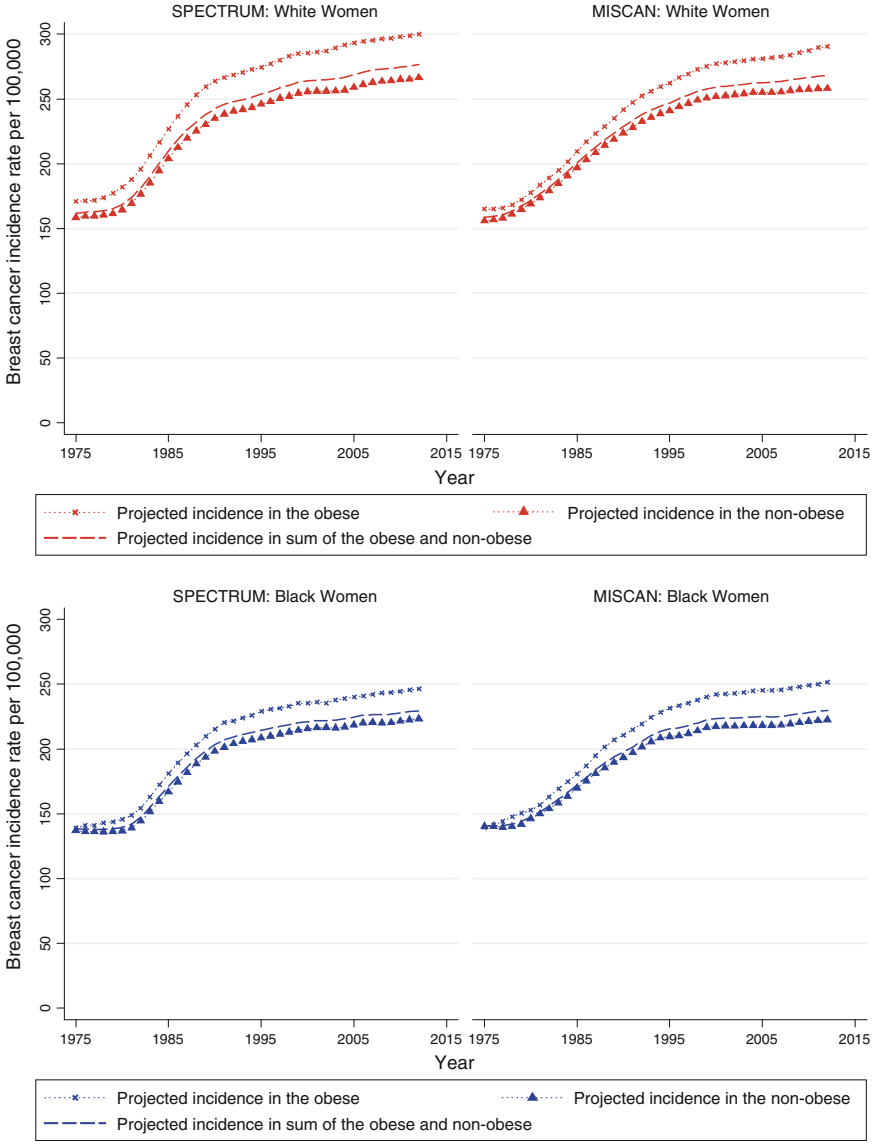


Figure 5.3. Model predicted age-adjusted breast cancer incidence rates over time by model, race and obesity for US women age 25+

Table 5.3. Projected breast cancer mortality in 2012 among women 25+ by race, obesity and model

	White			Black		
	Obese	Non-obese	All (weighted sum of obese + non-obese)	Obese	Non-obese	All (weighted sum of obese + non-obese)
SPECTRUM model						
Age- adjusted mortality rate per 100,000[a]	44.8	33.4	36.6	47.8	36.3	39.2
# of deaths from breast cancer[b]	14,363	24,398	38,761	2,059	2,872	4,931
Proportion of deaths in the US population attributable to obesity[c]			9.2%			8.4%
MISCAN model						
Age- adjusted mortality rate per 100,000[a]	35.8	30.9	32.2	41.1	36.0	37.3
# of deaths from breast cancer[b]	10,780	21,052	31,832	1,846	2,863	4,709
Proportion of deaths in the US population attributable to obesity[c]			4.4%			3.1%

[a] For comparability the model outputs for both race groups are age adjusted using the standard US million-population

[b] The number of deaths was calculated from the model projected age- and race-specific rates , the age- and race-specific population distribution projected for 2010, and the age and race-specific rates of prevalence of obesity

[c] The attributable fraction of deaths in the overall US population that are due to obesity is estimated by prevalence of obesity*(mortality in obese- mortality in non-obese) / (prevalence of obesity*mortality in obese+(1-prevalence of obesity)*mortality in non-obese)

Impact of obesity on breast cancer mortality

Obesity increases mortality rates for both race groups and accounts for 4.4-9.2% and 3.1-8.4% of the total deaths for Whites and Blacks, respectively across models (Table 5.3; Figure 5.4). This translates into 1,400-3,552 deaths in Whites and 148-412 deaths in Blacks that could be avoided each year if obesity were eliminated. Among women <50 years, obesity decreases death rates given the large decrease in incidence associated with its protective effects. For women age 50+ of both race groups, obesity accounts for about 5.1-11.5% of the deaths in the overall US population.

Obesity and impact on black-white differences in rates

The observed age-adjusted mortality rate was 33.9/100,000 in Whites and 48.1/100,000 in Blacks from 2007 to 2009 (latest years available). In Table 5.4 these values are compared to model predictions based on sequential replacement of parameter values in the White model by those from the Black model to test how much the higher prevalence of obesity in Black compared to White women affects the differences in mortality. As can be seen in Step 6, there is no net effect of race differences in obesity prevalence on mortality disparities. This result occurs because the higher prevalence of obesity among Black versus White women <50 years decreases cases and mortality, but increases these outcomes from ages 50 to 74. As White women have a higher prevalence of obesity than Blacks after age 74, substituting the

Black obesity prevalence (vs. White) decreases cases and deaths among the oldest women. Thus, differences in obesity prevalence do not account for the net age-adjusted mortality disparities between Blacks and Whites.

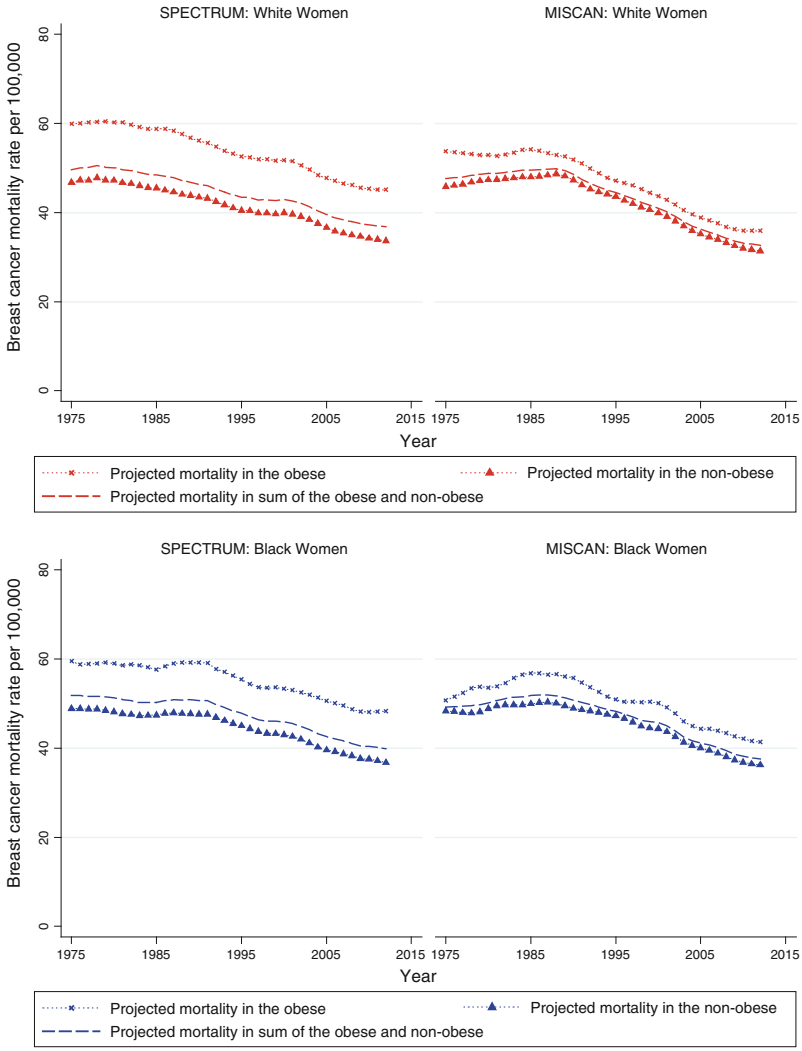


Figure 5.4. Model projected age-adjusted breast cancer mortality rates over time by model, race and obesity in US women age 25+

Table 5.4. Model predictions of breast cancer mortality differences between Black and White women age 25+ in the US in 2012 – impact of prevalence of obesity

	SEER observed 2007-2009 (White)	Model predicted (White)	Step 1: **demography, incidence**	Step 2: demography, incidence, **natural history**	Step 3: demography, incidence, natural history, **screening**	Step 4: demography, incidence, natural history, **treatment**	Step 5: demography, incidence, natural history, **screening, treatment**	Step 6: demography, incidence, natural history, screening, treatment, **obesity**	Model predicted (Black)	SEER observed 2007-2009 (Black)
	35.4									**48.1**
SPECTRUM model										
Mortality per 100,000		36.6	30. 7	36.9	37.5	38.9	39.6	39.4	39.4	
Difference, (observed-predicted)			17.4	11.2	10.6	9.2	8.5	8.7		
% explained by replaced value[a]				35.6%	3.6%	11.7%	15.5%	no net effect		
Total explained									50.2%	
MISCAN-Fadia model										
Mortality per 100,000		32.2	28.5	32.0	33.2	35.6	37.5	37.3	37.3	
Difference, (observed-predicted)			19.6	16.1	14.9	12.5	10.6	10.8		
% explained by replaced value[a]				17.8%	6.1%	18.6%	28.0%	no net effect		
Total explained									44.9%	

White value replaced with Black value (in **bold**). We used model output for 2007-2009 for these comparisons because this is the most recent year that data are available from SEER. In other analyses and tables, we include model projected rates to 2012

[a] Calculated as the ratio of reduction of the difference between observed and predicted mortality rate and the *maximum difference*. So, in SPECTRUM substituting Black natural history parameters into the White model, after adjusting for the lower incidence in Blacks than Whites explains 35.61% of the Black-White differences based on a reduction in the difference from 17.43 to 11.22 per 100,000, or 6.21 per 100,000 of the 17.43, i.e. 35.61%. If we consider the black screening rates in the white model, given black natural history and incidence, we see that the difference is drops from 11.22 to 10.60, or 0.62 per 100,000 of the 17.43%, or 3.59%. Treatment results in Step 4 are compared to Step 2. Step 5 is also compared to step 2. In the final step, Step 6 is compared to Step 5. No net effect occurs because the higher prevalence of obesity among Black versus White women under age 50 causes a net decrease in cases and mortality, a net increase from ages 50-74 and a net decrease from age 75+, since White women have a higher prevalence of obesity than Blacks at the oldest ages

DISCUSSION

This is the first study to use collaborative modeling to evaluate the impact of obesity on breast cancer incidence and mortality in White and Black women and to assess whether differences in obesity prevalence account for race disparities in mortality. We found that obesity accounts for about 3-4% of the cases and 3-9% of the deaths in both race groups. Variations in obesity prevalence have no net effect on the mortality differences between Blacks and Whites.

The overall modest impact of obesity represents the balance of an increase in cases/deaths among a large number of post-menopausal women and a decrease among a smaller number of pre-menopausal women. The obesity attributable fraction of 5.3-8.1% of cases among White and Black women 50 years and older we observed is similar to, but lower than prior US (8.9%) [15] and UK estimates (8.7%) because those included overweight and obese women.[57] If obesity were eliminated we could avoid more than 12,000 cases among White women and 1,000 in Black women. There are few measures that can prevent so many breast cancer cases, except perhaps Tamoxifen use by high-risk women.[58]

Obesity accounts for <10% of breast deaths across race groups. For colorectal cancer elimination of obesity and other risk factors could reduce mortality by up to 16%.[59] Decreasing obesity has other important effects on health, such as reductions in risk of other cancers and heart disease, and could lower health care costs,[21,22,60] making it an important public health target. However, while we can easily eliminate obesity in our model "laboratory", it is very difficult to treat in actual practice.[61,62] Thus, the modest impact projected by our models may not be achievable, but provide an upper estimate of the impact of obesity control efforts on breast cancer.

The conclusions of both models were very similar, but MISCAN uses a cure model that generates lower breast cancer mortality rates than SPECTRUM that applies a hazard of breast cancer death over time.[24,27,28] MISCAN also projects less mortality difference between obese/non-obese women due to use of continuous tumor size rather than discrete tumor stages, yielding lower obesity attributable mortality fractions. These differences capture some uncertainty and provide users with a range of results.

Obesity is a modest but potentially meaningful target in reducing the burden of breast cancer, but it does not appear to account for any net racial differences in age-adjusted mortality. This conclusion is consistent with the recent finding of Lu et al.[63] that obesity did not affect breast cancer survival in Blacks ages 35-64, although it had a modest impact for Whites. Unfortunately, that study did not have information on treatment. Others have found that adjusting for obesity does not remove Black-White differences in deaths from breast cancer.[64,65]

Our approach builds on and extends prior modeling of the impact of obesity on cancer outcomes [59] by incorporating the impact of obesity on incidence, screening, and treatment parameters and examining results separately by race. Despite the strengths of our approach,

there are some caveats that should be considered in evaluating the results. First, although weight can change over the life course,[25] we modeled obesity as constant after onset given the unavailability of data to link changes in weight to the multiple input parameters. This may underestimate the impact of obesity because a transition from non-obese to obese around menopause, which is a common pattern, maximizes risk during both periods. Next, as in other models we consider two categories of obesity [59] but do not consider an overweight category or body fat distribution, given the paucity of consistent epidemiological data across parameters for a wider range of characterizations. Others have not found a difference in estimates of risk [66] or survival [67] based on different categorizations of weight and there is no consistent evidence to suggest that level of obesity interacts with race in its effects on cancer incidence.[66] Obesity is also the net result of diet and physical activity, and these factors may affect survival.[68,69] However, it remains difficult to separately estimate the impact of these components (or their molecular correlates) [70-73] on the prevention, screening, treatment, and survival parameters included in the models. This is an important area for future research.[25] Until then, our results are useful as obesity is the most robustly related to breast cancer outcomes among measures of energy balance.[66,67]

Overall, these results represent the first comprehensive examination using comparative modeling of obesity and how it affects the entire breast cancer process from risk of disease to the type of disease at presentation and treatment impact in the context of competing mortality. The results indicate that obesity exerts a modest impact on breast cancer and does not explain net race disparities in outcome. However, obesity remains one of the few known modifiable risk factors for this and other diseases, increasing its relevance as a public health target for all women.

Acknowledgements

We thank the Breast Cancer Surveillance Consortium (BCSC) investigators for the data they have provided for this study. A list of the BCSC investigators and procedures for requesting BCSC data for research purposes are provided at http://breastscreening.cancer.gov/. The collection of cancer data used in this study was supported in part by several state public health departments and cancer registries throughout the US. For a full description of these sources, please see http://www.breastscreening.cancer.gov/work/acknowledgement.html. We thank the National Comprehensive Cancer Network investigators for use of their data on treatment dissemination. This work was supported by funding from the National Cancer Institute at the National Institutes of Health (Grant number U01CA088283, U01CA152958 and KO5CA96940 to JSM; grant number P01CA154292 to JSM and CBS; grant number R21CA149996 to LLA; and grant number UO1CA152956 to DL). Breast Cancer Surveillance Consortium data collection and sharing was supported by the National Cancer Institute (grant numbers U01CA63740, U01CA86076, U01CA86082, U01CA63736, U01CA70013, U01CA69976, U01CA63731, U01CA70040, and HHSN261201100031C).

REFERENCES

1. American Cancer Society. Breast Cancer Facts & Figures 2011–2012. http://www.cancer.org/Research/CancerFactsFigures/BreastCancerFactsFigures/breast-cancer-facts-and-figures-2011-2012. Accessed May 1, 2012.
2. Berry DA, Cronin KA, Plevritis SK, Fryback DG, Clarke L et al. Effect of screening and adjuvant therapy on mortality from breast cancer. *N Engl J Med* 2005;353:1784–1792.
3. Surveillance, Epidemiology, and End Results (SEER) Program (www.seer.cancer.gov) SEER*Stat Database: Mortality-All COD, Aggregated with state, total US (1969–2009) \Katrina/Rita Population Adjustment[-linked to county attributes-total US, 1969–2010 counties, National Cancer Institute, DCCPS. Accessed May 2012.
4. Centers for Disease Control and Prevention (CDC): Behavioral Risk Factor Surveillance Systems survey data. Atlanta, Georgia: US Department of Health and Human Services, Centers for Disease Control and Prevention. www.cdc.gov/brfss. Accessed May 2012.
5. Surveillance, Epidemiology, and End Results (SEER) Program (www.seer.cancer.gov) SEER*Stat Database: Incidence—SEER 9 Regs Research Data, Nov 2011 Sub, Vintage 2009 Pops (1973–2009) \Katrina/Rita Population Adjustment[-Linked To County Attributes-Total US, 1969–2010 Counties, National Cancer Institute, DCCPS. Accessed May 2012.
6. Flegal K, Carroll M, Ogden C, Johnson C. Prevalence of obesity and trends in the distribution of body mass index among US adults, 1999–2010. *JAMA* 2012;307:491–497.
7. NHANES, Centers for Disease Control and Prevention (CDC), National Center for Health Statistics (NCHS), National Health and Nutrition Examination Survey Data (2009). US Department of Health and Human Services, Centers for Disease Control and Prevention, Hyattsville, Maryland. www.cdc.gov/nchs/nhanes.htm. Accessed May 2012.
8. Wang YC, Graubard BI, Rosenberg MA, Kuntz KM, Zauber A, Kahle L, et al. Derivation of background mortality by smoking and obesity in cancer stimulation models. *Med Decis Making* 2012 [in press].
9. Carmichael AR, Bates T. Obesity and breast cancer: a review of the literature. *Breast* 2004;13:85–92.
10. van den Brandt PA, Spiegelman D, Yaun SS et al. Pooled analysis of prospective cohort studies on height, weight, and breast cancer risk. *Am J Epidemiol* 2000;152:514–527.
11. Michels KB, Terry KL, Willett WC. Longitudinal study on the role of body size in premenopausal breast cancer. *Arch Intern Med* 2006;166:2395–2402.
12. Yang XR, Chang-Claude J, Goode EL, Couch FJ, Nevanlinna H, Milne RL et al. Associations of breast cancer risk factors with tumor subtypes: a pooled analysis from the Breast Cancer Association Consortium studies. *J Natl Cancer Inst* 2011;103: 250–263.
13. Canchola AJ, Anton-Culver H, Bernstein L, Clarke CA, Henderson K, Ma H et al. Body size and the risk of postmenopausal breast cancer subtypes in the California Teachers Study cohort. *Cancer Causes Control* 2012;23:473–485.
14. Suzuki R, Orsini N, Saji S, Key TJ, Wolk A. Body weight and incidence of breast cancer defined by estrogen and progesterone receptor status—a meta-analysis. *Int J Cancer* 2009;124:698–712.
15. Kerlikowske K, Walker R, Miglioretti DL, Desai A, Ballard-Barbash R, Buist DS. Obesity, mammography use and accuracy, and advanced breast cancer risk. *J Natl Cancer Inst* 2008;100:1724–1733.
16. Hahn KM, Bondy ML, Selvan M et al. Factors associated with advanced disease stage at diagnosis in a population-based study of patients with newly diagnosed breast cancer. *Am J Epidemiol* 2007;166:1035–1044.
17. Cui Y, Whiteman M, Langenberg P, Sexton M, Tkaczuk K, Flaws J. Can obesity explain the racial difference in stage of breast cancer at diagnosis between black and white women? *J Womens Health Gend Based Med* 2004:11:527–536.
18. Cui Y, Whiteman M, Flaws J, Langenberg P, Tkaczuk K, Bush T et al. Body mass and stage of breast cancer at diagnosis. *Int J Cancer* 2002;98:279–283.
19. Rosner GL, Hargis JB, Hollis DR, Budman DR, Weiss RB, Henderson IC et al. Relationship between toxicity and obesity in women receiving adjuvant chemotherapy for breast cancer: results from cancer and leukemia group B study 8541. *J Clin Oncol* 1996;14:3000–3008.
20. Colleoni M, Li S, Gelber RD, Price KN, Coates AS, Castiglione-Gertsch M et al. Relation between chemotherapy dose, oestrogen receptor expression, and body-mass index. *Lancet* 2005;366:1108–1110.
21. Peeters A, Barendregt JJ, Willekens F, Mackenbach JP, Al Mamum A, Bonneux L. Obesity in adulthood and its consequences for life expectancy: a life-table analysis. *Ann Intern Med* 2003;138:24–32.
22. Flegal KM, Graubard BI, Williamson DF, Gail MH. Cause-specific excess deaths associated with underweight, overweight, and obesity. *JAMA* 2007;298:2028–2037.

23. Kumanyika SK, Parker L, Sim LJ. Bridging the evidence gap in obesity prevention: a framework to inform decision making. Institute of Medicine. The National Academies Press, Washington, DC, 2010.
24. van Ravesteyn NT, Schechter CB, Near AM, Heijnsdijk EA, Stoto MA, Draisma G et al. Race-specific impact of natural history, mammography screening, and adjuvant treatment on breast cancer mortality rates in the United States. *Cancer Epidemiol Biomarkers Prev* 2011;20:112–122.
25. Levy DT, Mabry PL, Wang YC, Gortmaker S, Huang TT, Marsh T et al. Simulation models of obesity: a review of the literature and implications for research and policy. *Obes Rev* 2011;12:378–394.
26. CISNET Cancer Intervention and Surveillance Modeling Network. http://cisnet.cancer.gov. Accessed May 2012.
27. Tan SY, van Oortmarssen GJ, de Koning HJ, Boer R, Habbema JD. The MISCAN-Fadia continuous tumor growth model for breast cancer. *J Natl Cancer Inst Monogr* 2006;36:56–65.
28. Mandelblatt J, Schechter CB, Lawrence W, Yi B, Cullen J. The SPECTRUM population model of the impact of screening and treatment on US breast cancer trends from 1975 to 2000: principles and practice of the model methods. *J Natl Cancer Inst Monogr* 2006;36:47–55.
29. Mandelblatt JS, Cronin KA, Bailey S, Berry DA, de Koning HJ, Draisma G et al. Effects of mammography screening under different screening schedules: model estimates of potential benefits and harms. *Ann Intern Med* 2009;151:738–747.
30. Carter SB, Gartner SS, Haines MR et al. Historical statistics of the united states, volume one: population. Cambridge University Press, New York, 2006.
31. Holford TR, Cronin KA, Mariotto AB, Feuer EJ. Changing patterns in breast cancer incidence trends. *J Natl Cancer Inst Monogr* 2006;36:19–25.
32. Ravdin PM, Cronin KA, Howlader N et al. The decrease in breast-cancer incidence in 2003 in the United States. *N Engl J Med* 2007;356:1670–1674.
33. Kelsey JL. Breast cancer epidemiology: summary and future directions. *Epidemiol Rev* 1993;15:256–263 .
34. Cronin KA, Yu B, Krapcho M, Miglioretti DL, Fay MP, Izmirlian G et al. Modeling the dissemination of mammography in the United States. *Cancer Causes Control* 2005;16:701–712.
35. Cronin KA, Mariotto AB, Clarke LD, Feuer EJ. Additional common inputs for analyzing impact of adjuvant therapy and mammography on US mortality. *J Natl Cancer Inst Monogr* 2006;36:26–29.
36. Cronin KA, Miglioretti DL, Krapcho M, Yu B, Geller BM, Carney PA et al. Bias associated with self-report of prior screening mammography. *Cancer Epidemiol Biomarkers Prev* 2009;18:1699–1705.
37. Cronin KA, Miglioretti DL, Krapcho M, Yu B, Geller BM, Carney PA, et al CEBP focus on cancer surveillance: bias associated with self-report of prior screening mammography— appendix: modeling mammography screening usage by race and ethnicity. Technical Report #2012-02. http://surveillance.cancer.gov/reports/. Accessed January 2010.
38. Zhu K, Wu H, Jatoi I, Potter J, Shriver C. Body mass index and use of mammography screening in the United States. *Prev Med* 2006;42:381–385.
39. Breast Cancer Surveillance Consortium. Performance Measures for 3,884,059 Screening Mammography Examinations from 1996 to 2007 by Age—based on BCSC data as of 2008. http://breastscreening.cancer.gov/data/performance/screening/2008/perf_age.html. Accessed April 2008.
40. Altekruse SF, Kosary CL, Krapcho M, Neyman N, Aminou R, Waldron W. SEER Cancer Statistics Review, 1975–2007, National Cancer Institute. Bethesda, MD. http://seer.cancer. gov/csr/1975-2007/ based on November 2009 SEER data submission, posted to the SEER web site, 2010. Accessed January 2011.
41. Cust AE, Stocks T, Lukanova A, Lundin E, Hallmans G, Kaaks R et al. The influence of overweight and insulin resistance on breast cancer risk and tumour stage at diagnosis: a prospective study. *Breast Cancer Res Treat* 2009;113:567–576.
42. Bekele BN, Nieto-Barajas LE, Munsell MF. Analysis of partially incomplete tables of breast cancer patients' characteristics with an ordinal variable. *J Stat Theory Practice* 2012 [in press].
43. Mariotto AB, Feuer EJ, Harlan LC, Abrams J. Dissemination of adjuvant multiagent chemotherapy and tamoxifen for breast cancer in the United States using estrogen receptor information: 1975–1999. *J Natl Cancer Inst Monogr*2006;36:7–15.
44. Mariotto A, Feuer EJ, Harlan LC, Wun LM, Johnson KA, Abrams J. Trends in use of adjuvant multi-agent chemotherapy and tamoxifen for breast cancer in the United States: 1975–1999. *J Natl Cancer Inst* 2002;94:1626–1634.
45. Brewster AM, Etzel C, Zhou R, Wong Y, Edge S, Blayney DW et al. The impact of obesity on receipt of adjuvant chemotherapy for breast cancer in the National Comprehensive Cancer Network (NCCN) centers. *Breast Cancer Res Treat* 2011;130:897–904.

46. National Comprehensive Cancer Network. NCCN Clinical Practice Guidelines in oncology v.2 (2008). www.nccn.org/professional/physician_gls/f_guidelines.asp. Accessed September 2009.
47. Early Breast Cancer Trialists' Collaborative Group (EBCTCG). Effects of chemotherapy and hormonal therapy for early breast cancer on recurrence and 15-year survival: an overview of the randomised trials. *Lancet* 2005;365:1687–1717.
48. Early Breast Cancer Trialists' Collaborative Group (EBCTCG). Polychemotherapy for early breast cancer: an overview of the randomised trials. *Lancet* 1998;352:930–942.
49. Early Breast Cancer Trialists' Collaborative Group (EBCTCG). Tamoxifen for early breast cancer: an overview of the randomised trials. *Lancet* 1998;351:1451–1467.
50. Clarke M, Coates AS, Darby SC, Davies C, Gelber RD, Godwin J et al. Adjuvant chemotherapy in oestrogen-receptor-poor breast cancer: patient-level meta-analysis of randomised trials. *Lancet* 2008;371:29–40.
51. Dignam JJ. Efficacy of systemic adjuvant therapy for breast cancer in African-American and Caucasian women. *J Natl Cancer Inst Monogr* 2001;30:36–43.
52. Griggs JJ, Sorbero ME, Lyman GH. Undertreatment of obese women receiving breast cancer chemotherapy. *Arch Intern Med* 2005;165:1267–1273.
53. Griggs JJ, Culakova E, Sorbero ME et al. Effect of patient socioeconomic status and body mass index on the quality of breast cancer adjuvant chemotherapy. *J Clin Oncol* 2007;25:277–284.
54. National Center for Health Statistics, United States, 2000 Public Health Service, Hyattsville. http://www.cdc.gov/nchs/products/life_tables.htm. Accessed May 2012.
55. Human Mortality Database. Centers for Disease Control and Prevention, Publications and Information Products, Life Tables. http://www.mortality.org. Accessed August 2011.
56. Projected population by single year of age, sex, race, and Hispanic origin for the United States: July 1, 2000 to July 1, 2050. Released 2008. http://www.census.gov/population/projections/data/. Accessed May 2012.
57. Parkin DM, Boyd L, Walker LC. The fraction of cancer attributable to lifestyle and environmental factors in the UK in 2010. *Br J Cancer* 2011;105(Suppl 2):S77–S81.
58. Fisher B, Costantino JP, Wickerham DL, Redmond CK, Kavanah M, Cronin WM et al. Tamoxifen for prevention of breast cancer: report of the National Surgical Adjuvant Breast and Bowel Project P-1 Study. *J Natl Cancer Inst* 1998;90:1371–1388.
59. Vogelaar I, Van Ballegooijen M, Schrag D, Boer R, Winawer SJ, Habbema JD et al. How much can current interventions reduce colorectal cancer mortality in the US? Mortality projections for scenarios of risk-factor modification, screening, and treatment. *Cancer* 2006;107:1624–1633.
60. Finkelstein EA, Trogdon JG, Cohen JW, Dietz W. Annual medical spending attributable to obesity: payer-and service-specific estimates. *Health Aff* (Millwood) 2009;28:w822–w831.
61. National Institutes of Health. Clinical guidelines on the identification, evaluation, and treatment of overweight and obesity in adults—the evidence report. *Obes Res* 1998;6(Suppl 2):51S–209S.
62. Fogelholm M, Kukkonen-Harjula K, Nenonen A, Pasanen M. Effects of walking training on weight maintenance after a very-low-energy diet in premenopausal obese women: a randomized controlled trial. *Arch Intern Med* 2000;160:2177–2184.
63. Lu Y, Ma H, Malone KE, Norman SA, Sullivan-Halley J, Strom BL et al. Obesity and survival among black women and white women 35 to 64 years of age at diagnosis with invasive breast cancer. *J Clin Oncol* 2011;29:3358–3365.
64. McCullough ML, Feigelson HS, Diver WR, Patel AV, Thun MJ, Calle EE. Risk factors for fatal breast cancer in African-American women and White women in a large US prospective cohort. *Am J Epidemiol* 2005;162:734–742.
65. Conroy SM, Maskarinec G, Wilkens LR, White KK, Henderson BE, Kolonel LN. Obesity and breast cancer survival in ethnically diverse postmenopausal women: the Multiethnic Cohort Study. *Breast Cancer Res Treat* 2011;129:565–574.
66. Sexton KR, Franzini L, Day RS, Brewster A, Vernon SW, Bondy ML. A review of body size and breast cancer risk in Hispanic and African American women. *Cancer* 2011;117:5271–5281.
67. Protani M, Coory M, Martin JH. Effect of obesity on survival of women with breast cancer: systematic review and meta-analysis. *Breast Cancer Res Treat* 2010;123:627–635.
68. Beasley JM, Kwan ML, Chen WY, Weltzien EK, Kroenke CH, Lu W et al. Meeting the physical activity guidelines and survival after breast cancer: findings from the after breast cancer pooling project. *Breast Cancer Res Treat* 2012;131:637–643.
69. Kwan ML, Chen WY, Kroenke CH, Weltzien EK, Beasley JM, Nechuta SJ et al. Pre-diagnosis body mass index and survival after breast cancer in the After Breast Cancer Pooling Project. *Breast Cancer Res Treat* 2012;132:729–739.

70. Esfahlan RJ, Zarghami N, Esfahlan AJ, Mollazadeh M, Nejati K, Nasiri M. The possible impact of obesity on androgen, progesterone and estrogen receptor (ER-alpha and ER-beta) gene expression in breast cancer patients. *Breast Cancer* (Auckl) 2011;5:227–237.
71. Chlebowski RT. Obesity and breast cancer outcome: adding to the evidence. *J Clin Oncol* 2012;30:126–128 .
72. Goodwin PJ, Ennis M, Pritchard KI, Trudeau ME, Koo J, Taylor SK et al. Insulin- and obesity-related variables in early-stage breast cancer: correlations and time course of prognostic associations. *J Clin Oncol* 2012;30:164–171.
73. Gu JW, Young E, Patterson SG, Makey KL, Wells J, Huang M et al. Postmenopausal obesity promotes tumor angiogenesis and breast cancer progression in mice. *Cancer Biol Ther* 2011;11:910–917.
74. Wang Y, Beydoun MA. The obesity epidemic in the United States—gender, age, socioeconomic, racial/ethnic, and geographic characteristics: a systematic review and meta-regression analysis. *Epidemiol Rev* 2007;29:6–28.
75. Wang YC, Colditz GA, Kuntz KM. Forecasting the obesity epidemic in the aging US population. *Obesity* (Silver Spring) 2007;15:2855–2865.
76. Lee S, Zelen M. A stochastic model for predicting the mortality of breast cancer. *J Natl Cancer Inst Monogr* 2006;36:79–8.

PART 2

Predicting the effects of different screening strategies

Chapter 6

Effects of mammography screening under different screening schedules: model estimates of potential benefits and harms

Jeanne S. Mandelblatt, Kathleen A. Cronin, Stephanie Bailey, Donald A. Berry, Harry J. de Koning, Gerrit Draisma, Hui Huang, Sandra J. Lee, Mark Munsell, Sylvia K. Plevritis, Peter Ravdin, Clyde B. Schechter, Bronislava Sigal, Michael A. Stoto, Natasha K. Stout, Nicolien T. van Ravesteyn, John Venier, Marvin Zelen, Eric J. Feuer; for the Breast Cancer Working Group of the Cancer Intervention and Surveillance Modeling Network (CISNET)

Ann Intern Med. 2009 Nov 17;151(10):738-47

ABSTRACT

Background: Despite trials of mammography and widespread use, optimal screening policy is controversial.

Objective: To evaluate U.S. breast cancer screening strategies.

Design: 6 models using common data elements.

Data Sources: National data on age-specific incidence, competing mortality, mammography characteristics, and treatment effects.

Target Population: A contemporary population cohort.

Time Horizon: Lifetime.

Perspective: Societal.

Interventions: 20 screening strategies with varying initiation and cessation ages applied annually or biennially.

Outcome Measures: Number of mammograms, reduction in deaths from breast cancer or life-years gained (vs. no screening), false-positive results, unnecessary biopsies, and overdiagnosis.

Results of Base Case: The 6 models produced consistent rankings of screening strategies. Screening biennially maintained an average of 81% (range across strategies and models, 67% to 99%) of the benefit of annual screening with almost half the number of false-positive results. Screening biennially from ages 50 to 69 years achieved a median 16.5% (range, 15% to 23%) reduction in breast cancer deaths versus no screening. Initiating biennial screening at age 40 years (vs. 50 years) reduced mortality by an additional
3% (range, 1% to 6%), consumed more resources, and yielded more false-positive results. Biennial screening after age 69 years yielded some additional mortality reduction in all models, but overdiagnosis increased most substantially at older ages.

Sensitivity Analysis Results: Varying test sensitivity or treatment patterns did not change conclusions.

Limitations: Results do not include morbidity from false-positive results, knowledge of earlier diagnosis, or unnecessary treatment.

Conclusion: Biennial screening achieves most of the benefit of annual screening with less harm. Decisions about the best strategy depend on program and individual objectives and the weight placed on benefits, harms and resource considerations.

Primary Funding Source: National Cancer Institute

INTRODUCTION

In 2009, an estimated 193 370 women in the United States will develop invasive breast cancer and about 40 170 of them will die of this disease.[1] Randomized trials of mammography [2-4] have demonstrated reductions in breast cancer mortality associated with screening from ages 50 to 74 years. Trial results for women aged 40 to 49 years and women aged 74 years or older were not conclusive, and the trials [4,5] had some problems with design, conduct, and interpretation. However, it is not feasible to conduct additional trials to get more precise estimates of the mortality benefits from extending screening to women younger than 50 years or older than 74 years or to test different screening schedules.

We developed models of breast cancer incidence and mortality in the United States. These models are ideally suited for estimating the effect of screening under a variety of policies.[6,7] Modeling has the advantage of being able to hold selected conditions (for example, screening intervals or test sensitivity) constant, which facilitates comparison of strategies. Because all models make assumptions about unobservable events, use of several models provides a range of plausible effects and can illustrate the effects of differences in model assumptions.[7]

We used 6 established models to estimate the outcomes across 20 mammography screening strategies that vary by age of initiation and cessation and by screening interval among a cohort of U.S. women. The results are intended to contribute to practice and guideline policy debates.

METHODS

The 6 models were developed independently within the Cancer Intervention and Surveillance Modeling Network (CISNET) of the National Cancer Institute (NCI) [7,8] and were exempt from institutional review board approval. The models have been described elsewhere.[7, 9-15] Briefly, they share common features and inputs but differ in some ways (Appendix Table 6.1, available at www.annals.org). Model E (Erasmus Medical Center, Rotterdam, the Netherlands), model G (Georgetown University Medical Center, Washington, DC, and Albert Einstein College of Medicine, Bronx, New York), model M (M.D. Anderson Cancer Center, Houston, Texas), and model W (University of Wisconsin, Madison, Wisconsin, and Harvard Medical School, Boston, Massachusetts) include ductal carcinoma in situ (DCIS). Models E and W specifically assume that some portions of DCIS are nonprogressive and do not result in death. Model W also assumes that some cases of small invasive cancer are nonprogressive. Model S (Stanford University, Palo Alto, California) and model D (Dana-Farber Cancer Institute, Boston, Massachusetts) include only invasive cancer. Some groups model breast cancer in stages, but 3 (models E, S, and W) use tumor size and tumor growth. The models also differ by whether treatment affects the hazard for death from breast cancer (models G, S, and D), results in a

cure for some fraction of cases (models E and W), or both (model M). Despite these differences, in previous collaborations [7] all the models came to similar qualitative estimates of the relative contributions of screening and treatment to observed decreases in deaths from breast cancer.

Model overview

We used the 6 models to estimate the benefits, resource use (as measured by number of mammograms), and harms of 20 alternative screening strategies varying by starting and stopping age and by interval (annual and biennial) (Table 6.1). The models begin with estimates of breast cancer incidence and mortality trends without screening and treatment and then overlay screening use and improvements in survival associated with treatment.[7] We use a cohort of women born in 1960 and follow them beginning at age 25 years for their entire lives. Breast cancer is generally depicted as having a preclinical, screening-detectable period (sojourn time) and a clinical detection point. On the basis of mammography sensitivity (or thresholds of detection), screening identifies disease in the preclinical screening-detection period and results in the identification of earlier-stage or smaller tumors than might be identified by clinical detection, resulting in reduction in breast cancer mortality. Age, estrogen receptor status, and tumor size– or stage–specific treatment have independent effects on mortality. Women can die of breast cancer or of other causes.

Table 6.1. Breast cancer screening strategies*

No screening
Screen from age 40 to 69 y
Screen from age 40 to 79 y
Screen from age 40 to 84 y
Screen from age 45 to-69 y
Screen from age 50 to 69 y
Screen from age 50 to 74 y
Screen from age 50 to 79 y
Screen from age 50 to 84 y
Screen from age 55 to 69 y
Screen from age 60 to 69 y

* Each strategy was evaluated by using an annual or biennial schedule, for a total of 20 strategies; we include no screening for comparison.

Model data variables

All 6 modeling groups use a common set of age-specific variables for breast cancer incidence, mammography test characteristics, treatment algorithms and effects, and nonbreast cancer competing causes of death (Appendix Table 6.2, available at www.annals.org). In addition to these common variables, each model includes model-specific inputs (or intermediate outputs) to represent preclinical detectable times, lead time, dwell time within stages of disease, and stage distribution in unscreened versus screened women on the basis of their specific model structure.[7, 9-15]

We use an age–period– cohort model to estimate what breast cancer incidence rates would have been without screening.[16] This approach considers the effect of age, temporal trends in risk by cohort, and time period. Because we do not have data on future incidence of breast cancer, we extrapolate forward assuming that future age-specific incidence increases as women age, as observed in 2000. To isolate the effect of technical effectiveness of screening and to assess the effect of screening on mortality while holding treatment constant, models assume 100% adherence to screening and indicated treatment.

Three groups use the age-specific mammography sensitivity (and specificity) values observed in the Breast Cancer Surveillance Consortium (BCSC) program for detection of all cases of breast cancer (invasive and in situ). Separate values are used for initial and subsequent mammography performed at either annual or biennial intervals.[17] Two of the models (D and G) use these data directly as input variables,[10,14] and 1 model (S) uses the data to calibrate the model.[13] The other 3 models (E, M, and W) use the BCSC data as a guide and to fit sensitivity estimates from this and other sources.[9,11,15]

All women who have estrogen receptor–positive invasive tumors receive hormonal treatment (tamoxifen if women aged <50 years at diagnosis and anastrozole if ≥50 years) and nonhormonal treatment with an anthracyclinebased regimen. Women with estrogen receptor–negative invasive tumors receive nonhormonal therapy only. Women with DCIS who have estrogen receptor–positive tumors receive hormonal therapy only.[18] Treatment effectiveness is based on a synthesis of recent clinical trials and is modeled as a proportionate reduction in mortality risk or the proportion cured.[19,20]

Benefits

We estimated the cumulative probability of unscreened women dying of breast cancer from age 40 years to death. Screening benefit is then calculated as the percentage of reduction in breast cancer mortality (vs. no screening). We also examined life-years gained because of averted or delayed breast cancer death. Benefits are cumulated over the lifetime of the cohort to capture reductions in breast cancer mortality (or life-years gained) occurring years after the start of screening, after considering nonbreast cancer mortality.[21,22]

Harms

As measures of the burden that a regular screening program imposes on a population, 3 different potential screening harms were examined: false-positive mammograms, unnecessary biopsies, and overdiagnosis. We define the rate of false-positive mammograms as the number of mammograms read as abnormal or needing further follow-up in women without cancer divided by the total number of positive screening mammograms based on the specificity reported in the BCSC.[17] We define unnecessary biopsies post hoc as the proportion of women with false-positive screening results who receive a biopsy.[23] We define overdiagnosis as the proportion of cases in each strategy that would not have clinically surfaced in a woman's lifetime (because of lack of progressive potential or death from another cause) among all cases arising from age 40 years onward.

Base-case analysis

We compared model results for the 20 strategies to select the most efficient approach. In a decision analysis, we considered a new intervention more efficient than a comparison intervention if it results in gains in health outcomes, such as life-years gained or deaths averted, while consuming fewer resources (or costs). If the new intervention results in worse outcomes and requires a greater investment, it is inefficient and would not be considered for further use. In economic analysis, inefficient strategies are said to be "dominated" when this occurs. To rank the screening strategies, we first look at the results of each model independently. For a particular model, a strategy that requires more mammographies (our measure of resource use) but has a lower relative percentage of mortality reduction (or life-years gained) is considered inefficient or dominated by other strategies. To evaluate strategies on the basis of results from all 6 models together, we classify them as follows: If a strategy is dominated in all or in 5 of 6 of the models, we considered it dominated overall. If a strategy is not dominated in any of the models, we classified it as efficient. For a strategy with mixed results across the models, we classified it as borderline.

After all dominated strategies were eliminated, the remaining strategies were represented as points on a graph plotting the average number of mammograms versus the percentage of mortality reduction (or life-years gained) for each model. We obtained the efficiency frontier for each graph by identifying the sequence of points that represent the largest incremental gain in percentage of mortality reduction (or life-years gained) per additional screening mammography. Screening strategies that fall on this frontier are the most efficient (that is, no alternative exists that provides more benefit for fewer mammographies performed).

Sensitivity analysis

We conducted a sensitivity analysis to see whether our conclusions about the ranking of strategies change when we vary input variables. First, we investigate the effect of assuming that mammography sensitivity for a given age, screening round, and screening interval is 10 percentage points less than that observed. Second, we examine whether ranking of strate-

gies varies if treatment includes newer hormonal and nonhormonal adjuvant regimens (for example, taxanes). Third, because adjuvant therapy is unlikely to reach 100% of women as modeled in our base-case analysis, we reassess the ranking of strategies if we assume that actual observed current treatment patterns apply to the cohort.[24]

Model validation and uncertainty

Each model has a different structure and assumptions and some varying input variables, so no single method can be used to validate results against an external gold standard. For instance, because some models used results from screening trials (or SEER [Surveillance, Epidemiology and End Results] data) for calibration or as input variables, we cannot use comparisons of projected mortality reductions to trial results to validate all of the models. In addition, we cannot directly compare the results of this analysis, which uses 100% actual screening for all women at specified intervals, with screening trial results in which invitation to screening and participation varied. In our previous work, [7, 9-11, 13-15] results of each model accurately projected independently estimated trends in the absence of intervention and closely approximated modern stage distributions and observed mortality trends. Overall, using 6 models to project a range of plausible screening outcomes provides implicit cross-validation, with the range of results from the models as a measure of uncertainty.

Role of the funding source

This work was done under contracts from the Agency for Healthcare Research and Quality (AHRQ) and NCI and grants from the NCI. Staff from the NCI provided some data and technical assistance, and AHRQ staff reviewed the manuscript. Model results are the sole responsibility of the investigators.

RESULTS

In an unscreened population, the models predict a cumulative probability of breast cancer developing over a woman's lifetime starting at age 40 years ranging from 12% to 15%. Without screening, the median probability of dying of breast cancer after age 40 years is 3.0% across the 6 models. Thus, if a particular screening strategy leads to a 10% reduction in breast cancer mortality, then the probability of breast cancer mortality would be reduced from 3.0% to 2.7%, or 3 deaths averted per 1000 women screened.

Benefits

The 6 models produce consistent results on the ranking of the strategies (Appendix Table 6.3, available at www.annals.org). Eight approaches are "efficient" in all models (that is, not dominated, because they provide additional mortality reductions for added use of mam-

mography); 7 of these have a biennial interval, and all but 2 start at age 50 years. Figure 6.1 shows these results, and again we see that most strategies on the efficiency frontier have a biennial interval. Screening every other year from ages 50 to 69 years is an efficient strategy for reducing breast cancer mortality in all models. In all models, biennial screening starting at age 50 years and continuing through ages 74, 79, or 84 years are of fairly similar efficiency.

In examining benefits in terms of life-years gained (Appendix Table 6.4, available at www.annals.org), 6 of the 8 consistently nondominated strategies have a biennial interval. In contrast to results for mortality reduction, half of the nondominated strategies include screening initiation at age 40 years. Annual screening strategies that include screening until age 79 or 84 years are on the efficiency frontier (Appendix Figure 6.1, available at www.annals.org), but are less resource-efficient than biennial approaches for increasing life-years gained.

As another way to examine the effect of screening interval, we calculated for each screening strategy and model the proportion of the annual benefit (in terms of mortality reduction) that could be achieved by biennial screening (Table 6.2). Biennial screening maintains an average of 81% (range across strategies and models, 67% to 99%) of the benefits achieved by annual screening.

Table 6.2. Percentage of reduction in breast cancer mortality maintained when moving from an annual screening interval to a biennial interval, by screening strategy and model

Model*	Maintained Reduction in Breast Cancer Mortality, by Screening Strategy, %†									
	Ages 50-69 y	Ages 40-69 y	Ages 45-69 y	Ages 40-79 y	Ages 40-84 y	Ages 55-69 y	Ages 60-69 y	Ages 50-74 y	Ages 50-79 y	Ages 50-84 y
D	76	75	78	79	82	83	79	81	78	83
E	75	73	74	75	75	75	73	76	75	76
G	85	86	91	87	88	91	86	89	88	89
M	90	96	97	97	99	92	84	95	93	95
S	74	73	78	76	77	80	74	79	85	79
W	68	67	70	70	71	71	70	72	70	73

* Model group abbreviations: D = Dana-Farber Cancer Institute; E = Erasmus Medical Center; G = Georgetown University; M = M.D. Anderson Cancer Center; S =Stanford University; W University of Wisconsin/Harvard.

† Differences in the range of results reflect differences in modeling approaches. For example, the benefit of screening in model M is modeled through stage shift, as with most other models, but also includes a "beyond stage shift" factor based on a cure fraction for small tumors. However, because many of these "cures" occur among women with invasive cancer that is not fatal, finding such cancer 1 year earlier confers very little mortality advantage to annual (vs. biennial) screening.

We also examined the incremental benefits gained by extending screening from ages 50 to 69 years to either earlier or later ages of initiation and cessation (Table 6.3). Continuing screening to age 79 years (vs. 69 years) results in a median increase in percentage of mortality reduction of 8% (range, 7% to 11%) and 7% (range, 6% to 10%) under annual and biennial intervals, respectively. If screening begins at age 40 years (vs. 50 years) and continues to age 69 years, all models project additional, albeit small, reductions in breast cancer mortality (3%

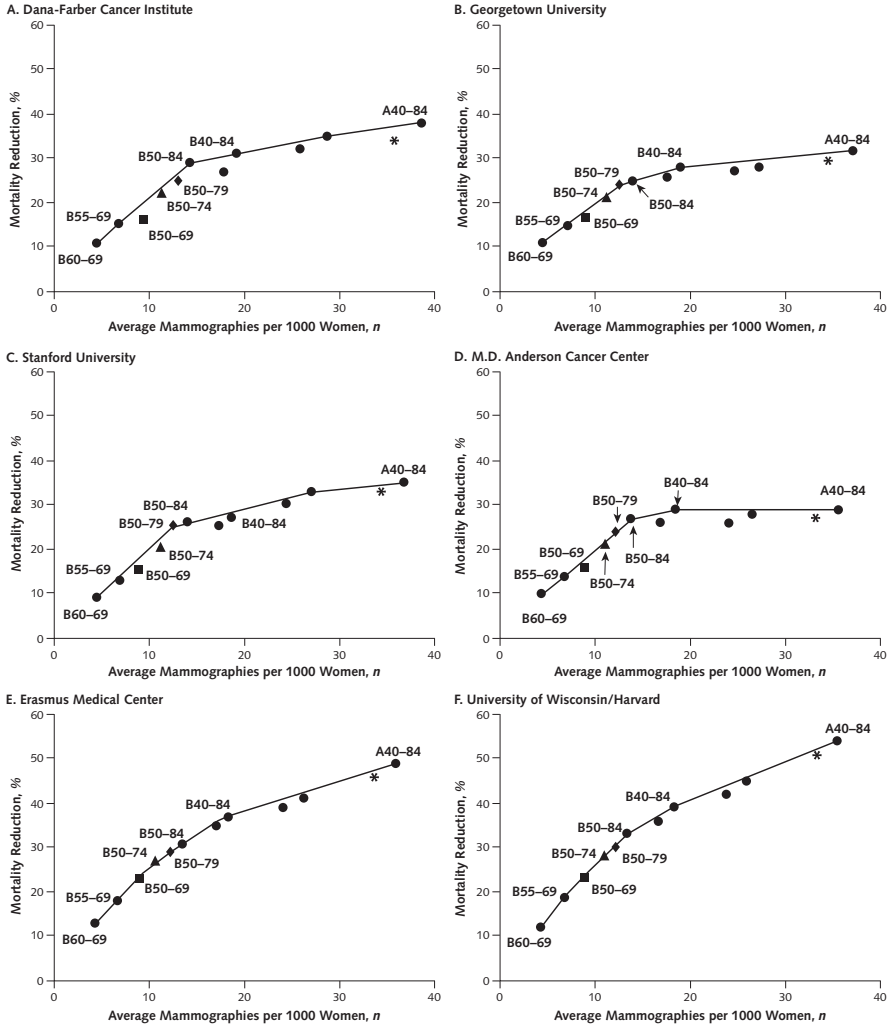


Figure 6.1. Percentage of breast cancer mortality reduction versus number of mammographies performed per 1000 women, by model and screening strategy.

The panels show an efficiency frontier graph for each model. The graph plots the average number of mammographies performed per 1000 women against the percentage of mortality reduction for each screening strategy (vs. no screening). Strategies are denoted as annual (A) or biennial (B) with starting and stopping ages. We plot efficient strategies (that is, those in which increases in use of mammography resources result in greater mortality reduction than the next least-intensive strategy) in all 6 models. We also plot "borderline" strategies (approaches that are efficient in some models but not others). The line between strategies represents the "efficiency frontier." Strategies on this line would be considered efficient because they achieve the greatest gain per use of mammography resources compared with the point (or strategy) immediately below it. Points that fall below the line are not considered as efficient as those on the line. When the slope in the efficiency frontier plot levels off, the additional reductions in mortality per unit increase in use of mammography are small relative to the previous strategies and could indicate a point at which additional investment (use of screening) might be considered as having a low return (benefit).

Table 6.3. Incremental changes in percentage of reduction in breast cancer mortality and life-years gained per 1000 women, by age of screening initiation and cessation

Model*	Start at Age 40 y vs. 50 y†						Stop at Age 79 y vs. 69 y‡					
	Difference in Percentage of Reduction in Breast Cancer Mortality		Difference in Breast Cancer Deaths Averted per 1000 Women		Difference in Life-Years Gained per 1000 Women		Difference in Percentage of Reduction in Breast Cancer Mortality		Difference in Breast Cancer Deaths Averted per 1000 Women		Difference in Life-Years Gained per 1000 Women	
	Annual	Biennial	Annual	Biennial	Annual	Biennial	Annual	Biennial	Annual	Biennial	Annual	Biennial
D	3	2	1	1	25	20	11	9	3	3	28	26
E	8	5	2	1	58	40	8	6	2	2	18	15
G	3	3	1	1	34	29	7	7	2	2	27	25
M	2	3	1	1	11	18	7	7	2	2	21	21
S	2	1	1	1	32	21	10	10	4	4	38	31
W	10	6	2	1	57	37	8	6	2	1	19	15
Median	3	3	1	1	33	25	8	7	2	2	24	23.5

* Model group abbreviations: D = Dana-Farber Cancer Institute; E = Erasmus Medical Center; G = Georgetown University; M = M.D. Anderson Cancer Center; S = Stanford University; W= University of Wisconsin/Harvard.
† Incremental difference between screening from 40 to 69 y versus 50 to 69 y.
‡ Incremental difference between screening from 50 to 79 y versus 50 to 69 y.

median reduction with either annual or biennial intervals) (Table 6.3). This translates into a median of 1 additional breast cancer death averted (range, 1 to 2 deaths) per 1000 women screened under a strategy of annual screening from age 40 to 69 years (vs. 50 to 69 years). Thus, greater mortality reductions could be achieved by stopping screening at an older age than by initiating screening at an earlier age.

However, when life-years gained is the outcome measure, 3 of the models conclude that benefits are greater from extending screening to the younger rather than the older age group (Table 6.3). For instance, starting annual screening at age 40 years (vs. 50 years) and continuing annually to age 69 years yields a median of 33 (range, 11 to 58) life-years gained per 1000 women screened, whereas extending annual screening to age 79 years (vs. 69 years) yields a median of only 24 (range, 18 to 38) life-years gained per 1000 women screened.

Harms

All the models project similar rates of false-positive mammograms over the lifetime of screened women across the screening strategies; Table 6.4 summarizes results for an exemplar model. More false-positive results occur in strategies that include screening from ages 40 to 49 years than in those that initiate screening at age 50 years or later and those that include annual screening rather than biennial screening. For instance, annual screening from ages 40 to 69 years yields 2250 false-positive results for every 1000 women screened over this period, almost twice as many as that of biennial screening in this age group. The proportion of biopsies that occur because of these false-positive results that are retrospectively deemed

Table 6.4. Benefits and harms comparison of different starting and stopping ages using the exemplar model*

Strategy	Average Screenings per 1000 Women	Potential Benefits (vs. No Screening)			Potential Harms (vs. No Screening)†	
		Percentage of Mortality Reduction	Cancer Deaths Averted per 1000 Women	Life-Years Gained per 1000 Women	False-Positive Results per 1000 Women	Unnecessary Biopsies per 1000 Women
Comparison of different starting ages						
Biennial screening						
40–69 y	13865	16‡	6.1	120‡	1250	88
45–69 y	11771	17‡	6.2	116‡	1050	74
50–69 y	8944	15	5.4	99	780	55
55–69 y	6941	13	4.9	80	590	41
60–69 y	4246	9	3.4	52	340	24
Annual screening						
40–69 y	27583	22‡	8.3	164‡	2250	158
45–69 y	22623	22‡	8	152‡	1800	126
50–69 y	17759	20‡	7.3	132‡	1350	95
55–69 y	13003	16‡	6.1	102‡	950	67
60–69 y	8406	12‡	4.6	69‡	600	42
Comparison of different stopping ages						
Biennial screening						
50–69 y	8944	15	5.4	99	780	55
50–74 y	11109	20	7.5	121	940	66
50–79 y	12347	25	9.4	130	1020	71
50–84 y	13836	26	9.6	138	1130	79
Annual screening						
50–69 y	17759	20‡	7.3	132‡	1350	95
50–74 y	21357	26‡	9.5	156‡	1570	110
50–79 y	24439	30	11.1	170	1740	122
50–84 y	26913	33	12.2	178	1880	132

* Results are from model S (Stanford University). Model S was chosen as an exemplar model to summarize the balance of benefits and harms associated with screening 1000 women under a particular screening strategy.

† Overdiagnosis is another significant harm associated with screening. However, given the uncertainty in the knowledge base about ductal carcinoma in situ and small invasive tumors, we felt that the absolute estimates are not reliable. In general, overdiagnosis increases with age across all age groups but increases more sharply for women who are screened in their 70s and 80s.

‡ Strategy is dominated by other strategies; the strategy that dominates may not be in this table.

unnecessary (that is, the woman did not have cancer) is about 7%; therefore, many more women will undergo unnecessary biopsies under annual screening than biennial screening.

Of the 6 models, 5 estimated rates of overdiagnosis. They showed an increase in the risk for overdiagnosis as age increases (data not shown). Although the increase with age occurs over the entire age range considered in the different screening strategies, the rate of increase accelerates in the older age groups, mostly because of increasing rates of competing causes of mortality. Rates of overdiagnosis were higher for DCIS than for invasive disease, proportionately affecting younger women more because more cases of DCIS are diagnosed at younger ages. However, overall, initiating screening at age 40 years (vs. 50 years) had a smaller effect on overdiagnosis than did extending screening beyond age 69 years. Biennial strategies decrease the rate of overdiagnosis, but by much less than one half. The absolute estimate of overdiagnosis varied between models depending on whether DCIS was or was not included and on the assumptions related to progression of DCIS and invasive disease, reflecting the uncertainty in the current knowledge base.

Sensitivity analysis

The overall conclusions are robust across the 6 models under different assumptions about mammography sensitivity, treatment patterns, and treatment effectiveness (data not shown).

DISCUSSION

This study uses 6 established models that use common inputs but different approaches and assumptions to extend previous randomized mammography screening trial results to the U.S. population and to age groups in whom trial results are less conclusive. All 6 modeling groups concluded that the most efficient screening strategies are those that include a biennial screening interval. Conclusions about the optimal starting ages for screening depend more on the measure chosen for evaluating outcomes. If the goal of a national screening program is to reduce mortality in the most efficient manner, then programs that screen biennially from age 50 years to age 69, 74, or 79 years are among the most efficient on the basis of the ratio of benefits to the number of screening examinations. If the goal of a screening program is to efficiently maximize the number of life-years gained, then the preferred strategy would be to screen biennially starting at age 40 years. Decisions about the best starting and stopping ages also depend on tolerance for false-positive results and rates of overdiagnosis.

The conclusion of this modeling analysis—that biennial intervals are more efficient and provide a better balance of benefits and harms than annual intervals—is contrary to some current practices in the United States.[25-27] However, our result that biennial screening is more efficient than annual screening is consistent with previous modeling research [28-32] and screening trials, most of which used 2-year intervals.[2-5] The model results also agree with reports

showing similar intermediate cancer outcomes (for example, stage distribution) between programs using annual and biennial screening, especially among women aged 50 years or older.[33-37] In addition, we demonstrated substantial increases in false-positive results and unnecessary biopsies associated with annual intervals, and these harms are reduced by almost 50% with biennial intervals. Our results are also consistent with current knowledge of disease biology. Slow-growing tumors are much more common than fast-growing tumors, and the ratio of slow- to fast-growing tumors increases with age,[38] so that little survival benefit is lost between screening every year versus every other year. For the small subset of women with aggressive, fast-growing tumors, even annual screening is not likely to confer a survival advantage. Guidelines in other countries [4] include biennial screening. However, whether it will be practical or acceptable to change the existing U.S. practice of annual screening cannot be addressed by our models.

In all models, some reductions in breast cancer mortality, albeit small, were seen with strategies that started screening at age 40 years versus 50 years. Because models can represent millions of observations, they are well-suited to detect small differences in a group over time that might not be seen in even the largest clinical trial with a 10- to 15-year follow-up.[4,39-42] If program benefits are measured in life-years, the measure most commonly used in cost-effectiveness analysis, then our results suggest that initiating screening at age 40 years saves more life-years than extending screening past age 69 years (albeit at the cost of increasing the number of false-positive mammograms).

Previous recommendations on breast cancer screening have suggested an upper age limit for screening cessation because of decreasing program efficiency due to competing mortality.[26,43] Our result that screening strategies that include an upper age limit beyond age 69 years remain on the efficiency frontier (albeit with low incremental gains over strategies that stop screening at earlier ages and with greater harms) is consistent with previously reported results of screening benefit from observational and modeled data.[31,32,44-47] However, the observational data reports may have been confounded by the inability to capture lead time and length biases.[48-50] Any benefits of screening older women must be balanced against possible harms. For instance, the probability of overdiagnosis increases with age and increases more dramatically for the oldest age groups. Model estimates for the oldest age groups also have more uncertainty compared with estimates for ages 50 to 74 years because of the lack of primary data on natural history of breast cancer and the absence of screening trial data after age 74 years. With the demographic pressure of an aging society, more research will be needed to fully understand the natural history of this disease and the balance of risks and benefits of screening and treatment in the older age groups.[38,50]

Our results also highlight the need for better primary data on the natural history of DCIS and small invasive cancer to draw reliable conclusions on the absolute magnitude of overdiagnosis associated with different screening schedules.[37,51] Clinical investigation,[52] follow-up

in screening trials,[53] epidemiologic trends in incidence,[54] and previous modeling efforts [9,55] all indicated that some DCIS cases will not progress,[56,57] but how many is not known.

The collaboration of 6 groups with different modeling philosophies and approaches to estimate the same end points by using a common set of data provides an excellent opportunity to cross-replicate data generated from modeling, represent uncertainty related to modeling assumptions and structure, and give insight into which results are consistent across modeling approaches and which are dependent on model assumptions. The resulting conclusions about the ranking of screening strategies were very robust and should provide greater credibility than inferences based on 1 model alone.

Despite our consistent results, our study had some limitations.[58] First, our models provide estimates of the average benefits and harms expected across a cohort of women and do not reflect personal data for individual women. Also, although our models project mortality reductions similar to those observed in clinical trials, the range of results includes higher mortality reductions than that achieved in the trials because we model lifetime screening and assume adherence to all screening and treatment. The trials followed women for limited numbers of years and have some nonadherence. The models also do not capture differences in outcomes among certain risk subgroups, such as women with BRCA1 or BRCA2 genetic susceptibility mutations, women who are healthier or sicker than average, or black women who seem to have more disease at younger ages than white women.[59]

Second, the outcomes considered do not capture morbidity associated with surgery for screening-detected disease [60] or decrements in quality of life associated with false-positive results, living with earlier knowledge of a cancer diagnosis, or overdiagnosis.[61]

Third, in estimating lifetime results, we projected breast cancer trends from background incidence rates of a 1960 birth cohort extrapolated forward in time. However, future background incidence (and mortality) may change as the result of several different forces, such as changes in patterns of reproduction; less use of hormone replacement therapy after 2002 or prescription of tamoxifen or other agents for primary disease prevention; increasing rates of obesity; and further advances in treatment (for example, trastuzumab).[62] Although most models portray known differences in biology by age (for example, distribution of estrogen receptor–positive tumors, sensitivity of screening, and length of the preclinical sojourn times), some aspects of the natural history of disease are not known or cannot be fully captured.

We assumed 100% adherence to screening and treatment to evaluate program efficacy. Benefits will always fall short of the projected results because adherence is not perfect. If actual adherence varies systematically by age or other factors, the ranking of strategies could change. In addition, we did not consider "mixed" strategies (for example, screening annually from age 40 to 49 years and then biennially from age 50 to 79 years) as was done in some trials [5] and other analyses.[36,63] We found that the benefits of screening from ages 40 to 49 years were small. Benefits in this age group were also associated with harms in terms of false-positive results and unnecessary biopsies. Thus, although strategies that include annual

screening from ages 40 to 49 years might be efficient, this would be largely driven by the more favorable balance of benefits and harms after age 50 years. In addition, we judged that mixed strategies are very difficult to communicate to consumers and implement in public health practice.

Finally, we did not discount benefits or include costs in our analysis, although the average number of mammograms per woman (and false-positive results) provides some proxy of resource consumption. Even with these acknowledged limitations, the models demonstrate meaningful, qualitatively similar outcomes despite variations in structure and assumptions.

Overall, the evaluation of screening strategies by the 6 models suggests that optimal program design is based on biennial intervals. Choices about optimal ages of initiation and cessation will ultimately depend on program goals, resources, weight attached to the presence of trial data, the balance of harms and benefits, and considerations of efficiency and equity.

Acknowledgement

The authors thank the BCSC investigators, participating mammography facilities, and radiologists for the data they provided that were used to inform some of our model data input variables. A list of the BCSC investigators and procedures for requesting BCSC data for research purposes is at http://breastscreening.cancer.gov/. The authors also thank Mary Barton, MD, MPP, and William Lawrence, MD, MSc, from AHRQ; members of the U.S. Preventive Services Task Force; the Oregon Evidence-based Practice Center; Ann Zauber, PhD; and Karla Kerlikowske, MD, for helpful comments and review of earlier versions of this article. The authors thank Jackie Ford and Aimee Near for manuscript preparation.

REFERENCES

1. Jemal A, Siegel R, Ward E, Hao Y, Xu J, Thun MJ. Cancer statistics, 2009. *CA Cancer J Clin* 2009;59:225-49.
2. Nyström L, Andersson I, Bjurstam N, Frisell J, Nordenskjöld B, Rutqvist LE. Long-term effects of mammography screening: updated overview of the Swedish randomised trials. *Lancet* 2002;359:909-19.
3. Tabár L, Vitak B, Chen HH, Duffy SW, Yen MF, Chiang CF, et al. The Swedish Two-County Trial twenty years later. Updated mortality results and new insights from long-term follow-up. *Radiol Clin North Am* 2000;38:625-51.
4. Vainio H, Bianchini F, eds. Breast Cancer Screening. International Agency for Research on Cancer Handbook on Cancer Prevention, Report No. 7. Lyon, France: International Agency for Research on Cancer; 2002.
5. Moss SM, Cuckle H, Evans A, Johns L, Waller M, Bobrow L; Trial Management Group. Effect of mammographic screening from age 40 years on breast cancer mortality at 10 years' follow-up: a randomised controlled trial. *Lancet* 2006;368:2053-60.
6. Mandelblatt JS, Fryback DG, Weinstein MC, Russell LB, Gold MR. Assessing the effectiveness of health interventions for cost-effectiveness analysis. Panel on Cost-Effectiveness in Health and Medicine. *J Gen Intern Med* 1997;12: 551-8.
7. Berry DA, Cronin KA, Plevritis SK, Fryback DG, Clarke L, Zelen M, et al. Effect of screening and adjuvant therapy on mortality from breast cancer. *N Engl J Med* 2005;353:1784-92.
8. Cancer Intervention and Surveillance Modeling Network. Accessed at http://cisnet.cancer.gov /breast/profiles. html on 15 September 2008.
9. Fryback DG, Stout NK, Rosenberg MA, Trentham-Dietz A, Kuruchittham V, Remington PL. The Wisconsin Breast Cancer Epidemiology Simulation Model. *J Natl Cancer Inst Monogr* 2006:37-47.
10. Mandelblatt J, Schechter CB, Lawrence W, Yi B, Cullen J. The SPECTRUM population model of the impact of screening and treatment on U.S. breast cancer trends from 1975 to 2000: principles and practice of the model methods. *J Natl Cancer Inst Monogr* 2006:47-55.
11. Berry DA, Inoue L, Shen Y, Venier J, Cohen D, Bondy M, et al. Modeling the impact of treatment and screening on U.S. breast cancer mortality: a Bayesian approach. *J Natl Cancer Inst Monogr* 2006:30-6.
12. Clarke LD, Plevritis SK, Boer R, Cronin KA, Feuer EJ. A comparative review of CISNET breast models used to analyze U.S. breast cancer incidence and mortality trends. *J Natl Cancer Inst Monogr* 2006:96-105.
13. Plevritis SK, Sigal BM, Salzman P, Rosenberg J, Glynn P. A stochastic simulation model of U.S. breast cancer mortality trends from 1975 to 2000. *J Natl Cancer Inst Monogr* 2006:86-95.
14. Lee S, Zelen M. A stochastic model for predicting the mortality of breast cancer. *J Natl Cancer Inst Monogr* 2006:79-86.
15. Tan SY, van Oortmarssen GJ, de Koning HJ, Boer R, Habbema JD. The MISCAN-Fadia continuous tumor growth model for breast cancer. *J Natl Cancer Inst Monogr* 2006:56-65.
16. Holford TR, Cronin KA, Mariotto AB, Feuer EJ. Changing patterns in breast cancer incidence trends. *J Natl Cancer Inst Monogr* 2006:19-25.
17. Breast Cancer Surveillance Consortium. Performance Measures for 3,884,059 Screening Mammography Examinations from 1996 to 2007 by Age & Time (Months) Since Previous Mammography. Accessed at http:// breastscreening.cancer.gov/data/performance/screening/ perf_age_time.html on 7 October 2009.
18. National Comprehensive Cancer Network. NCCN Clinical Practice guidelines in oncology v.2.2008. Accessed at www.nccn.org/professionals/physician_gls /f_guidelines.asp on 22 September 2009.
19. Clarke M, Coates AS, Darby SC, Davies C, Gelber RD, Godwin J, et al; Early Breast Cancer Trialists' Collaborative Group (EBCTCG). Adjuvant chemotherapy in oestrogen-receptor-poor breast cancer: patient-level meta-analysis of randomised trials. *Lancet* 2008;371:29-40.
20. Early Breast Cancer Trialists' Collaborative Group (EBCTCG). Effects of chemotherapy and hormonal therapy for early breast cancer on recurrence and 15-year survival: an overview of the randomised trials. *Lancet* 2005;365:1687-717.
21. Rosenberg MA. Competing risks to breast cancer mortality. *J Natl Cancer Inst Monogr* 2006:15-9.
22. Cronin KA, Feuer EJ, Clarke LD, Plevritis SK. Impact of adjuvant therapy and mammography on U.S. mortality from 1975 to 2000: comparison of mortality results from the CISNET breast cancer base case analysis. *J Natl Cancer Inst Monogr* 2006:112-21.
23. Rosenberg RD, Yankaskas BC, Abraham LA, Sickles EA, Lehman CD, Geller BM, et al. Performance benchmarks for screening mammography. *Radiology* 2006;241:55-66.

24. Mariotto AB, Feuer EJ, Harlan LC, Abrams J. Dissemination of adjuvant multiagent chemotherapy and tamoxifen for breast cancer in the United States using estrogen receptor information: 1975-1999. *J Natl Cancer Inst Monogr* 2006:7-15.
25. Smith RA, Saslow D, Sawyer KA, Burke W, Costanza ME, Evans WP 3rd, et al; American Cancer Society High-Risk Work Group. American Cancer Society guidelines for breast cancer screening: update 2003. *CA Cancer J Clin* 2003;53:141-69.
26. National Cancer Institute. NCI Statement on Mammography Screening [press release]. Bethesda, MD: National Cancer Institute; 31 January 2002. Accessed at www.cancer.gov/newscenter/mammstatement31jan02 on 22 September 2009.
27. Preventive Services: Breast Cancer Screening. Accessed at www.medicare.gov/Health/Mammography.asp on 22 September 2009.
28. Salzmann P, Kerlikowske K, Phillips K. Cost-effectiveness of extending screening mammography guidelines to include women 40 to 49 years of age. *Ann Intern Med* 1997;127:955-65.
29. Stout NK, Rosenberg MA, Trentham-Dietz A, Smith MA, Robinson SM, Fryback DG. Retrospective cost-effectiveness analysis of screening mammography. *J Natl Cancer Inst* 2006;98:774-82.
30. Lee S, Huang H, Zelen M. Early detection of disease and scheduling of screening examinations. *Stat Methods Med Res* 2004;13:443-56.
31. Mandelblatt JS, Schechter CB, Yabroff KR, Lawrence W, Dignam J, Extermann M, et al. Toward optimal screening strategies for older women. Costs, benefits, and harms of breast cancer screening by age, biology, and health status. *J Gen Intern Med* 2005;20:487-96.
32. Kerlikowske K, Salzmann P, Phillips KA, Cauley JA, Cummings SR. Continuing screening mammography in women aged 70 to 79 years: impact on life expectancy and cost-effectiveness. *JAMA* 1999;282:2156-63.
33. Hofvind S, Vacek PM, Skelly J, Weaver DL, Geller BM. Comparing screening mammography for early breast cancer detection in Vermont and Norway. *J Natl Cancer Inst* 2008;100:1082-91.
34. Smith-Bindman R, Chu PW, Miglioretti DL, Sickles EA, Blanks R, Ballard-Barbash R, et al. Comparison of screening mammography in the United States and the United kingdom. *JAMA* 2003;290:2129-37.
35. Smith-Bindman R, Ballard-Barbash R, Miglioretti DL, Patnick J, Kerlikowske K. Comparing the performance of mammography screening in the USA and the UK. *J Med Screen* 2005;12:50-4.
36. White E, Miglioretti DL, Yankaskas BC, Geller BM, Rosenberg RD, Kerlikowske K, et al. Biennial versus annual mammography and the risk of late-stage breast cancer. *J Natl Cancer Inst* 2004;96:1832-9.
37. Wai ES, D'yachkova Y, Olivotto IA, Tyldesley S, Phillips N, Warren LJ, et al. Comparison of 1- and 2-year screening intervals for women undergoing screening mammography. *Br J Cancer* 2005;92:961-6.
38. Fracheboud J, Groenewoud JH, Boer R, Draisma G, de Bruijn AE, Verbeek AL, et al. Seventy-five years is an appropriate upper age limit for population-based mammography screening. *Int J Cancer* 2006;118:2020-5.
39. Miller AB, To T, Baines CJ, Wall C. The Canadian National Breast Screening Study-1: breast cancer mortality after 11 to 16 years of follow-up. A randomized screening trial of mammography in women age 40 to 49 years. *Ann Intern Med* 2002;137:305-12.
40. Elmore JG, Armstrong K, Lehman CD, Fletcher SW. Screening for breast cancer. *JAMA* 2005;293:1245-56.
41. Elmore JG, Reisch LM, Barton MB, Barlow WE, Rolnick S, Harris EL, et al. Efficacy of breast cancer screening in the community according to risk level. *J Natl Cancer Inst* 2005;97:1035-43.
42. Norman SA, Russell Localio A, Weber AL, Coates RJ, Zhou L, Bernstein L, et al. Protection of mammography screening against death from breast cancer in women aged 40-64 years. *Cancer Causes Control* 2007;18:909-18.
43. U.S. Preventive Services Task Force. Screening for breast cancer: recommendations and rationale. *Ann Intern Med* 2002;137:344-6.
44. McCarthy EP, Burns RB, Freund KM, Ash AS, Shwartz M, Marwill SL, et al. Mammography use, breast cancer stage at diagnosis, and survival among older women. *J Am Geriatr Soc* 2000;48:1226-33.
45. Lash TL, Fox MP, Buist DS, Wei F, Field TS, Frost FJ, et al. Mammography surveillance and mortality in older breast cancer survivors. *J Clin Oncol* 2007;25:3001-6.
46. Badgwell BD, Giordano SH, Duan ZZ, Fang S, Bedrosian I, Kuerer HM, et al. Mammography before diagnosis among women age 80 years and older with breast cancer. *J Clin Oncol* 2008;26:2482-8.
47. Boer R, de Koning HJ, van Oortmarssen GJ, van der Maas PJ. In search of the best upper age limit for breast cancer screening [Abstract]. *Eur J Cancer* 1995;31A:2040-3.
48. Berry DA, Baines CJ, Baum M, Dickersin K, Fletcher SW, Gøtzsche PC, et al. Flawed inferences about screening mammography's benefit based on observational data [Letter]. *J Clin Oncol* 2009;27:639-40; author reply 641-2.

49. Schonberg MA, McCarthy EP. Mammography screening among women age 80 years and older: consider the risks [Letter]. *J Clin Oncol* 2009;27:640-1; author reply 641-2.
50. Mandelblatt JS, Silliman R. Hanging in the balance: making decisions about the benefits and harms of breast cancer screening among the oldest old without a safety net of scientific evidence [Editorial]. *J Clin Oncol* 2009;27:487-90.
51. Bryan BB, Schnitt SJ, Collins LC. Ductal carcinoma in situ with basal-like phenotype: a possible precursor to invasive basal-like breast cancer. *Mod Pathol*. 2006;19:617-21.
52. Kerlikowske K, Molinaro A, Cha I, Ljung BM, Ernster VL, Stewart K, et al. Characteristics associated with recurrence among women with ductal carcinoma in situ treated by lumpectomy. *J Natl Cancer Inst* 2003;95:1692-702.
53. Moss S. Overdiagnosis and overtreatment of breast cancer: overdiagnosis in randomised controlled trials of breast cancer screening. *Breast Cancer Res* 2005; 7:230-4.
54. Feuer EJ, Etzioni R, Cronin KA, Mariotto A. The use of modeling to understand the impact of screening on U.S. mortality: examples from mammography and PSA testing. *Stat Methods Med Res* 2004;13:421-42.
55. de Koning HJ, Draisma G, Fracheboud J, de Bruijn A. Overdiagnosis and overtreatment of breast cancer: microsimulation modelling estimates based on observed screen and clinical data. *Breast Cancer Res* 2006;8:202.
56. Burstein HJ, Polyak K, Wong JS, Lester SC, Kaelin CM. Ductal carcinoma in situ of the breast. *N Engl J Med* 2004;350:1430-41.
57. Jones JL. Overdiagnosis and overtreatment of breast cancer: progression of ductal carcinoma in situ: the pathological perspective. *Breast Cancer Res* 2006; 8:204.
58. Weinstein MC, O'Brien B, Hornberger J, Jackson J, Johannesson M, Mc-Cabe C, et al; ISPOR Task Force on Good Research Practices—Modeling Studies. Principles of good practice for decision analytic modeling in health-care evaluation: report of the ISPOR Task Force on Good Research Practices—Modeling Studies. *Value Health* 2003;6:9-17.
59. Mandelblatt JS, Liang W, Sheppard VB, Wang J, Isaacs C. Breast cancer in minority women. In: Harris J, Lippman M, Morrow M, Osborne CK, eds. Diseases of the Breast. 4th ed. Philadelphia: Lippincott Williams & Wilkin; 2009.
60. El-Tamer MB, Ward BM, Schifftner T, Neumayer L, Khuri S, Henderson W. Morbidity and mortality following breast cancer surgery in women: national benchmarks for standards of care. *Ann Surg* 2007;245:665-71.
61. Bonomi AE, Boudreau DM, Fishman PA, Ludman E, Mohelnitzky A, Cannon EA, et al. Quality of life valuations of mammography screening. *Qual Life Res* 2008;17:801-14.
62. Ravdin PM, Cronin KA, Howlader N, Berg CD, Chlebowski RT, Feuer EJ, et al. The decrease in breast-cancer incidence in 2003 in the United States. *N Engl J Med* 2007;356:1670-4.
63. Buist DS, Porter PL, Lehman C, Taplin SH, White E. Factors contributing to mammography failure in women aged 40-49 years. *J Natl Cancer Inst* 2004; 96:1432-40.

APPENDIX

Appendix Table 6.1. Summary of model features

Feature	Model*					
	D	**E**	**G**	**M**	**S**	**W**
Includes DCIS	No	Yes	Yes	Yes	No	Yes
Includes ER status	Yes	Yes	Yes	Yes	Yes	Yes
How treatment affects mortality	Hazard reduction	Cure fraction	Hazard reduction	Hazard reduction and cure fraction based on mode of diagnosis†	Hazard reduction	Cure fraction
Calibrated to mortality?	No	No	No	Yes	No	Yes‡
Calibrated to incidence?	No	Yes	Yes	Yes	Yes	Yes
Factors affecting screening benefits§	Stage shift, age shift	Size (larger or smaller than fatal diameter)	Stage shift, age shift	Stage shift, age shift	Stage shift, size within stage, age shift	Effectiveness of treatment by stage and age shifts
Factors affecting treatment benefits (independent of screening)	ER status, age, calendar year	ER status, age	ER status, age	ER status, age, calendar year (and improvements in care)	ER status, age	ER status, age, calendar year (which affect cure probability)

DCIS = ductal carcinoma in situ; ER = estrogen receptor.

* Model group abbreviations: D = Dana-Farber Cancer Institute; E = Erasmus Medical Center; G = Georgetown University; M = M.D. Anderson Cancer Center; S =Stanford University; W = University of Wisconsin/Harvard.

† If cancer is clinically detected in model M, a hazard reduction is applied to the survival function. If cancer is detected by screening, then a cure fraction is applied for cases diagnosed in stages 1 and 2a. If cancer is detected by screening in stages 2b, 3, or 4, a similar hazard reduction is applied as for the clinically detected cases. This results in screening benefits due to stage shift and better prognosis for screening-detected versus clinically detected cases within early-stage disease. The use of a cure fraction for early-stage screening-detected cancer is a modification of the model published elsewhere.[7,11]

‡ Model W is calibrated only to mortality for a subset of the cure fraction variables after the natural history model was calibrated to incidence.

§ Note that all models use age-specific inputs for sensitivity of mammography screening. Sensitivity, in turn, has a small effect on screening benefits.

Appendix Table 6.2. Summary of base-case input data sources*

Model Inputs	Data Sets			
	BCSC	SEER 9 Registry	Connecticut Tumor Registry	Berkeley Mortality Database
Secular breast cancer incidence	No	Yes	Yes	No
Mammography test characteristics	Yes	No	No	No
Other cause of death	No	No	No	Yes
Breast cancer survival in 1975	No	Yes	No	No
Breast cancer prevalence in 1975	No	Yes	Yes	No

BCSC = Breast Cancer Surveillance Consortium; SEER 9 = Surveillance, Epidemiology, and End Results 9.

* For this analysis, we assume that 100% of women are screened and that all women detected with cancer are treated as per current practice guidelines.

Appendix Table 6.3. Average number of screening examinations and percentage of reduction in breast cancer mortality, by screening strategy

Screening Strategy	Average Screenings per 1000 Women*	Reduction in Breast Cancer Mortality (vs. No Screening), by Model, %†					
Efficient strategies (not dominated in 6 of 6 models)		**D**	**E**	**G**	**M**	**S**	**W**
Biennial screening, ages 60–69 y	4263	11	13	11	10	9	12
Biennial screening, ages 55–69 y	6890	15	18	15	14	13	19
Biennial screening, ages 50–69 y	8947	16	23	17	16	15	23
Biennial screening, ages 50–74 y	11 066	22	27	21	21	20	28
Biennial screening, ages 50–79 y	12 366	25	29	24	24	25	30
Biennial screening, ages 50–84 y	13 837	29	31	25	27	26	33
Biennial screening, ages 40–84 y	18 708	31	37	28	29	27	39
Annual screening, ages 40–84 y	36 550	38	49	32	29‡	35	54
Borderline strategies (dominated in 2–3 of 6 models)							
Biennial screening, ages 40–79 y	17 241	27§	35	26	26§	25§	36
Annual screening, ages 50–79 y	24 419	32	39	27§	26§	30	42
Annual screening, ages 50–84 y	26 905	35	41	28§	28§	33	45
Annual screening, ages 40–79 y	34 078	34§	46	30	27§	33§	51
Inefficient/dominated strategies (dominated in all 6 models)							
Annual screening, ages 60–69 y	8438	14§	18§	13§	12§	12§	17§
Biennial screening, ages 45–69 y	11 694	18§	26§	20§	19§	17§	27§
Annual screening, ages 55–69 y	13 009	18§	25§	17§	15§	16§	26§
Biennial screening, ages 40–69 y	13 831	18§	28§	20§	19§	16§	29§
Annual screening, ages 50–69 y	17 733	21§	31§	20§	18§	20§	33§
Annual screening, ages 50–74 y	21 330	27§	35§	24§	22§	26§	38§
Annual screening, ages 45–69 y	22 546	23§	35§	22§	20§	22§	39§
Annual screening, ages 40–69 y	27 428	24§	39§	23§	20§	22§	43§

* Average number of mammograms across models. Not all possible mammograms in the age group are obtained in strategies that continue to the oldest age groups, because many women die of other causes before screening would occur.

† Model group abbreviations: D = Dana-Farber Cancer Institute; E = Erasmus Medical Center; G = Georgetown University; M = M.D. Anderson Cancer Center; S =Stanford University; W = University of Wisconsin/Harvard.

‡ Because of rounding, this strategy seems to be dominated, but the actual result is 29.4.

§ Strategy is dominated ("inefficient") within the specific model. A strategy is classified as dominated if another strategy (from the efficient, borderline, or inefficient/dominated category) results in an equal or higher percentage of mortality reduction with fewer average screening examinations.

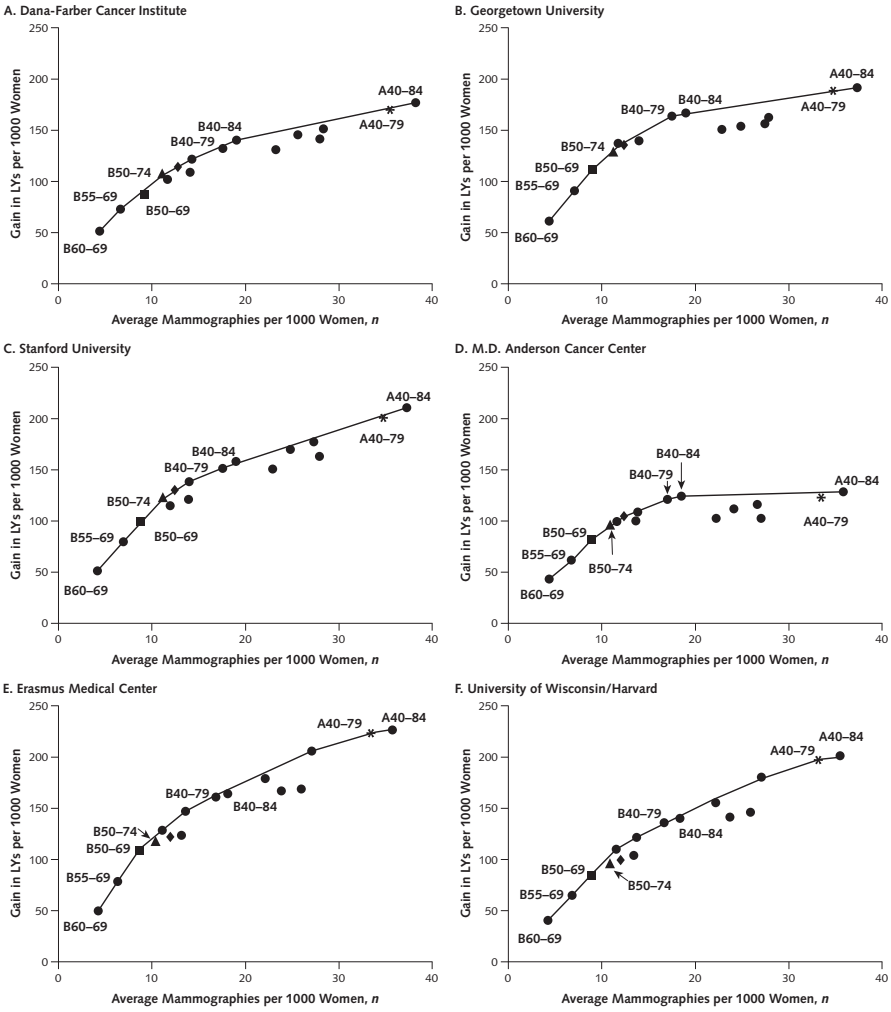
Appendix Table 6.4. Average number of screening examinations and life-years gained, by screening strategy

Screening Strategy	Average Screenings per 1000 Women*	Life-Years Gained per 1000 Women (vs. No Screening), by Model, %†					
Efficient strategies (not dominated in 5 or 6 of 6 models)		**D**	**E**	**G**	**M**	**S**	**W**
Biennial screening, ages 60–69 y	4263	51	49	61	43	52	39
Biennial screening, ages 55–69 y	6890	73	78	91	62	80	64
Biennial screening, ages 50–69 y	8947	88	107	111	82	99	84
Biennial screening, ages 50–74 y	11 066	106	116	128	96	121	95
Biennial screening, ages 40–79 y	17 241	133	161	164	122	151	136
Biennial screening, ages 40–84 y	18 708	140	164	167	126	158	140
Annual screening, ages 40–79 y	34 078	170	224	188	123‡	202	198
Annual screening, ages 40–84 y	36 550	177	227	192	128	210	202
Borderline strategies (dominated in 2–4 of 6 models)							
Biennial screening, ages 45-69 y	11 694	102‡	129	136	99	116‡	109
Biennial screening, ages 50–79 y	12 366	114	122‡	136‡	103	130	99
Biennial screening, ages 50–84 y	13 837	121	124‡	139‡	108	138	103
Biennial screening, ages 40–69 y	13 831	108‡	147	140	101‡	120‡	121
Annual screening, ages 45–69 y	22 546	131‡	179	152‡	103‡	152‡	155
Annual screening, ages 50–79 y	24 419	145	166‡	154‡	112‡	170	142‡
Annual screening, ages 50–84 y	26 905	152	169‡	157‡	116‡	178	146‡
Annual screening, ages 40–69 y	27 428	142‡	206	162‡	103‡	164‡	180
Inefficient or dominated strategies (dominated in all 6 models)							
Annual screening, ages 60–69 y	8438	65‡	69‡	71‡	53‡	69‡	56‡
Annual screening, ages 55–69 y	13 009	91‡	107‡	100‡	68‡	102‡	90‡
Annual screening, ages 50–69 y	17 733	117‡	148‡	128‡	91‡	132‡	123‡
Annual screening, ages 50–74 y	21 330	134‡	160‡	144‡	104‡	156‡	135‡

* Average number of mammograms across models. Not all possible mammograms in the age group are obtained in strategies that continue to the oldest age groups, because many women die of other causes before screening would occur.

† Model group abbreviations: D = Dana-Farber Cancer Institute; E = Erasmus Medical Center; G = Georgetown University; M = M.D. Anderson Cancer Center; S = Stanford University; W = University of Wisconsin/Harvard.

‡ Strategy is dominated within a specific model. Strategy is classified as dominated if another strategy (from the efficient, borderline or inefficient/dominated category) results in an equal or higher gain in life-years with fewer average screening examinations.



Appendix Figure 6.1. Life-years gained versus number of mammographies performed per 1000 women, by model and screening strategy.

The panels show an efficiency frontier graph for each model. The graph plots the average number of mammographies performed per 1000 women against LYs gained for each screening strategy (vs. no screening). Strategies are denoted as annual (A) or biennial (B) with starting and stopping ages. We plot efficient strategies (that is, those in which increases in use of mammography resources result in greater LYs gained than the next least-intensive strategy) in all 6 models. We also plot "borderline" strategies (approaches that are efficient in some models but not others). The line between strategies represents the "efficiency frontier." Strategies on this line would be considered efficient because they achieve the greatest gain per use of mammography resources compared with the point (or strategy) immediately below it. Points that fall below the line are not considered as efficient as those on the line. When the slope in the efficiency frontier plot levels off, the additional LYs gained per unit increase in use of mammography are small relative to the previous strategies and could indicate a point at which additional investment (use of screening) might be considered as having a low return (benefit). LY = life-year.

Chapter 7

Tipping the balance of benefits and harms to favor screening mammography starting at age 40 years: a comparative modeling study of risk

Nicolien T. van Ravesteyn, Diana L. Miglioretti, Natasha K. Stout, Sandra J. Lee, Clyde B. Schechter, Diana S.M. Buist, Hui Huang, Eveline A.M. Heijnsdijk, Amy Trentham-Dietz, Oguzhan Alagoz, Aimee M. Near, Karla Kerlikowske, Heidi D. Nelson, Jeanne S. Mandelblatt*, and Harry J. de Koning*

** dual senior authors*

Ann Intern Med. 2012 May 1;156:609-617

ABSTRACT

Background: Timing of initiation of screening for breast cancer is controversial in the United States.

Objective: To determine the threshold relative risk (RR) at which the harm–benefit ratio of screening women age 40 to 49 years equals that of biennial screening for women age 50 to 74 years.

Design: Comparative modeling study.

Data Sources: Surveillance, Epidemiology, and End Results program, Breast Cancer Surveillance Consortium, and medical literature.

Target Population: A contemporary cohort of women eligible for routine screening.

Time Horizon: Lifetime.

Perspective: Societal.

Interventions: Mammography screening starting at age 40 versus 50 years with different screening methods (film, digital) and screening intervals (annual, biennial).

Outcome Measures: Benefits: life-years gained, breast cancer deaths averted; harms: false-positive mammography findings; harm–benefit ratios: false-positive findings/life-year gained, false-positive findings/death averted.

Results of Base-Case Analysis: Screening average-risk women age 50 to 74 years biennially yields the same false-positive/life-year gained as biennial screening with digital mammography starting at age 40 for women with a 2-fold increased risk above average (median threshold RR, 1.9 [range across models, 1.5 to 4.4]). The threshold RRs are higher for annual screening with digital mammography (median, 4.3 [range, 3.3 to 10]) and higher when false-positive findings/death averted is used as outcome measure instead of false-positive findings/life-year gained. The harm–benefit ratio for film mammography is more favorable than for digital mammography because film has a lower false-positive rate.

Results of Sensitivity Analysis: The threshold RRs changed slightly when a more comprehensive measure of harm was used and were relatively insensitive to lower adherence assumptions.

Limitation: Risk was assumed to influence onset of disease without influencing screening performance.

Conclusion: Women age 40 to 49 years with a 2-fold increased risk have similar harm–benefit ratios for biennial screening mammography as average-risk women age 50 to 74 years. Threshold RRs required for favorable harm–benefit ratios vary by screening method, interval, and outcome measure.

Primary Funding Source: National Cancer Institute.

INTRODUCTION

Breast cancer is the most frequently diagnosed noncutaneous cancer among women in the United States, where it is second only to lung cancer as a cause of cancer deaths. Mammography screening has been shown to reduce breast cancer mortality rates in randomized trials [1,2] and nationwide screening programs.[3]

The U.S. Preventive Services Task Force (USPSTF) recommends biennial breast cancer mammography screening for women age 50 to 74 years on the basis of a comprehensive evaluation of current research indicating a favorable balance of benefits and harms.[4,5] Individual trials have not demonstrated significant breast cancer mortality reductions from screening women in their forties,[6,7] but a meta-analysis of 8 trials demonstrated a 15% mortality reduction.[8] The absolute benefits (for example, number of deaths prevented) are smaller than for older women because of the lower incidence of breast cancer and lower sensitivity of mammography in women age 40 to 49 years. At the same time, screening in this age group is accompanied by more harm (false-positive results and unnecessary biopsies) as a result of lower screening specificity. As a result, the USPSTF concluded that the "decision to start regular, biennial screening mammography before the age of 50 years should be an individual one and take into account patient context, including the patient's values regarding specific benefits and harms".[4,9]

Technology that improves screening test performance might influence the balance of benefits and harms of screening in younger women. Digital mammography has rapidly replaced film mammography in most areas of the United States.[10] Younger women are more likely than older women to have dense breasts, and screening regimens using digital mammography in women 40 to 49 might have a different balance of benefits and harms than strategies that use film mammography. Digital has a higher test sensitivity than film mammography in women younger than age 50,[11] detects more cases of ductal carcinoma in situ, and leads to more false-positive results.[12] Thus, it is uncertain whether initiating screening at age 40 with digital mammography would yield a more favorable balance of benefits and harms.

Another factor that changes the balance of benefits and harms is risk for developing breast cancer. Clearly, the absolute benefits of screening before age 50 will be larger for women with an increased risk for breast cancer than for average-risk women. A more risk-based screening approach might therefore be appropriate.[13-17] To implement a risk-based screening approach, it is crucial to know the magnitude of the relative risk (RR) that would tip the balance of benefits and harms to recommend screening for women age 40 to 49 (that is, threshold RR) and which risk factors lead to that elevated risk.

This study sought to determine the threshold risk at which the harm–benefit ratio of starting screening at age 40 equals the harm–benefit ratio of currently recommended biennial screening for average-risk women starting at age 50. It also evaluates the effect of screening method (film, digital) and screening interval (annual, biennial) on the threshold RR by using 4 simulation models .[18-21]

METHODS

Model overview

We used 4 microsimulation models developed as part of the Cancer Intervention and Surveillance Modeling Network (CISNET), which is an international collaborative modeling consortium funded by the National Cancer Institute.[22,23] The 4 models were model D (Dana-Farber Cancer Institute, Boston, Massachusetts), model E (Erasmus University Medical Center, Rotterdam, the Netherlands), model G-E (Georgetown University Medical Center, Washington, DC, and Albert Einstein College of Medicine, Bronx, New York), and model W (University of Wisconsin, Madison, Wisconsin, and Harvard Medical School, Boston, Massachusetts). The models have been described in detail elsewhere,[18-21] and information about the models can be found online (http://cisnet.cancer.gov/). Briefly, the models simulated life histories for individual women. After estimating breast cancer incidence and mortality in the absence of screening and treatment, the models overlaid screening use and improvements in survival associated with treatment advances.[22] The Appendix Figure (available at www.annals.org) shows the influence of breast cancer screening on (simulated) life histories. Supplement 1 (available at www.annals.org) outlines the main model differences and assumptions.

Model parameters

The model inputs and assumptions were based on assumptions previously used in a supporting article for the most recent USPSTF recommendation.[23] We used a common set of age-specific variables for breast cancer incidence, survival, and competing non–breast cancer causes of death (Supplement 2, available at www.annals.org). A cohort of women born in 1960 was simulated and followed throughout their entire lifetime. We assumed 100% adherence with screening and adjuvant treatment guidelines.

The Breast Cancer Surveillance Consortium (BCSC) provided data on recent performance (between 2001 and 2007) of film and digital mammography (Table 7.1). The BCSC collects prospective data on breast imaging in community practice in 5 mammography registries and 2 affiliated sites in the United States (http://breastscreening.cancer.gov). Each registry obtains annual approval from its institutional review board for consenting processes or a waiver of consent, enrollment of participants, and ongoing data linkages for research purposes. All registries have received Federal Certificates of Confidentiality that protect the identities of research participants. The 4 models used BCSC data inputs for sensitivity; specificity; and stage distribution by age, screening method (film, digital), and screening interval (annual, biennial) (http://breastscreening.cancer.gov/data/elements.html).

Screening strategies

The effects of 5 screening scenarios were estimated per 1000 average-risk women age 40 followed over their lifetimes. These scenarios included biennial screening for women age 50 to

Table 7.1. Sensitivity and specificity of screening mammography (digital and film) by age and screening interval used in the cancer intervention and surveillance modeling network models based on data from the breast cancer surveillance consortium, 2001–2007*

Test Characteristic According to Age	Film Mammography		Digital Mammography	
	Annual	Biennial	Annual	Biennial
Sensitivity				
Age 40–44 y	69.5 (62.6–75.9)	68.0 (57.8–77.1)	78.4 (64.4–90.9)†	90.7 (76.8–100)†
Age 45–49 y	73.8 (69.6–77.7)	80.6 (75.1–85.3)	79.8 (67.6–87.7)†	91.6 (73.0–99.0)†
Specificity				
Age 40–44 y	91.0 (90.8–91.2)	90.1 (89.8–90.3)	87.5 (87.0–88.0)	87.7 (86.8–88.6)
Age 45–49 y	90.9 (90.7–91.0)	89.8 (89.6–90.0)	87.9 (87.5–88.3)	87.4 (86.6–88.2)

* Values in parentheses are 95% CIs.

† The models used extrapolated values for the sensitivity of digital mammography because the 95% CIs of the sensitivity of digital mammography were wide. The extrapolated values were derived by using regression models based on the sensitivity values in other age groups (data not shown). The 4 extrapolated numbers are shown and were very close to the Breast Cancer Surveillance Consortium data (80.0 and 78.9 for annual screening at ages 40 to 44 and 45 to 49 y, respectively, and 91.7 for biennial screening at age 45 to 49 y), except for the sensitivity of biennial screening with digital mammography for the 40- to 44-y age group, for whom the Breast Cancer Surveillance Consortium data indicated a sensitivity of 100.

74 extended with 4 screening scenarios for women age 40 to 49 varying by screening interval (annual and biennial) and screening method (film and digital): 1) biennial film screening at age 50 to 74; 2a) biennial film screening at age 50 to 74 and biennial film screening at age 40 to 49; 2b) biennial film screening at age 50 to 74 and biennial digital screening at age 40 to 49; 3a) biennial film screening at age 50 to 74 and annual film screening at age 40 to 49; and 3b) biennial film screening at age 50 to 74 and annual digital screening at age 40 to 49.

The incremental effects of each scenario were determined by comparing it to the previous, less intensive scenario. To determine the effects of adding biennial screening for women in their forties, scenarios 2a and 2b were compared with scenario 1; for annual screening, scenarios 3a and 3b were compared with scenarios 2a and 2b, respectively.

Benefit, harms, and harm–benefit ratios

We estimated the effect of each screening strategy on the number of breast cancers detected, the number of breast cancer deaths averted, the number of life-years gained, and the number of false-positive results on mammography screening. The time horizon for calculating effects was from age 40 until death of all simulated women.

First, the effects (benefits and harms) of biennial screening between ages 50 and 74 were determined. We defined harm as the number of false-positive results and benefits as the number of breast cancer deaths averted and number of life-years gained. Then, the additional effects of screening between ages 40 and 49 were assessed for 2 screening intervals (annual and biennial) and 2 screening methods (film and digital). For each screening strategy,

we determined the harm–benefit ratios by dividing the incremental harm by the incremental benefits.

We then implemented different RRs in the models. The higher risk was modeled over the entire lifetime of the simulated women. We calculated the incremental harms, benefits, and harm–benefit ratios of the 5 screening scenarios for women at increased risk. We used the harm–benefit ratio of biennial screening for average-risk women between ages 50 and 74 (scenario 1) as a threshold value. Then, we determined for the 4 screening scenarios (scenarios 2a and 2b and 3a and 3b) that started at age 40 how high the RR needed to be to have the same harm–benefit ratio as the threshold value.

Sensitivity analysis

To evaluate how the harm–benefit ratios and threshold RRs were influenced by certain assumptions and parameter values, we performed several sensitivity analyses. First, we explored the effect of reduced adherence (70%). Second, we considered alternative screening test characteristics of digital mammography (best-case scenario) by using an improved sensitivity and specificity (using the upper limit of the 95% CI). Third, we assessed the influence of a broader measure of harm on the threshold RRs by calculating quality-adjusted life-years (QALYs) lost. We applied quality-of-life decrements due to undergoing mammography and diagnostics [24] and life-years with breast cancer by stage of disease at diagnosis [25] (Supplement 3, available at www.annals.org). The number of QALYs lost incorporates harm from undergoing mammography, having a false-positive screening test result, and harm from overdiagnosis because with more overdiagnosis, more life-years are spent in disease stages and, thus, more QALYs are lost when there is a large amount of overdiagnosis. All 4 models incorporate overdiagnosis because in all models there is a chance that breast cancer is detected in the presence of screening but would not have become symptomatic and would not have caused breast cancer death during a woman's lifetime if no screening had taken place. Overdiagnosis estimates vary across models because of different underlying assumptions (for example, on the rates of progression from ductal carcinoma in situ to invasive disease and on the possibility of limited malignant potential of invasive disease). We used the harm–benefit ratio QALYs lost/life-year gained to re-estimate the threshold RRs.

Role of the funding source

This work was funded by a supplement from the National Cancer Institute. The funding sources had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, and approval of the manuscript; or decision to submit this manuscript for publication.

RESULTS

Average risk

In the absence of screening, the models estimate that a median of 153 cases of breast cancer would be diagnosed (range across models, 152 to 158) and 25 deaths from breast cancer would occur (range, 19 to 35) among 1000 women 40 years of age followed over their lifetimes (data not shown).

If these women underwent biennial film mammography between the ages of 50 and 74, the models predict that 10 610 (range, 10 529 to 10 660) mammograms would be obtained over 12 screening rounds; 6.3 (range, 5.3 to 6.9) breast cancer deaths would be avoided; 109 (range, 93 to 113) life-years would be gained; and 883 (range, 795 to 939) mammography results would be read as false-positive. The harm–benefit ratios are estimated as 8.3 (range, 7.8 to 8.7) false-positive findings/life-year gained and 146 (range, 128 to 151) false-positive findings/death averted (Table 7.2).

In all models, the harm–benefit ratios for adding screening between ages 40 and 49 are less favorable than those for biennial screening starting at age 50 (Table 7.3). In all models, adding annual to biennial screening leads to slight increases in additional life-years gained and breast cancer deaths averted, but at the expense of greater increases in incremental harm. Adding annual screening to biennial screening starting at age 40 therefore has a less favorable harm–benefit ratio than adding biennial screening starting at age 40 to biennial screening from ages 50 to 74 in all 4 models (Table 7.3).

With digital mammography screening, more life-years are gained and more breast cancer deaths averted than with plain film mammography, but because of the lower specificity of

Table 7.2. Effects of biennial screening between ages 50 and 74 years per 1000 women*

Model†	Mammography Screenings, *n*	Breast Cancer Cases Detected, *n*	Benefits		Harms	Harm–Benefit Ratios	
			Breast Cancer Deaths Averted, *n*	Life-Years Gained	False-Positive Findings, *n*	False-Positive Findings/Breast Cancer Deaths Averted	False-Positive Findings/Life-Years Gained
D	10 529	156	6.9	113	877	128	7.8
E	10 655	164	5.3	93	795	151	8.5
G-E	10 566	163	6.6	108	939	142	8.7
W	10 660	232	5.9	111	890	151	8.0
Median	10 610	163	6.3	109	883	146	8.3

* Screening equivalent to 12.5 screening rounds in 25 y.

† Model D: Dana-Farber Cancer Institute, Boston, Massachusetts; model E: Erasmus University Medical Center, Rotterdam, the Netherlands; model G-E: Georgetown University Medical Center, Washington, DC, and Albert Einstein College of Medicine, Bronx, New York; model W: University of Wisconsin, Madison, Wisconsin, and Harvard Medical School, Boston, Massachusetts.

digital mammography for women in their forties, it also yields more false-positive results. In all 4 models, there is greater harm relative to benefit from digital than from film mammography in women age 40 to 49 (Table 7.3).

Table 7.3. Incremental changes in the benefits and harms of starting screening at age 40 years instead of at age 50 years, per 1000 women

Model*	Benefits				Harms		Harm-Benefit Ratios			
	Additional Breast Cancer Deaths Averted, *n*		Additional Life-Years Gained		Additional False-Positive Findings, *n*		False-Positive Findings/Breast Cancer Deaths Averted		False-Positive Findings/Life-Years Gained	
	Film	Digital	Film	Digital	Film	Digital	Film	Digital	Film	Digital
Biennial screening for women starting at age 40 y†										
D	0.5	0.5	17	16	490	607	896	1166	29	39
E	1.2	1.3	37	39	470	582	381	450	13	15
G	1.0	1.1	29	32	406	579	406	526	14	18
W	1.3	1.6	41	49	486	603	363	370	12	12
Median	1.1	1.2	33	36	478	592	393	488	13	16
Annual screening for women starting at age 40 y‡										
D	0.2	0.3	8	11	393	593	1579	1732	50	54
E	0.6	0.6	18	19	376	568	641	940	20	30
G	0.3	0.2	7	7	426	577	1420	2884	60	84
W	0.7	0.7	21	21	390	589	565	841	18	28
Median	0.4	0.5	13	15	391	583	1030	1336	35	42

* Model D: Dana-Farber Cancer Institute, Boston, Massachusetts; model E: Erasmus University Medical Center, Rotterdam, the Netherlands; model G-E: Georgetown University Medical Center, Washington, DC, and Albert Einstein College of Medicine, Bronx, New York; model W: University of Wisconsin, Madison, Wisconsin, and Harvard Medical School, Boston, Massachusetts.
† Incremental to biennial screening for women starting at age 50 y.
‡ Incremental to biennial screening for women starting at age 40 y.

Increased risk

In all models, screening women with increased risk for breast cancer leads to a more favorable harm–benefit ratio. Screening women with increased risk results in more life-years gained and more breast cancer deaths averted with approximately the same number of false-positive results (see Supplement 4 and 5, available at www.annals.org, for data on women with a 2-fold increased risk for cancer). The models predict that annual screening with digital mammography for women age 40 to 49 with a 4-fold increased risk above average (median threshold RR, 4.3 [range across models, 3.3 to 10]) would yield similar false-positive findings/life-year gained estimates as biennial screening for average-risk women age 50 to 74 years. To find similar false-positive findings/life-year gained estimates for biennial screening with

Table 7.4. Threshold relative risks estimated by the Cancer Intervention and Surveillance Modeling Network models for the different screening strategies*

Model†	Relative Risks Estimated by Using False-Positive Findings/Breast Cancer Deaths Averted				Relative Risks Estimated by Using False-Positive Findings/Life-Years Gained			
	Biennial		Annual		Biennial		Annual	
	Film	Digital	Film	Digital	Film	Digital	Film	Digital
D	5.1	5.7	6.4	6.6	3.7	4.4	4.9	5.1
E	2.4	2.8	3.8	5.6	1.5	1.7	2.3	3.3
G-E	2.9	3.7	>10	>10	1.6	2.1	7.0	10
W	2.2	2.3	3.4	5.1	1.5	1.5	2.2	3.4
Median	2.7	3.3	5.1	6.1	1.6	1.9	3.6	4.3

* The threshold relative risks represent the risk at which the harm–benefit ratio of starting biennial or annual screening at age 40 y would equal that of currently recommended biennial screening starting at age 50 y.
† Model D: Dana-Farber Cancer Institute, Boston, Massachusetts; model E: Erasmus University Medical Center, Rotterdam, the Netherlands; model G-E: Georgetown University Medical Center, Washington, DC, and Albert Einstein College of Medicine, Bronx, New York; model W: University of Wisconsin, Madison, Wisconsin, and Harvard Medical School, Boston, Massachusetts.

digital mammography for women age 40 to 49, the threshold RRs are lower in all models (median threshold RR, 1.9 [range across models, 1.5 to 4.4]).

For screening with film mammography, the threshold RRs are predicted to be somewhat lower in all models than those estimated for digital mammography (median threshold RR for biennial screening, 1.6 [range, 1.5 to 3.7]; Table 7.4). When deaths averted was considered as an outcome measure instead of life-years gained, all models estimated higher threshold RRs of 3.3 (range, 2.3 to 5.7) for biennial screening and even higher threshold RRs for annual screening (Table 7.4). The incremental changes in the benefits and harms of starting screening at age 40 instead of at age 50 per 1000 women for the threshold RRs are listed in Supplement 6 (available at www.annals.org).

Sensitivity analysis

When 70% adherence was assumed instead of 100%, the harms diminished by 30%, whereas the benefits diminished less (by 26% for biennial and 5% for annual screening). Therefore, more favorable harm–benefit ratios were found for biennial screening (for example, for film mammography in model E: 12.2 false-positive findings/life-year gained versus 12.7 when 100% attendance was assumed) and for annual screening (in model E: 14.9 false-positive findings/life-year gained versus 20.4 when 100% attendance was assumed). However, adding annual screening to biennial screening starting at age 40 still had a less favorable harm–benefit ratio than adding biennial screening.

Changing screening test characteristics of digital mammography had only a very small influence on the harm-benefit ratios. The best-case scenario resulted in somewhat more benefits and fewer harms, but the differences in harm–benefit ratios were very small (<8%).

Use of the more comprehensive measure of harm (QALYs lost) led to somewhat higher predicted median threshold RRs (for example, 2.1 [range, 1.4 to 2.9] for biennial screening with digital mammography [versus 1.9 when false-positive findings were used]).

DISCUSSION

For women with approximately 2-fold increased risk for breast cancer, the balance of benefits and harms (life-year gained vs. false-positive results) of starting biennial screening at age 40 approximates that of biennial screening for average-risk women starting at age 50. The models consistently showed that the additional benefits of adding annual screening are small and that there is greater harm relative to benefit from digital than from film mammography in women age 40 to 49 years. To obtain harm–benefit ratios similar to those that result from currently recommended screening, the false-positive rates for biennial screening with digital mammography would have to decrease substantially among women in their forties.

The model results on the difference between annual and biennial screening are largely in line with previous work. A retrospective study found that women screened annually versus biennially had similar distributions of prognostic factors (such as tumor size, lymph node status, and histologic grade).[26] However, another study found that specifically among women age 40 to 49 years, those undergoing biennial screening, were more likely to have late-stage disease at diagnosis than those undergoing annual screening.[27] It has been suggested that for younger women annual screening would be more beneficial than biennial screening because of the faster tumor growth rates in this age group.[28-30] This is reflected in the model outcomes showing that adding annual screening to biennial screening in the 40- to 49-year age group is somewhat more beneficial than in older age groups. For example, a previous study showed that 72% to 89% of the mortality benefit is maintained in these 4 models when women age 50 to 74 move from annual to biennial screening scenarios.[23] The present study shows that in the 40- to 49-year age group the percentage of mortality benefit maintained is somewhat lower (66% to 77%).

All 4 models found only small differences between film and digital mammography with regard to the benefits of screening, which is in line with results of a study reporting that improvements in sensitivity did not markedly affect breast cancer mortality.[31] However, digital mammography did result in substantially more false-positive results than did film mammography. This translated into greater harm relative to benefits for digital than for film mammography in younger women. Therefore, it seems unlikely that data on the performance of digital mammography, if it had been available, would have led the USPSTF to recommend screening women starting at age 40. A recent study in the Netherlands found that referral and false-positive rates first increased after the implementation of digital mammography, but then these rates started to decrease over time and stabilized at a somewhat higher level

than film mammography after a little more than 1 year.[32] However, in a recent U.S. study comparing the screening performance of digital and film mammography, excluding the first year after the transition to digital mammography did not influence results.[33] Another recent study found that the availability of comparison mammography halved the false-positive recall probability.[34] It remains to be investigated whether false-positive rates in the United States can be reduced without also decreasing sensitivity or detection rates.

The results of the models are consistent regarding differences between outcome measures, predicting considerably higher threshold RRs when breast cancer deaths averted are used instead of life-years gained because in the 40- to 49-year age group there are more life-years to gain by averting a death than in older age groups. Life-years gained may be considered preferable because as a summary measure it incorporates both the number of lives saved and the number of life-years gained per life saved. Our results indicate that the outcome measure used is a main determinant of the screening strategy that will be chosen for women age 40 to 49, highlighting the importance of taking into account preferences of individual women for specific benefits and harms.

Several limitations of this study should be mentioned. Of note, we calculated the harm–benefit ratios for women age 50 to 74 who were screened biennially and used these as threshold values for younger women. However, it might be that younger women have different concerns and preferences than do older women and that these preferences vary between individual women. In calculating the harm–benefit ratio, we included only false-positive screening mammography results as the harm. Ideally, all harms and all benefits are taken into account in determining optimal screening scenarios. In addition to false-positive results, harms of screening mammography include unnecessary biopsies, radiation exposure, false reassurance, pain related to the procedure, overdiagnosis (the detection of lesions that would not have become clinically apparent without screening), overtreatment, and the burden of performing medical tests on healthy individuals. Several studies have shown that the risk for radiation is minimal,[35,36] and false reassurance has been found to play only a minor role in breast cancer screening.[37] Although many women experience pain during the procedure (range, 1% to 77%), very few consider this a deterrent from future screening.[8,38] Estimates of overdiagnosis vary widely, ranging up to 54%.[39] Although a proportion of invasive cancer diagnosed by mammography may never have presented clinically, the proportion is likely to be small for women age 40 to 49, ranging up to 7%.[39] For ductal carcinoma in situ, this proportion might be larger but is surrounded by uncertainty. For these reasons we chose to focus on false-positive examinations as the main harm for women in their forties. We did, however, perform a sensitivity analysis in which we considered a more comprehensive measure of harm, QALYs lost. Although this measure is more comprehensive, capturing disutility of false-positive results and the effect of overdiagnosis, it is less transparent than the number of false-positive results, and the preferences of individual women might diverge from the assumed societal utilities.

Another limitation is that the models differed for some outcomes. For biennial film screening, 3 models (E, G-E, and W) found similar threshold RRs (1.5 to 1.6), whereas 1 model (D) estimated a threshold RR of 3.7. This discrepancy relates to differences in the estimated benefits, reflecting differences in model structures. In model D [19], the stage distribution data are directly incorporated in constructing breast cancer–specific survival. Thus, small incremental changes in stage shifts between annual/biennial or between film/digital led to smaller incremental benefits. The other models used a combination of sensitivity values and stage distribution to calibrate parameters and show larger benefits. Additionally, the models differed regarding the predicted incremental benefit of adding annual to biennial screening (range of life-years gained, 7 to 21) because the models make different assumptions for unobservable variables, such as sojourn time, which is the duration of the preclinical, screen-detectable phase of the tumor. To our knowledge, no randomized, controlled trials have directly compared annual and biennial screening. The range in model outcomes thus reflects uncertainties in current knowledge of the incremental benefits of screening women age 40 to 49 and about shortening the screening interval.

Finally, model outcomes largely depended on the inputs and assumptions. One assumption was that the higher risk influenced only the incidence (onset of disease) and not the screening performance (sensitivity, specificity) or natural history of disease (such as the tumor growth rate and breast cancer survival). However, at least some risk factors, including breast density and family history, influence both breast cancer risk and screening performance.[40,41] If this is taken into account, the harm–benefit ratio could change for women with risk factors that influence performance. The psychological effect of false-positive results might also differ by risk group. For example, the amount of anxiety or distress might be higher for younger women and for women with a family history of breast cancer than for average-risk women.[42]

Our finding that women with increased risks for breast cancer have similar harm–benefit ratios from starting biennial screening mammography at age 40 is in line with studies finding that breast cancer risk or detection for women with a first-degree relative is similar to that for women a decade older without such a history.[40] Several other countries have risk-based screening guidelines. For example, guidelines in the Netherlands state that women with a moderately increased risk, defined as an RR of 2 to 3, should be offered annual screening starting at age 40. Similarly, in Australia, guidelines specify that a starting age younger than 50 or more frequent examinations should be considered individually for women with moderately increased risk, defined as an RR of 1.5 to 3.

A systematic review and meta-analysis of risk factors and their prevalence rates in women age 40 to 49 years in the United States was conducted jointly with the present study.[43] Two risk factors were associated with a 2-fold or higher RR: having a first-degree relative with breast cancer (9% of women in the United States) and extremely dense breasts on mammography (13% of women with Breast Imaging Reporting and Data System category 4 breast density). Results of these two studies imply that women with these characteristics could ben-

efit from biennial screening starting at age 40 and that for them the balance of benefits and harms of screening would be similar to the balance of benefits and harms for average-risk women starting screening at age 50. In addition to these single risk factors, combinations of risk factors could potentially reach this risk threshold.[16,17] A potential difficulty with including breast density in screening recommendations is that breast density is not uniformly reported and requires baseline mammography examinations to determine breast density, introducing additional potential screening harms.

Our results provide important information toward more individualized, risk-based screening, suggesting that starting biennial screening at age 40 for women with an increased risk for breast cancer (RR ≥ 1.9) has a balance of benefits and harms similar to that of biennial screening for average-risk women age 50 to 74 years. For women below this level of risk, the harm–benefit ratio of starting screening at age 40 is less favorable than that of biennial screening between ages 50 and 74. Reducing the false-positive rate is crucial to improving the balance of benefits and harms for screening regimens for women of all ages.

Acknowledgments

The authors thank Drs. Kathleen Cronin and Brian Sprague for their valuable advice and consultation on this project. They also thank the BCSC investigators, participating women, mammography facilities, and radiologists for the data they have provided for this study. The BCSC investigators and procedures for requesting BCSC data for research purposes are listed at: http://breastscreening.cancer.gov/.

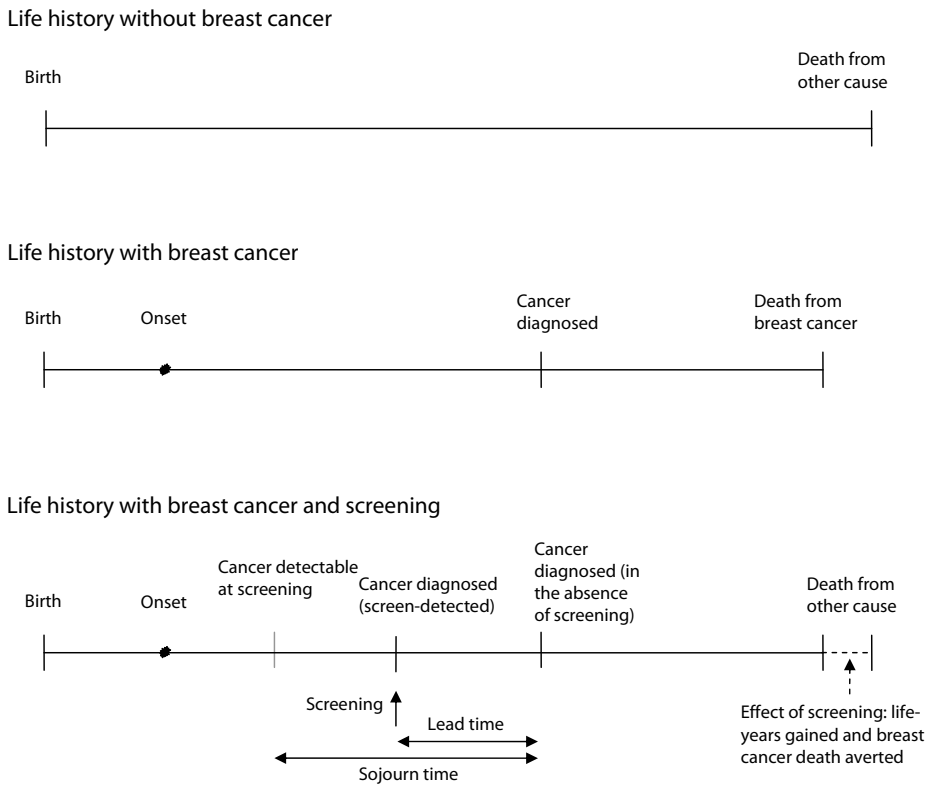
REFERENCES

1. Nyström L, Andersson I, Bjurstam N, Frisell J, Nordenskjöld B, Rutqvist LE. Long-term effects of mammography screening: updated overview of the Swedish randomised trials. *Lancet* 2002;359:909-19.
2. Tabár L, Vitak B, Chen HH, Duffy SW, Yen MF, Chiang CF, et al. The Swedish Two-County Trial twenty years later. Updated mortality results and new insights from long-term follow-up. *Radiol Clin North Am* 2000;38:625-51.
3. Otto SJ, Fracheboud J, Looman CW, Broeders MJ, Boer R, Hendriks JH, et al; National Evaluation Team for Breast Cancer Screening. Initiation of population-based mammography screening in Dutch municipalities and effect on breast-cancer mortality: a systematic review. *Lancet* 2003;361:1411-7.
4. U.S. Preventive Services Task Force. Screening for breast cancer: U.S. Preventive Services Task Force recommendation statement. *Ann Intern Med* 2009;151:716-26, W-236.
5. Nelson HD, Tyne K, Naik A, Bougatsos C, Chan B, Nygren P, et al. Screening for beast cancer: systematic evidence review update for the US Preventive Services Task Force. Bethesda, MD: Agency for Healthcare Research and Quality; 2009. Report No.: 10-05142-EF-1.
6. Moss SM, Cuckle H, Evans A, Johns L, Waller M, Bobrow L; Trial Management Group. Effect of mammographic screening from age 40 years on breast cancer mortality at 10 years' follow-up: a randomised controlled trial. *Lancet* 2006;368:2053-60.
7. Miller AB, To T, Baines CJ, Wall C. The Canadian National Breast Screening Study-1: breast cancer mortality after 11 to 16 years of follow-up. A randomized screening trial of mammography in women age 40 to 49 years. *Ann Intern Med* 2002;137:305-12.
8. Nelson HD, Tyne K, Naik A, Bougatsos C, Chan BK, Humphrey L; U.S. Preventive Services Task Force. Screening for breast cancer: an update for the U.S. Preventive Services Task Force. *Ann Intern Med* 2009;151:727-37, W237-42.
9. U.S. Preventive Services Task Force. Screening for breast cancer. Accessed at http://www.uspreventiveservicestaskforce.org/uspstf/uspsbrca.htm on 29 August 2011
10. U.S. Food and Drug Administration. Mammography Quality Standards Act. 2011 Scorecard Statistics. Accessed at http://www.fda.gov/Radiation-EmittingProducts/MammographyQualityStandardsActandProgram/DocumentArchives/ucm241654.htm on 29 July 2011
11. Pisano ED, Gatsonis C, Hendrick E, Yaffe M, Baum JK, Acharyya S, et al; Digital Mammographic Imaging Screening Trial (DMIST) Investigators Group. Diagnostic performance of digital versus film mammography for breast-cancer screening. *N Engl J Med* 2005;353:1773-83.
12. Karssemeijer N, Bluekens AM, Beijerinck D, Deurenberg JJ, Beekman M, Visser R, et al. Breast cancer screening results 5 years after introduction of digital mammography in a population-based screening program. *Radiology* 2009;253:353-8.
13. Qaseem A, Snow V, Sherif K, Aronson M, Weiss KB, Owens DK; Clinical Efficacy Assessment Subcommittee of the American College of Physicians. Screening mammography for women 40 to 49 years of age: a clinical practice guideline from the American College of Physicians. *Ann Intern Med* 2007;146:511-5.
14. Partridge AH, Winer EP. On mammography---more agreement than disagreement. *N Engl J Med* 2009;361:2499-501.
15. Kerlikowske K. Evidence-based breast cancer prevention: the importance of individual risk [Editorial]. *Ann Intern Med* 2009;151:750-2.
16. Schousboe JT, Kerlikowske K, Loh A, Cummings SR. Personalizing mammography by breast density and other risk factors for breast cancer: analysis of health benefits and cost-effectiveness. *Ann Intern Med* 2011;155:10-20.
17. Tice JA, Cummings SR, Smith-Bindman R, Ichikawa L, Barlow WE, Kerlikowske K. Using clinical factors and mammographic breast density to estimate breast cancer risk: development and validation of a new predictive model. *Ann Intern Med* 2008;148:337-47.
18. Fryback DG, Stout NK, Rosenberg MA, Trentham-Dietz A, Kuruchittham V, Remington PL. The Wisconsin Breast Cancer Epidemiology Simulation Model. *J Natl Cancer Inst Monogr* 2006:37-47.
19. Lee S, Zelen M. A stochastic model for predicting the mortality of breast cancer. *J Natl Cancer Inst Monogr* 2006:79-86.
20. Mandelblatt J, Schechter CB, Lawrence W, Yi B, Cullen J. The SPECTRUM population model of the impact of screening and treatment on U.S. breast cancer trends from 1975 to 2000: principles and practice of the model methods. *J Natl Cancer Inst Monogr* 2006:47-55.
21. Tan SY, van Oortmarssen GJ, de Koning HJ, Boer R, Habbema JD. The MISCAN-Fadia continuous tumor growth model for breast cancer. *J Natl Cancer Inst Monogr* 2006:56-65.

22. Berry DA, Cronin KA, Plevritis SK, Fryback DG, Clarke L, Zelen M, et al. Effect of screening and adjuvant therapy on mortality from breast cancer. *N Engl J Med* 2005;353:1784-92.
23. Mandelblatt JS, Cronin KA, Bailey S, Berry DA, de Koning HJ, Draisma G, et al; Breast Cancer Working Group of the Cancer Intervention and Surveillance Modeling Network. Effects of mammography screening under different screening schedules: model estimates of potential benefits and harms. *Ann Intern Med* 2009;151:738-47.
24. de Haes JC, de Koning HJ, van Oortmarssen GJ, van Agt HM, de Bruyn AE, van Der Maas PJ. The impact of a breast cancer screening programme on quality-adjusted life-years. *Int J Cancer* 1991;49:538-44.
25. Stout NK, Rosenberg MA, Trentham-Dietz A, Smith MA, Robinson SM, Fryback DG. Retrospective cost-effectiveness analysis of screening mammography. *J Natl Cancer Inst* 2006;98:774-82.
26. Wai ES, D'yachkova Y, Olivotto IA, Tyldesley S, Phillips N, Warren LJ, et al. Comparison of 1- and 2-year screening intervals for women undergoing screening mammography. *Br J Cancer* 2005;92:961-6.
27. White E, Miglioretti DL, Yankaskas BC, Geller BM, Rosenberg RD, Kerlikowske K, et al. Biennial versus annual mammography and the risk of late-stage breast cancer. *J Natl Cancer Inst* 2004;96:1832-9.
28. Moskowitz M. Breast cancer: age-specific growth rates and screening strategies. *Radiology* 1986;161:37-41.
29. Tabar L, Fagerberg G, Chen HH, Duffy SW, Smart CR, Gad A, et al. Efficacy of breast cancer screening by age. New results from the Swedish Two-County Trial. *Cancer* 1995;75:2507-17.
30. Buist DS, Porter PL, Lehman C, Taplin SH, White E. Factors contributing to mammography failure in women aged 40-49 years. *J Natl Cancer Inst* 2004;96:1432-40.
31. Taylor P. Modelling the impact of changes in sensitivity on the outcomes of the UK breast screening programme. *J Med Screen* 2010;17:31-6.
32. Bluekens AM, Karssemeijer N, Beijerinck D, Deurenberg JJ, van Engen RE, Broeders MJ, et al. Consequences of digital mammography in population-based breast cancer screening: initial changes and long-term impact on referral rates. *Eur Radiol* 2010;20:2067-73.
33. Kerlikowske K, Hubbard RA, Miglioretti DL, Geller BM, Yankaskas BC, Lehman CD. Comparative effectiveness of digital versus film-screen mammography in community practice in the United States: a cohort study. *Ann Intern Med* 2011;155:493-502.
34. Hubbard RA, Kerlikowske K, Flowers CI, Yankaskas BC, Zhu W, Miglioretti DL. Cumulative probability of false-positive recall or biopsy recommendation after 10 years of screening mammography: a cohort study. *Ann Intern Med* 2011;155:481-92.
35. Yaffe MJ, Mainprize JG. Risk of radiation-induced breast cancer from mammographic screening. *Radiology* 2011;258:98-105.
36. de Gelder R, Draisma G, Heijnsdijk EA, de Koning HJ. Population-based mammography screening below age 50: balancing radiation-induced vs prevented breast cancer deaths. *Br J Cancer* 2011;104:1214-20.
37. de Gelder R, van As E, Tilanus-Linthorst MM, Bartels CC, Boer R, Draisma G, et al. Breast cancer screening: evidence for false reassurance? *Int J Cancer* 2008;123:680-6.
38. Armstrong K, Moye E, Williams S, Berlin JA, Reynolds EE. Screening mammography in women 40 to 49 years of age: a systematic review for the American College of Physicians. *Ann Intern Med* 2007;146:516-26.
39. Biesheuvel C, Barratt A, Howard K, Houssami N, Irwig L. Effects of study methods and biases on estimates of invasive breast cancer overdetection with mammography screening: a systematic review. *Lancet Oncol* 2007;8:1129-38.
40. Kerlikowske K, Carney PA, Geller B, Mandelson MT, Taplin SH, Malvin K, et al. Performance of screening mammography among women with and without a first-degree relative with breast cancer. *Ann Intern Med* 2000;133:855-63.
41. Carney PA, Miglioretti DL, Yankaskas BC, Kerlikowske K, Rosenberg R, Rutter CM, et al. Individual and combined effects of age, breast density, and hormone replacement therapy use on the accuracy of screening mammography. *Ann Intern Med* 2003;138:168-75
42. Gilbert FJ, Cordiner CM, Affleck IR, Hood DB, Mathieson D, Walker LG. Breast screening: the psychological sequelae of false-positive recall in women with and without a family history of breast cancer. *Eur J Cancer* 1998;34:2010-4.
43. Nelson H, Zakher B, Cantor A, Fu R, Griffin J, O'Meara E, et al. Risk factors for breast cancer for women age 40 to 49: a systematic review and meta-analysis. *Ann Intern Med* 2012; 156:635-48.
44. Lee SJ, Zelen M. Modelling the early detection of breast cancer. *Ann Oncol* 2003;14:1199-202.
45. van Ravesteyn NT, Heijnsdijk EA, Draisma G, de Koning HJ. Prediction of higher mortality reduction for the UK Breast Screening Frequency Trial: a model-based approach on screening intervals. *Br J Cancer* 2011;105:1082-8.

46. Holford TR, Cronin KA, Mariotto AB, Feuer EJ. Changing patterns in breast cancer incidence trends. *J Natl Cancer Inst Monogr* 2006:19-25.
47. Rosenberg MA. Competing risks to breast cancer mortality. *J Natl Cancer Inst Monogr* 2006:15-9.
48. Cronin KA, Mariotto AB, Clarke LD, Feuer EJ. Additional common inputs for analyzing impact of adjuvant therapy and mammography on U.S. mortality. *J Natl Cancer Inst Monogr* 2006:26-9.

APPPENDIX



Appendix Figure 7.1. Schematic overview of simulated life histories and the effect of screening.

The italicized words in the descriptions below refer to the words outlined in the figure. *Sojourn time* is the duration of the preclinical, screen-detectable phase of the tumor and *lead time* is the interval from screen detection to the time of clinical diagnosis, when the tumor would have surfaced without screening. **Model D** is a state transition model where potential benefit from early detection arises because of a stage shift. The natural history of breast cancer is modeled analytically using stochastic models. The model assumes breast cancer (invasive) progresses from a no-disease state (S0) to pre-clinical (Sp) state and to clinical state (Sc). Some cases will continue to the disease-specific death (Sd) state. Death due to other causes is treated as a competing risk. The Sp state begins when *cancer is detectable at screening* and Sc begins *when cancer is diagnosed in absence of screening*. For a given birth cohort, age-specific invasive breast cancer incidence rate and age-dependent *sojourn time* in pre-clinical state (published values) are used to estimate the transition probabilities from S0 to Sp. The transition probabilities from Sp to Sc are estimated based on the age-specific breast cancer incidence rate. The other basic assumption is that any reduction in mortality associated with screening is from the stage-shift. That is *screen-detected cases* have a better stage distribution with a higher proportion of cases in earlier stages. The stage distribution data for *screen-detected* cases are obtained from BCSC and directly incorporated in constructing breast cancer specific survival. Also *the lead time* for screen-detected cases is treated as a random variable and adjusted in constructing the breast cancer specific survival *for screen-detected cases*. When cancer is diagnosed, a

treatment is applied by age, stage and ER status and treatment reduces the hazard of breast cancer specific mortality by age, stage and ER status. **Model E** is a microsimulation model based on continuous tumor growth. The natural history of breast-cancer is modeled as a continuously growing tumor from *onset* of cancer (starting with a tumor diameter of 0.1 mm). The moments that events happen are determined by tumor sizes. The screening threshold diameter determines the moment that *the cancer is detectable at screening*, and the diameter of clinical detection determines when the *cancer will be diagnosed in the absence of screening*. Each tumor has a size (the fatal diameter, which differs between tumors) at which diagnosis and treatment will no longer result in cure given available treatment options. If the tumor is diagnosed (either on the basis of clinical presentation with symptoms or by screening) and treated before the tumor reaches the fatal diameter, the woman will be cured and will die of non-breast cancer causes (*death from other causes*). Variation between tumors is modeled by probability distributions of parameters. Screening might detect tumors at a smaller tumor size with a larger probability of cure (because the tumor has not yet reached the fatal diameter) than when the cancer is diagnosed in the absence of screening. **Model G-E** is an event-driven continuous time state transition model. Based on birth cohort-specific incidence curves, the date at which progressive *breast cancer will appear clinically* (if ever) is sampled, and the stage, ER and HER2 are then sampled based on age-period specific stage distributions for these parameters. A *sojourn time* is sampled from an age-specific distribution and the beginning of the sojourn period is defined as the clinical incidence date minus the sojourn time. If a screening event takes place during the sojourn period, it may detect the tumor with probability equal to the age-specific mammogram sensitivity. If the tumor is *screen detected*, a stage at detection is sampled from a probability distribution calculated from the observed lead time, the distributions of dwell times in the clinical stages, and the stage at the clinical detection date. Whether clinically or screen detected, treatment is sampled from an age-stage-ER-HER2-period specific distribution of possible treatment regimens. Each particular treatment regimen reduces the hazard of breast cancer mortality by a ratio that depends on age and stage at diagnosis, ER, and HER2. The date of *breast cancer death* (which may turn out to be after the date of death from other causes) is then sampled from the corresponding age-stage-ER-HER2-treatment regimen-specific survival function. Simulated women who do not have progressive breast cancer may have limited malignant potential (LMP) breast cancer. LMP breast cancer is modeled as never being clinically detected, and is never fatal. But it is screen-detectable for five years and, if screen-detected, its stage is always DCIS. These screen-detected LMP DCIS are then treated the same way as progressive breast cancer diagnosed during the DCIS stage, but treatment has no effect on mortality because these LMP tumors are never fatal. **Model W** is a discrete-event, stochastic tumor growth simulation model. It simulates the natural history of breast-cancer using a continuous time growth model for tumor size and a Poisson process for tumor extent with a randomly assigned growth rate from a population level distribution. In the model breast cancer is assumed a progressive disease arising in the in situ stage. Model W further assumes a fraction of all tumors have "limited malignant potential" (LMP). This subtype is non-lethal, limited in size and stage to in situ and early localized disease and is predominately detected by screening mammography. If undetected for a fixed dwell period they are assumed to regress. Breast cancer can be detected by one of two methods: breast imaging (*screen-detected*), or *by symptoms*, where the likelihoods of detection are functions of a woman's age and tumor size. Upon detection, a woman will receive standard treatment and depending on calendar year, and woman and tumor level characteristics, may also receive adjuvant treatment. Treatment effectiveness, a function of treatment type, is independent of the method of detection and is modeled as a "cure/no-cure" process.

Chapter 8

Benefits and harms of mammography screening after age 74 years: estimates of overdiagnosis

Nicolien T. van Ravesteyn, Natasha K. Stout, Clyde B. Schechter, Eveline A.M. Heijnsdijk, Oguzhan Alagoz, Amy Trentham-Dietz, Jeanne S. Mandelblatt, Harry J. de Koning

Submitted

ABSTRACT

Purpose: To quantify the benefits and harms of mammography screening for women beyond age 74.

Method: Three previously validated microsimulation models were used to simulate a cohort of women born in 1960. All women received biennial screening starting at age 50 with cessation ages varying from 74 up to 96. We estimated the number of life-years gained (LYG), quality-adjusted life-years, breast cancer deaths averted and number of overdiagnosed women per 1,000 screens.

Results: The models predicted LYG per 1,000 screens with no upper age limit, but the gains decreased with advancing age from 8-11 LYG at age 74 (range across models) to 5-8 at age 80 and 1-2 at age 90. When adjusted for quality-of-life decrements, the LYG remained positive, but decreased by 5-13% at age 74 and 11-22% at age 80. At age 90-92 years, all LYG were counterbalanced by a loss in quality of life, mainly due to the increasing number of overdiagnosed breast cancers; 1-5 per 1,000 screens at age 74, 2-6 at age 80 and 4-8 at age 90. The age at which harms began to outweigh benefits shifted to a younger age when larger or longer utility losses due to a breast cancer diagnosis were assumed.

Conclusion: Screening benefits outweigh harms at age 74, but the balance becomes less favorable with advancing age. At age 90, harms outweigh benefits, largely as a consequence of overdiagnosis. This age was the same across the three models, despite important model differences in assumptions on ductal carcinoma *in situ*.

INTRODUCTION

Mammography screening has been shown to be effective in reducing breast cancer mortality in randomized trials and nationwide screening programs in women aged 50-74 years.[1-3] Benefits and harms of screening mammography in women aged 74 years and older are less well established and surrounded by uncertainty, because none of the randomized, controlled trials designed to evaluate screening mammography included sufficient numbers of women aged 74 years and older.

There are several factors that might influence the balance between benefits and harms of mammography screening in women older than 74 years. Because the incidence of breast cancer increases with age[4,5] and sensitivity is higher in the older age groups,[6] the benefits of screening may be larger for older than for younger women. On the other hand, the benefits of screening might be limited due to the higher death rate from competing causes with advancing age. Moreover, the harms of screening, in particular overdiagnosis, may on average increase with age due to the shorter remaining life-expectancy.

The extent to which overdiagnosis occurs is uncertain and widely debated, reflected in the wide range of estimates of up to 54% that have been published.[7-14] Overdiagnosis is generally defined as 'the detection of tumors that would not have been detected in a woman's lifetime in the absence of screening'. The difficulties associated with estimating the amount of overdiagnosis are reflected in this definition; once a screening program has been initiated it is impossible to know what would have happened in the absence of screening. An effective method to address this issue is to use microsimulation models, which represent tumor growth and/or transitions among cancer states and evaluate screening effects based on the synthesis of detailed data.

Despite the uncertainty around the benefits and harms, many women aged 75 years and older are being screened in the United States (U.S.). A recent study found that 62% of women aged 75 to 79 years and 50% of women over 80 years-old reported receiving a mammogram in the past two years.[15] Several studies found that a physician recommendation is a strong determinant of mammography screening[15-17] and 70%-86% of primary care physicians would recommend mammography for a 80-year old healthy woman.[18]

There is, however, no consensus in the U.S. on whether or not to recommend screening for women beyond age 74. For example, the American Cancer Society recommends mammography screening as long as women are in good health (without an upper age),[19] while the U.S. Preventive Services Task Force (USPSTF) recommends screening from age 50 to 74 years.[20] In their most recent recommendation, the USPSTF concluded that 'the current evidence is insufficient to assess the additional benefits and harms of screening mammography in women 75 years or older'.[20]

In the present study, we therefore quantified the additional benefits and harms of continuing mammography screening after age 74 years using three well-established and cross-

validated microsimulation models. Furthermore, we aimed to provide information about the harms and benefits of screening in a meaningful way as previous studies have found that women as well as primary care physicians may have difficulties in understanding cancer screening statistics.[21,22] Therefore, we presented the outcomes in two ways: benefits and harms were presented on the same scale (i.e., as absolute numbers per 1,000 screens), and combined in a single metric: quality-adjusted life-years (QALYs).[23] In particular, we estimated the amount of overdiagnosed invasive breast cancers and ductal carcinoma *in situ* (DCIS) associated with screening women beyond age 74. We also determined the age at which mammography screening no longer resulted in a positive number of QALYs, as a proxy for the age at which the harms began to outweigh the benefits of screening.

METHOD

Model overview

We used three microsimulation models developed as part of the Cancer Intervention and Surveillance Modeling Network (CISNET), which is an international collaborative modeling consortium funded by the National Cancer Institute. The three models were: model E (Erasmus MC, Erasmus University Medical Center, Rotterdam, the Netherlands), model G-E (Georgetown University Medical Center, Washington, DC, and Albert Einstein College of Medicine, Bronx, New York), and model W (University of Wisconsin, Madison, Wisconsin, and Harvard Medical School, Boston, Massachusetts). The models have been described in detail elsewhere[24-26] and information about the models can be found online (http://cisnet.cancer.gov/). Briefly, the models simulated life histories for individual women. After estimating breast cancer incidence and mortality in the absence of screening and treatment, the models overlaid screening use and improvements in survival associated with treatment advances. A schematic representation of the influence of breast cancer screening on (simulated) life histories is shown in Figure 8.1.

All three models include DCIS with three different types of preclinical DCIS: regressive DCIS, DCIS that is diagnosed clinically, and DCIS that progresses to invasive disease. Model W also considers some cases of small invasive cancer that are non-progressive and have limited malignant potential (LMP). In model E, age-specific transition rates of DCIS becoming invasive or clinically diagnosed and DCIS regression rates were estimated using incidence data by stage, age and calendar year (1975-1999) from the Surveillance, Epidemiology and End Results (SEER) Program. In model G-E, all invasive cancers are assumed to begin as DCIS. In addition, the model includes non-progressive lesions (NPL) which are screen-detectable for a period of five years, and if detected appear as DCIS. If not screen-detected during that window, the lesions regress. The proportion of NPL was interval-estimated by fitting results to SEER incidence and mortality rates from 1975-2000. For this analysis, to be conservative,

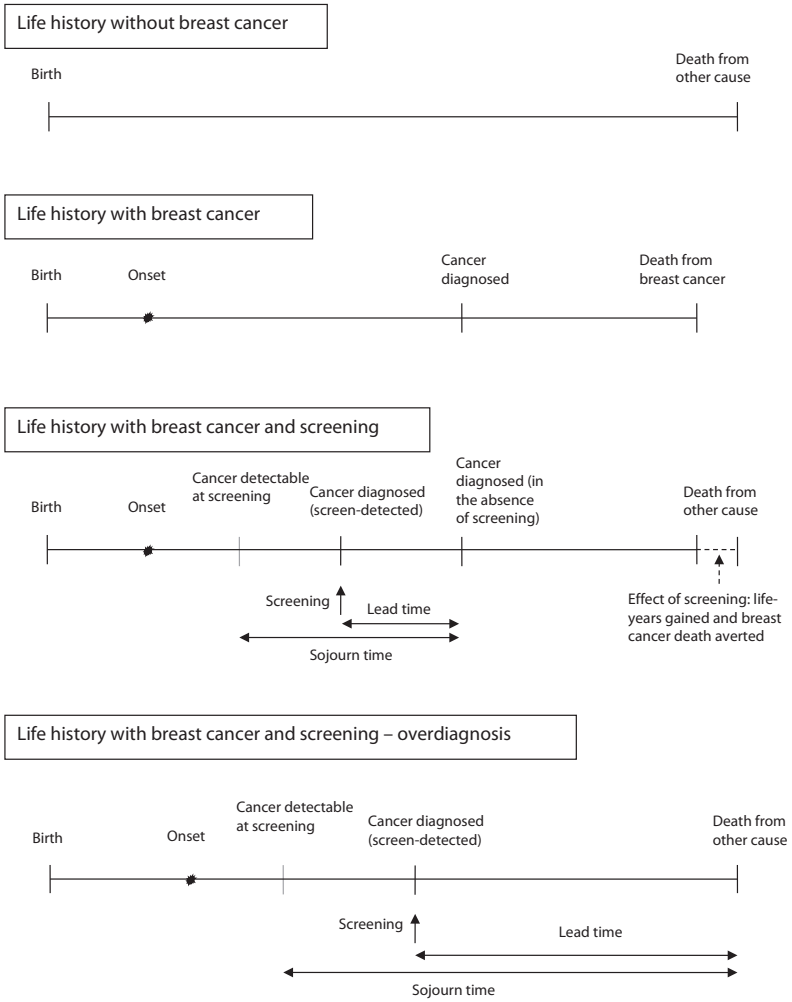


Figure 8.1. Schematic overview of simulated life histories and effect of screening. Sojourn time is the duration of the preclinical, screen-detectable phase of the tumor and lead time is the interval from screen detection to the time of clinical diagnosis, when the tumor would have been diagnosed without screening. If the tumor is screen-detected without a clinical diagnosis in the absence of screening, the detection represents overdiagnosis. Lead time represents additional years that are lived with breast cancer due to screening

the lower end of the interval of values was used. In Model W, breast cancer disease progression is based on a Gompertzian tumor growth model paired with a Poisson process for spread to lymph nodes. All natural history parameters including the proportion of LMP tumors were calibrated such that incidence trends matched SEER data from 1975-1999. All models have been previously validated and adequately reproduced SEER age-specific incidence (Figure 8.2) and mortality rates over time.[24-27]

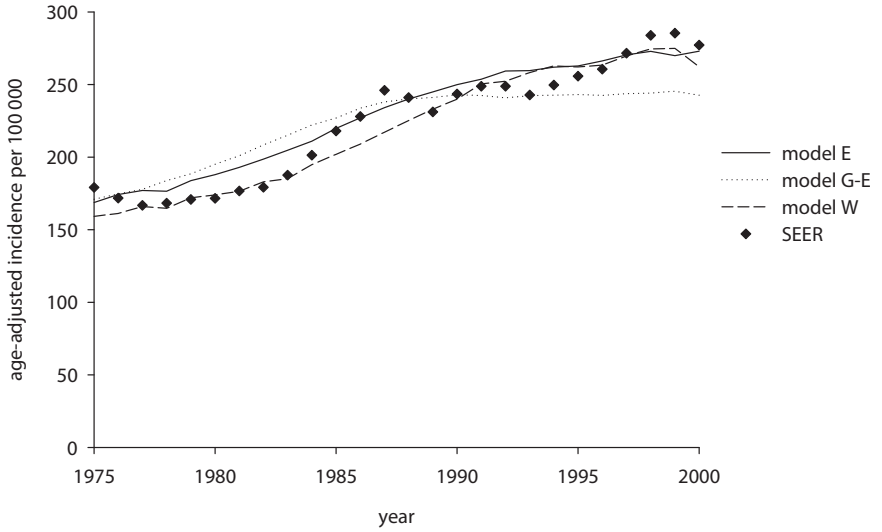


Figure 8.2. Age-adjusted breast cancer incidence rates from 1975-2000 predicted by the models vs. reported to SEER for women aged 30-79 years

Model parameters

We used a common set of age-specific model inputs[28] for breast cancer incidence,[29] breast cancer specific survival,[30] and competing non-breast cancer causes of death.[31] A cohort of women born in 1960 was simulated and followed throughout their entire lifetime. We assumed 100% adherence with screening and adjuvant treatment guidelines. In addition, we applied quality-of-life decrements to women undergoing a mammogram and diagnostics[32], and life-years with breast cancer by stage of disease at diagnosis[33] (Table 8.1) to estimate QALYs.

Table 8.1. Utility values and durations of different health states used in the simulation models of breast cancer

State	Utility	1-utility	Duration	Source
screening attendance	0.994	0.006	1 week	de Haes[32]
diagnostic phase	0.895	0.105	5 weeks	de Haes[32]
breast cancer by stage of disease at diagnosis				
local or DCIS	0.90	0.10	2 years	Stout[33]
regional	0.75	0.25	2 years	Stout[33]
distant	0.60	0.40	until death	Stout[33]

Screening scenarios

Biennial screening with 100% adherence started at age 50 with varying cessation ages of screening. First, we simulated screening policy the currently recommended by the USPSTF (biennial screening between age 50 and 74 years) and assessed the benefits and harms per 1,000 women alive at age 50, followed until death. We then determined the benefits and harms of the last screen (at age 74 years) and the additional benefits and harms of adding one more screen after the last screening test were estimated for increasing stopping ages of up to 96 years. We estimated the number of life-years gained (LYG), QALYs, breast cancer deaths averted, false positive exams and number of overdiagnosed women for each screening scenario per 1,000 screens. The number of overdiagnosed women was calculated as the difference in the predicted number of diagnosed women in the presence of screening and the predicted number of diagnosed women in the absence of screening.

Sensitivity analysis

To evaluate how the results were influenced by certain assumptions and parameter values, we performed several sensitivity analyses. In particular, we assessed the effect of using different utility decrements associated with a breast cancer diagnosis for DCIS and local disease as used in previous studies[34,35] (0.05 0.15, and 0.20, instead of 0.10). In addition, we assessed the effect of different durations for the utility decrements (5 years instead of 2 years for the effect of diagnosis at the DCIS, local or regional state).

RESULTS

The models estimated that if 1,000 50-year-old women underwent biennial screening from age 50 to 74 years, between 8.0 and 8.9 breast cancer deaths were prevented, depending on the model, and there were 132-142 LYG (Table 8.2). The models predicted that in the absence of screening there would be 139-154 breast cancers diagnosed. In the presence of screening this increased to 151-170 diagnoses. Thus, among these 1,000 women, 6-33 women were diagnosed in the presence of screening, but would not have been diagnosed in the absence of screening, and were thus overdiagnosed.

Extending screening beyond age 74 resulted in a steep increase in the number of overdiagnosed women (Figure 8.3A-C). The overdiagnosed breast cancers (DCIS + invasive) increased from 1.2-5.0 per 1,000 screens at age 74, 1.8-6.0 at age 80 and 3.7-7.5 at age 90.

If overdiagnosis is expressed as a percentage of screen-detected cancers it also increased steeply in all three models (Table 8.3). Screening women between ages 50 and 74 results in 5%-32% (range between the models) of the invasive breast cancers that were screen-detected being overdiagnosed, increasing to 14%-36% for a screen at age 80 and 28%-41% for a screen at age 90 (Table 8.3). For DCIS the percentages were higher in all three models,

Table 8.2. Benefits and harms of biennial mammography screening age 50-74, assuming 100% attendance

Outcomes per 1000 women alive at age 50	Model E	Model G-E	Model W
Number of mammograms	11,151	11,337	11,117
Number of breast cancers	151	159	170
Screen detected BCs (% DCIS)	84 (26%)	81 (19%)	95 (30%)
Life-years gained (per 1000 screens)	136 (12.2)	142 (12.6)	132 (11.8)
QALYs (per 1000 screens)	132 (12)	135 (12)	119 (11)
Reduction in LYG after adjustment for QoL	3%	6%	9%
BC deaths averted (per 1000 screens)	8.9 (0.8)	8.0 (0.7)	8.1 (0.7)
False positives (per 1000 screens)	865 (78)	1030 (91)	915 (82)
Overdiagnosis DCIS (per 1000 screens)	9.2 (0.8)	2.7 (0.2)	11.4 (1.0)
Overdiagnosis invasive BCs (per 1000 screens)	3.0 (0.3)	3.2 (0.3)	21.7 (2.0)
Overdiagnosis Total (per 1000 screens)	12.2 (1.1)	5.9 (0.5)	33.0 (3.0)
% of all BCs detected at ages 50+*	8%	4%	19%
% of BCs detected during screening #	11%	6%	27%
% of screen-detected BCs^	14%	7%	35%

* Number of excess cancers as a proportion of cancers diagnosed from age 50 to death

Number of excess cancers as a proportion of cancers diagnosed during the screening period (between age 50 and 74)

^ Number of excess cancers as a proportion of screen-detected cancers

Abbreviations
BC: breast cancer
DCIS: ductal carcinoma in situ
QALYs: quality-adjusted life-years
QoL: quality of life

and increased from 18%-41% for screening between age 50-74 to 35%-72% for a screen at age 80 and 53%-91% for a screen at age 90.

The models predicted that screening beyond age 74 resulted in benefits in terms of breast cancer deaths averted and LYG with no upper age limit, but the number of breast cancer deaths averted and LYG per 1,000 screens steadily declined with increasing age (Figure 3D-E). Screening women at age 74 results in 7.8-11.4 LYG per 1,000 screens, which decreased to 4.8-7.8 for a screen at age 80 and 1.4-2.4 for a screen at age 90.

The number of QALYs gained decreased steadily with increasing age from 7.1-9.9 at age 74 to 4.0-6.1 at age 80, and 2.4-3.7 at age 84 years. QALYs were still positive for screening up to age 90 (Figure 8.3F). The number of QALYs gained became negative at age 90 in models E and W and at age 92 in model G-E. In other words, at age 90-92 years, all LYG were counterbalanced by a loss in quality of life due to undergoing a screening test and diagnostics and additional life-years with disease which would have been spent in a healthy state in the absence of screening. The percentages of LYG that were counterbalanced by losses in quality of life increased steeply with increasing age at screening (Figure 8.3G).

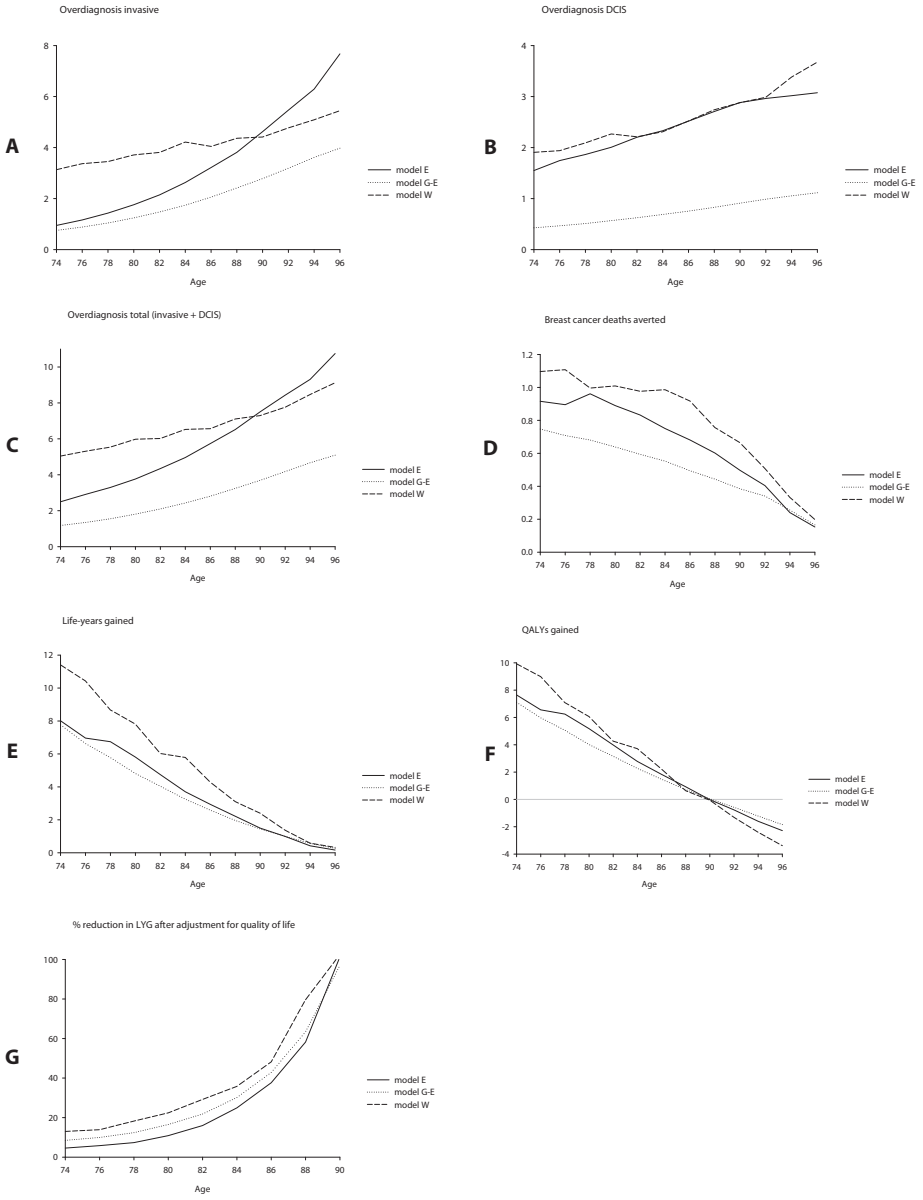


Figure 8.3. Benefits and harms of adding one screen after age 74 – number of excess invasive cancers per 1,000 screens (panel A), number of excess DCIS per 1,000 screens (panel B), number of excess total breast cancers per 1,000 screens (invasive + DCIS; panel C), number of breast cancer deaths averted per 1,000 screens (panel D), number of life-years gained per 1,000 screens (panel E), number of QALYs gained per 1,000 screens (panel F), and relative reduction in LYG after adjustment for quality of life (%) (panel G)

Table 8.3. Percentage of screen-detected breast cancers (invasive and DCIS) that are overdiagnosed^ by screening age and model

Stage at diagnosis	Screening age*	Model E	Model G-E	Model W
Invasive	50-74	5%	5%	32%
	74	11%	9%	35%
	80	17%	14%	36%
	90	37%	28%	41%
DCIS	50-74	41%	18%	40%
	74	61%	27%	47%
	80	72%	35%	52%
	90	91%	53%	60%
Total	50-74	14%	7%	35%
	74	22%	12%	39%
	80	29%	17%	41%
	90	48%	32%	47%

^ Number of excess cancers as a proportion of cancers detected at screening

* All screens with 100% attendance

Screening age 50-74: biennial screening starting at age 50 and ending at age 74. The percentage includes all excess breast cancers detected at screening between age 50 and 74 divided by all screen-detected breast cancers between age 50 and 74.

Screening age 74, 80, 90: all women have been screened biennially up to the screening age. The percentage includes all excess breast cancers detected at screening at age 74, 80, and 90 divided by all screen-detected breast cancers at age 74, 80, and 90, respectively

Sensitivity analysis

If a larger utility loss due to a breast cancer diagnosis is assumed for DCIS and local disease, the age at which QALYs become negative is shifted to a younger age in all three models. For instance, when the utility loss is 0.2 instead of 0.1, the age at which the number of QALYs gained become negative shifts to age 86 in model W and to age 88 in models E and G-E (Appendix Figure 8.1A). If instead of a 2-year duration, a 5-year duration for the utility decrements is assumed, the age shifts to age 86 in all three models (Appendix Figure 8.1B).

DISCUSSION

The model results were very consistent in estimating the age at which the harms began to outweigh the benefits of mammography screening. At age 90-92 years, all LYG were counterbalanced by a loss in quality of life, mainly due to the increasing amount of overdiagnosis. The consistency between models was remarkable, since the models included different assumptions on the natural history of DCIS.

Despite the consistency of our results some limitations have to be considered. We estimated the benefits and harms for a cohort of women born in 1960, who on average, have a remaining life expectancy at age 74 of 13-14 years. If life-expectancy for older women continues to increase in the future, then we might have underestimated the benefits and overestimated the harms of screening. In addition, life-expectancy varies by health status. Benefits are expected to be larger for healthy than for non-healthy women. Future research could evaluate the potential value of personalizing screening for older women if we were better able to predict life-expectancy.

To calculate QALYs we used utility values, i.e., we attached weights to certain health states. We found that adjusting for quality of life (thus using QALYs instead of LYG) has only a small effect for ages 50-74, but an increasing effect at older ages due to the increasing amount of overdiagnosis, and hence additional life-years with disease. In addition, we found that the age at which harms began to outweigh benefits was sensitive to the utility values, and shifted to a younger age when a larger disutility of the disease state or a longer duration for the utility decrements was assumed. This emphasizes the need for validated data on patient's utilities and durations for specific breast cancer disease states. The advantages and disadvantages of using utilities and QALYs have been widely discussed.[36] For the present study, the most important drawback is that individual preferences might diverge from the assumed values. For those women, looking at the benefits and harms per 1,000 screens might be more informative than looking at QALYs.

Our results on QALYs gained are largely in line with what has been previously reported. For example, a previous study found no reduction in the number of QALYs gained as the upper age of screening increases when optimistic assumptions about the preclinical durations were made. However, when pessimistic assumptions were made, the QALYs gained became negative when screening was continued beyond age 80 years.[37] The estimates from the present study are in between those from the optimistic and pessimistic scenarios.[37]

The amount of overdiagnosis has been the topic of intense debate, partly due to methodological issues. Overdiagnosis is overestimated when calculations are derived from the implementation period of a screening program and when there is insufficient follow-up to observe a reduction in breast cancer incidence.[8] Similarly, the range of overdiagnosis estimates is considerably smaller (1% to 10%) when only studies that adequately adjust for lead time and changes in breast cancer risk are included.[38]

The models estimated that 4%-19% of all breast cancers detected in women age 50 and over are overdiagnosed for biennial screening from age 50 to 74. A recent study on overdiagnosis in the U.S. estimated that 31% of breast cancers diagnosed in 2008 were overdiagnosed.[14] This estimate cannot be directly compared to the estimates presented here, as we estimated overdiagnosis for specific screening scenarios and not for the screening as observed in the U.S. The estimates from the models will, however, likely be lower, mainly

because the models incorporate an age-period-cohort model which incorporates a stronger increase in background incidence over time.[29]

The model results showed a large range in overdiagnosis estimates. For invasive disease, one model (model W) estimated markedly higher overdiagnosis than the other two models up to age 86 years. This difference between models is due to the fact that model W includes a subset of small invasive cancers with limited malignant potential. These cancers are assumed to grow only to a limited size and then disappear.

There was also a large difference in the predicted amount of overdiagnosis of DCIS between models, which likely reflects the continued uncertainty about DCIS natural history.[39] Little is known on the natural history of DCIS, because DCIS is usually removed as soon as it is detected. There is evidence for progression of DCIS from studies in women with low grade DCIS which is initially mistakenly diagnosed as benign. These studies report that 14%–60% of those women develop invasive cancer within 10–20 years.[40-42] There is, however, also evidence that not all DCIS become invasive, for example from autopsy studies that found a prevalence of DCIS of 0-15% in women not known to have had breast cancer.[43] Our results do not provide additional information on the natural history of DCIS, since all three models adequately replicated incidence trends, despite differences in the assumed natural history of DCIS. This finding is in line with a previous modeling study that found that two alternative models with extreme assumptions on progression and regression rates of DCIS fit the observed breast cancer incidence in the Netherlands equally well.[44]

The models estimate that at age 90 years, 53-91% of the screen-detected DCIS are overdiagnosed. In other words, for every 1,000 screens performed at age 90 years, 1 to 3 women are overdiagnosed with DCIS. Those women would not have been diagnosed with breast cancer in the absence of screening, but are diagnosed in the presence of screening and will probably also undergo treatment for their disease. It has been found that almost all women (97.5%) diagnosed with DCIS undergo a surgical procedure.[45] In addition, 61% of women diagnosed with DCIS receive radiotherapy,[46] and 47% receive adjuvant hormonal therapy.[47] Although older women tend to receive less aggressive treatment than younger women,[48] older women undergoing treatment may be exposed to more toxicity than younger women.[49] It is important to inform women about DCIS, as only 6% of U.S. women reported having heard of DCIS, but when informed, 60% wanted to take DCIS into account when deciding about mammography, and many indicated that the probability of DCIS becoming invasive would influence their treatment decision.[21] Future research on the biological behavior and predictors of risk for developing invasive disease is needed to be able to know which women need treatment and which women can forego treatment.[50]

In summary, the balance between benefits and harms of mammography becomes less favorable beyond age 74, due to the increasing amount of overdiagnosis. Beyond age 90 harms outweigh benefits. An upper age limit of breast cancer screening, therefore, seems appropriate. The appropriate upper age for an individual woman depends on the weight she

attaches to specific benefits and harms. From a societal perspective the willingness-to-pay for a QALY may also need to be taken into account. If we were better able to distinguish between subtypes of DCIS that progress and those that do not, harm from treating non-progressive disease can be prevented.

Acknowledgements

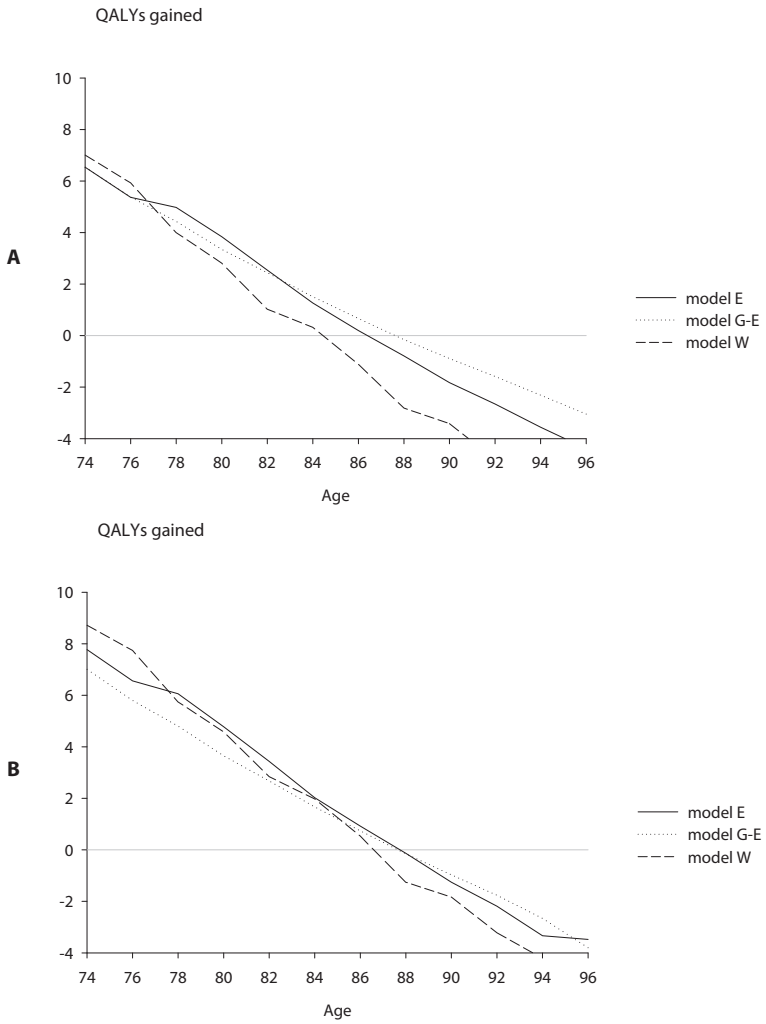
Data collection and sharing was supported by the National Cancer Institute-funded Breast Cancer Surveillance Consortium (U01CA63740, U01CA86076, U01CA86082, U01CA63736, U01CA70013, U01CA69976, U01CA63731, U01CA70040, HHSN261201100031C). A list of the BCSC investigators and procedures for requesting BCSC data for research purposes are provided at: http://breastscreening.cancer.gov/.

REFERENCES

1. Nystrom L, Andersson I, Bjurstam N, Frisell J, Nordenskjold B, Rutqvist LE. Long-term effects of mammography screening: updated overview of the Swedish randomised trials. *Lancet* 2002;359:909-19.
2. Tabar L, Vitak B, Chen HH, Duffy SW, Yen MF, Chiang CF, et al. The Swedish Two-County Trial twenty years later. Updated mortality results and new insights from long-term follow-up. *Radiol Clin North Am* 2000;38:625-51.
3. Otto SJ, Fracheboud J, Looman CW, Broeders MJ, Boer R, Hendriks JH, et al. Initiation of population-based mammography screening in Dutch municipalities and effect on breast-cancer mortality: a systematic review. *Lancet* 2003;361:1411-7.
4. Ferlay J, Shin H, Bray F, Forman D, Mathers C, Parkin D. GLOBOCAN 2008 v1.2, Cancer Incidence and Mortality Worldwide: IARC CancerBase No. 10 Lyon, France: International Agency for Research on Cancer, 2010.
5. SEER*Stat Database: Incidence - SEER 18 Regs Research Data + Hurricane Katrina Impacted Louisiana Cases. Linked To County Attributes - Total U.S., 1969-2010 Counties, National Cancer Institute, DCCPS, Surveillance Research Program, Surveillance Systems Branch, released April 2012, based on the November 2011 submission.
6. Sinclair N, Littenberg B, Geller B, Muss H. Accuracy of screening mammography in older women. *AJR Am J Roentgenol* 2011;197:1268-73.
7. Biesheuvel C, Barratt A, Howard K, Houssami N, Irwig L. Effects of study methods and biases on estimates of invasive breast cancer overdetection with mammography screening: a systematic review. *Lancet Oncol* 2007;8:1129-38.
8. de Gelder R, Heijnsdijk EA, van Ravesteyn NT, Fracheboud J, Draisma G, de Koning HJ. Interpreting overdiagnosis estimates in population-based mammography screening. *Epidemiol Rev* 2011;33:111-21.
9. Jorgensen KJ, Gotzsche PC. Overdiagnosis in publicly organised mammography screening programmes: systematic review of incidence trends. *BMJ* 2009;339:b2587.
10. Jorgensen KJ, Keen JD, Gotzsche PC. Is mammographic screening justifiable considering its substantial overdiagnosis rate and minor effect on mortality? *Radiology* 2011;260:621-7.
11. Kalager M, Adami HO, Bretthauer M, Tamimi RM. Overdiagnosis of invasive breast cancer due to mammography screening: results from the norwegian screening program. *Ann Intern Med* 2012;156:491-9.
12. Kopans DB, Smith RA, Duffy SW. Mammographic screening and "overdiagnosis". *Radiology* 2011;260:616-20.
13. Duffy SW, Tabar L, Olsen AH, Vitak B, Allgood PC, Chen TH, et al. Absolute numbers of lives saved and overdiagnosis in breast cancer screening, from a randomized trial and from the Breast Screening Programme in England. *J Med Screen* 2010;17:25-30.
14. Bleyer A, Welch HG. Effect of three decades of screening mammography on breast-cancer incidence. *N Engl J Med* 2012;367:1998-2005.
15. Bellizzi KM, Breslau ES, Burness A, Waldron W. Prevalence of cancer screening in older, racially diverse adults: still screening after all these years. *Arch Intern Med* 2011;171:2031-7.
16. Fox SA, Murata PJ, Stein JA. The impact of physician compliance on screening mammography for older women. *Arch Intern Med* 1991;151:50-6.
17. Meissner HI, Breen N, Taubman ML, Vernon SW, Graubard BI. Which women aren't getting mammograms and why? (United States). *Cancer Causes Control* 2007;18:61-70.
18. Leach CR, Klabunde CN, Alfano CM, Smith JL, Rowland JH. Physician over-recommendation of mammography for terminally ill women. *Cancer* 2012;118:27-37.
19. Smith RA, Cokkinides V, Brawley OW. Cancer screening in the United States, 2012: A review of current American Cancer Society guidelines and current issues in cancer screening. *CA Cancer J Clin* 2012;62:129-142.
20. U.S. Preventive Services Task Force. Screening for breast cancer: U.S. Preventive Services Task Force recommendation statement. *Ann Intern Med* 2009;151:716-26, W-236.
21. Schwartz LM, Woloshin S, Sox HC, Fischhoff B, Welch HG. US women's attitudes to false positive mammography results and detection of ductal carcinoma in situ: cross sectional survey. *BMJ* 2000;320:1635-40.
22. Wegwarth O, Schwartz LM, Woloshin S, Gaissmaier W, Gigerenzer G. Do physicians understand cancer screening statistics? A national survey of primary care physicians in the United States. *Ann Intern Med* 2012;156:340-9.
23. Harris R, Sawaya GF, Moyer VA, Calonge N. Reconsidering the criteria for evaluating proposed screening programs: reflections from 4 current and former members of the U.S. Preventive services task force. *Epidemiol Rev* 2011;33:20-35.
24. Fryback DG, Stout NK, Rosenberg MA, Trentham-Dietz A, Kuruchittham V, Remington PL. The Wisconsin Breast Cancer Epidemiology Simulation Model. *J Natl Cancer Inst Monogr* 2006:37-47.

25. Mandelblatt J, Schechter CB, Lawrence W, Yi B, Cullen J. The SPECTRUM population model of the impact of screening and treatment on U.S. breast cancer trends from 1975 to 2000: principles and practice of the model methods. *J Natl Cancer Inst Monogr* 2006:47-55.
26. Tan SY, van Oortmarssen GJ, de Koning HJ, Boer R, Habbema JD. The MISCAN-Fadia continuous tumor growth model for breast cancer. *J Natl Cancer Inst Monogr* 2006:56-65.
27. Berry DA, Cronin KA, Plevritis SK, Fryback DG, Clarke L, Zelen M, et al. Effect of screening and adjuvant therapy on mortality from breast cancer. *N Engl J Med* 2005;353:1784-92.
28. Mandelblatt JS, Cronin KA, Bailey S, Berry DA, Koning H J de, Draisma G, et al. Effects of mammography screening under different screening schedules: model estimates of potential benefits and harms *Ann Int Med* 2009;151:738-47.
29. Holford TR, Cronin KA, Mariotto AB, Feuer EJ. Changing patterns in breast cancer incidence trends. *J Natl Cancer Inst Monogr* 2006:19-25.
30. Cronin KA, Mariotto AB, Clarke LD, Feuer EJ. Additional common inputs for analyzing impact of adjuvant therapy and mammography on U.S. mortality. *J Natl Cancer Inst Monogr* 2006:26-9.
31. Rosenberg MA. Competing risks to breast cancer mortality. *J Natl Cancer Inst Monogr* 2006:15-9.
32. de Haes JC, de Koning HJ, van Oortmarssen GJ, van Agt HM, de Bruyn AE, van Der Maas PJ. The impact of a breast cancer screening programme on quality-adjusted life-years. *Int J Cancer* 1991;49:538-44.
33. Stout NK, Rosenberg MA, Trentham-Dietz A, Smith MA, Robinson SM, Fryback DG. Retrospective cost-effectiveness analysis of screening mammography. *J Natl Cancer Inst* 2006;98:774-82.
34. Mandelblatt JS, Schechter CB, Yabroff KR, Lawrence W, Dignam J, Muennig P, et al. Benefits and costs of interventions to improve breast cancer outcomes in African American women. *J Clin Oncol* 2004;22:2554-66.
35. Wong IO, Kuntz KM, Cowling BJ, Lam CL, Leung GM. Cost effectiveness of mammography screening for Chinese women. *Cancer* 2007;110:885-95.
36. Whitehead SJ, Ali S. Health outcomes in economic evaluation: the QALY and utilities. *Br Med Bull* 2010;96:5-21.
37. Boer R, de Koning HJ, van Oortmarssen GJ, van der Maas PJ. In search of the best upper age limit for breast cancer screening. *Eur J Cancer* 1995;31A:2040-3.
38. Puliti D, Duffy SW, Miccinesi G, de Koning H, Lynge E, Zappa M, et al. Overdiagnosis in mammographic screening for breast cancer in Europe: a literature review. *J Med Screen* 2012;19 Suppl 1:42-56.
39. Morrow M. The certainties and the uncertainties of ductal carcinoma in situ. *J Natl Cancer Inst* 2004;96:424-5.
40. Eusebi V, Feudale E, Foschini MP, Micheli A, Conti A, Riva C, et al. Long-term follow-up of in situ carcinoma of the breast. *Semin Diagn Pathol* 1994;11:223-35.
41. Feig SA. Ductal carcinoma in situ. Implications for screening mammography. *Radiol Clin North Am* 2000;38:653-68, vii.
42. Betsill WL, Jr., Rosen PP, Lieberman PH, Robbins GF. Intraductal carcinoma. Long-term follow-up after treatment by biopsy alone. *JAMA* 1978;239:1863-7.
43. Welch HG, Black WC. Using autopsy series to estimate the disease "reservoir" for ductal carcinoma in situ of the breast: how much more breast cancer can we find? *Ann Intern Med* 1997;127:1023-8.
44. de Gelder R, Fracheboud J, Heijnsdijk EA, den Heeten G, Verbeek AL, Broeders MJ, et al. Digital mammography screening: weighing reduced mortality against increased overdiagnosis. *Prev Med* 2011;53:134-40.
45. Baxter NN, Virnig BA, Durham SB, Tuttle TM. Trends in the treatment of ductal carcinoma in situ of the breast. *J Natl Cancer Inst* 2004;96:443-8.
46. Habel LA, Achacoso NS, Haque R, Nekhlyudov L, Fletcher SW, Schnitt SJ, et al. Declining recurrence among ductal carcinoma in situ patients treated with breast-conserving surgery in the community setting. *Breast Cancer Res* 2009;11:R85.
47. Livaudais JC, Hwang ES, Karliner L, Napoles A, Stewart S, Bloom J, et al. Adjuvant hormonal therapy use among women with ductal carcinoma in situ. *J Womens Health (Larchmt)* 2012;21:35-42.
48. Schonberg MA, Marcantonio ER, Li D, Silliman RA, Ngo L, McCarthy EP. Breast cancer among the oldest old: tumor characteristics, treatment choices, and survival. *J Clin Oncol* 2010;28:2038-45.
49. Crivellari D, Bonetti M, Castiglione-Gertsch M, Gelber RD, Rudenstam CM, Thurlimann B, et al. Burdens and benefits of adjuvant cyclophosphamide, methotrexate, and fluorouracil and tamoxifen for elderly patients with breast cancer: the International Breast Cancer Study Group Trial VII. *J Clin Oncol* 2000;18:1412-22.
50. Partridge AH, Elmore JG, Saslow D, McCaskill-Stevens W, Schnitt SJ. Challenges in ductal carcinoma in situ risk communication and decision-making: report from an American Cancer Society and National Cancer Institute workshop. *CA Cancer J Clin* 2012;62:203-10.

APPENDIX



Appendix Figure 8.1. The number of QALYs gained of adding one screen after age 74 – QALYs gained per 1,000 screens assuming different utility values (panel A) and durations (panel B).

A. QALYs gained - Utility decrement of 0.2 instead of 0.1 for DCIS and local disease

B. QALYs gained - Utility decrements for a breast cancer diagnosis of DCIS, local and regional disease for a duration of 5 years instead of 2 years

Chapter 9

Personalizing age of screening cessation based on comorbidity level: results of collaborative modeling of breast, prostate, and colorectal cancer

Iris Lansdorp-Vogelaar, Roman Gulati, Angela B. Mariotto, Clyde B. Schechter, Tiago M. de Carvalho, Amy B. Knudsen, Nicolien T. van Ravesteyn, Eveline A.M. Heijnsdijk, Marjolein van Ballegooijen, Carolyn M. Rutter, Karen M. Kuntz, Eric J. Feuer, Ruth Etzioni, Harry J. de Koning, Ann G. Zauber*, Jeanne S. Mandelblatt*

** dual senior authors*

Submitted

ABSTRACT

Background: US Preventive Services Task Force guidelines recently recommended cessation of breast and colorectal cancer screening after age 74; prostate cancer screening is no longer recommended. Upper limits on screening ages were based on a steeper rise in harms and decreasing benefits with advancing age. However, heterogeneity in comorbidity might shift the balance of harms and benefits towards cessation at younger or older ages.

Methods: We used seven well-established cancer simulation models and data on average and comorbidity level-specific life expectancy to project the incremental number of screening tests (measure of harms) and life-years gained (measure of benefits) of biennial mammography, biennial prostate-specific antigen testing, and annual fecal immunochemical testing from age 50 to a range of cessation ages from 66 to 90. The screening cessation age by comorbidity level (none, mild, moderate and severe) was defined as the age at which the balance between harms and benefits was as favorable as screening until age 74 among those with average life expectancy.

Results: Model projected screening cessation ages were highly consistent across models and cancer sites. Keeping the same harms-benefit ratio used by the USPSTF for the average population would lead to persons with no, mild, moderate and severe comorbidities being screened until median ages 77, 74, 72, and 68, respectively. Beyond these ages screening harms outweighed the benefits.

Conclusion: The balance of benefits and harms of cancer screening varies considerably by comorbidity level. This study provides clinicians with data to inform personalized decisions about when to stop cancer screening.

INTRODUCTION

The US Preventive Services Task Force (USPSTF) recently recommended that individuals older than age 74 do not need to continue colorectal cancer screening, that there is insufficient evidence to recommend continuation of breast cancer screening if they have been regularly screened prior to this age, and that there is sufficient evidence to recommend against prostate cancer screening after this age.[1-3] These recommendations were based on data indicating that the average gain in life-years associated with extending screening from age 75 to age 85 was small in comparison to the harms. The USPSTF does note that screening might be considered between ages 75 and 85 based on health status, but does not provide clinicians with further information to implement such a personalized approach.

There are over 13 million individuals between ages 75 and 85 in the US, and this number is expected to increase to almost 30 million by 2050.[4] Thus, clinicians will be caring for a large and growing number of individuals affected by the current recommendations. There is great heterogeneity in the health of these individuals. Previous decision analyses looking at health benefits of different cancer screening cessation ages by life expectancy [5,6] have limited clinical utility because they do not provide a framework for determining life expectancy.

In the current analysis, we determine the ages of cessation of breast, prostate, and colorectal cancer screening based on individual comorbidity level using seven well-established, independently developed models from three cancer sites that are part of the Cancer Intervention and Surveillance Modeling Network (CISNET). Although the USPSTF has since recommended against routine prostate cancer screening,[7] we included it in the analyses since many US groups continue to recommend it,[8-10] and we anticipate that prostate cancer screening will continue for the foreseeable future.

METHODS

We used microsimulation models to estimate the incremental benefits and harms of adding one more screen between the ages of 66 and 90 inclusive in cohorts that vary by age and comorbidity level. The balance of incremental benefits and harms for each cohort was compared to that of screening the entire population until age 76 (i.e. the first age screening was no longer recommended by the USPSTF) to define an "optimal" cessation age.

The models

The models used for this analysis are MISCAN-Fadia and SPECTRUM for breast cancer;[11,12] MISCAN-prostate and the FHCRC model for prostate cancer;[13,14] and CRC-SPIN, SimCRC and MISCAN-Colon for colorectal cancer.[15-17] Each model simulates the life histories of a large cohort of individuals from birth to death and tracks underlying disease in the absence

Table 9.1. Summary of model features by cancer site and model

	Breast Cancer		Prostate Cancer		Colorectal Cancer		
	MISCAN-Fadia	SPECTRUM	MISCAN-Prostate	FHCRC Prostate	MISCAN-Colon	CRC-SPIN	SimCRC
Includes pre-cancers [1]	Yes	Yes	No	No	Yes	Yes	Yes
Includes tumor biomarkers [2]	Yes	Yes	Yes	Yes	No	No	No
How treatment affects mortality	Cure fraction	Hazard reduction	Hazard reduction	Hazard reduction	Not explicitly modeled[3]	Not explicitly modeled[3]	Not explicitly modeled[3]
Calibrated to mortality?	No	No	No	No	No	No	No
Calibrated to incidence?	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Factors affecting screening benefits	Cancer diameter shift, age-shift	Stage-shift, age-shift	Cure fraction	Stage-shift, age-shift	Stage-shift, age-shift	Stage-shift, age-shift	Stage-shift, age-shift

[1] Pre-cancers include ductal carcinomas in situ for breast cancer and adenomas for colorectal cancer.
[2] Tumor biomarkers include estrogen receptor / human epidermal growth factor 2 status for breast cancer, and prostate-specific antigen level and Gleason score for prostate cancer
[3] The models use the latest relative survival estimates from SEER to model the probability of dying from colorectal cancer

of screening. These models have previously been applied to inform the USPSTF recommendations for breast and colorectal cancer screening [18,19] and to project outcomes from prostate cancer screening and treatment interventions.[20] Using multiple models per cancer site provides a credible range of results and serves as a sensitivity analysis on the impact of variations in underlying model structure and assumptions (Table 9.1). Briefly, screening extends life through either detection of disease at an earlier stage or a smaller size when it has better survival after treatment than in the absence of screening. Detailed descriptions of each model have been published elsewhere [11-17] and can be found in standardized profiles (available from http://cisnet.cancer.gov/profiles/).

Inputs were standardized across models within cancer site, including test characteristics, screening and follow-up assumptions, treatment distributions, and cancer-specific and other-cause survival. Sources for the model inputs have been described in prior publications.[18-20]

Population

The models simulate US cohorts who are 66-90 years old and alive in 2010 stratified by comorbidity level (none, mild, moderate, or severe). For comparison, we also simulated cohorts aged 74 years and 76 years (75 years for colorectal cancer) with average comorbidity levels. All individuals undergo regular screening starting at age 50 with biennial mammography, biennial prostate-specific antigen (PSA) testing or annual fecal immunochemical testing (FIT), respectively. We chose FIT over colonoscopy for CRC screening to get more precise estimates for CRC screening cessation ages. We made the simplifying assumption that comorbid-

Table 9.2. Overview of comorbidity levels, associated conditions, and life expectancies at age 74

Comorbidity group	Percent of population at age 74	Conditions included	Life expectancy in years at age 74	
			Men	Women
Entire population	100%	All	12.2 yr	14.2 yr
No comorbidity	69%	None	13.1 yr	15.1 yr
Mild comorbidity	2%	History of MI, acute MI, ulcer or rheumatologic disease	12.5 yr	13.1 yr
Moderate comorbidity	12%	(Cardio-)vascular disease; paralysis; diabetes; or combinations of diabetes with MI, ulcer, or rheumatologic disease	11.0 yr	12.4 yr
Severe comorbidity	17%	AIDS; COPD; mild or severe liver disease; chronic renal failure; dementia; congestive heart failure; or combinations of aforementioned diseases not categorized under moderate comorbidity	8.1 yr	9.8 yr

ity influenced non-cancer life-expectancy but not cancer risk or progression, treatment, or cancer-specific survival.

Comorbidity-specific life tables

Non-cancer life-expectancy was based on age-, gender-, and comorbidity-specific life tables developed at the National Cancer Institute (personal communication Cho, Mariotto, et al). Briefly, claims from a random 5% sample of beneficiaries continuously enrolled with Medicare Parts A and B from 1992-2005 without a diagnosis of cancer residing in the Surveillance Epidemiology, and End Results (SEER) areas were used to estimate comorbidity scores based on 16 conditions.[21,22] The Cox proportional hazard method was used to estimate non-cancer age-conditional life tables for each gender and age combination using comorbidity as a covariate. Comorbidity was then grouped into four levels: none, mild, moderate, and severe, each with its own life expectancy at a given age (Table 9.2). We used the weighted average of the comorbidity-specific life tables for the comparator "average" cohorts. We extrapolated beyond the 13 years of available data by assuming that life expectancy converged towards average US rates after that period.

Cancer screening scenarios

For each cohort, we compared projected outcomes with and without attending one more screen at the current age. So, for example, we examined the life-years gained among a cohort of 76-year olds associated with performing an additional mammogram at age 76 vs. stopping mammography at age 74. Diagnostic follow-up was modeled according to current recommendations. We assumed 100% adherence with screening and diagnostic follow-up.

Outcomes

Screening benefits were measured as the incremental life-years gained (LYG) and cancer deaths prevented (CDP) by screening one more time vs. stopping screening. The harms of

screening were expressed as the incremental numbers of screening tests, false-positive screens and overdiagnosed cases (i.e. cancer that would not have been diagnosed in the absence of screening) resulting from screening one more time vs. stopping.

Analysis

For the base case analysis, we used the incremental number needed to screen to gain one life-year (NNS/LYG) as the metric to measure the balance between harms and benefits. The NNS/LYG was calculated as the incremental number of screening tests needed to screen each cohort one more time at the current age divided by the life-years gained from that additional screening. Since the USPSTF recommended screening through age 74 for the entire population, we assumed that the NNS/LYG at age 74 yields an acceptable balance of harms and benefits, but that the NNS/LYG for screening at age 76 (75 for colorectal cancer) exceeds the acceptable balance. The NNS/LYG by comorbidity level should therefore be below this latter balance and we used this balance as our threshold. Thus, we first calculated the NNS/LYG for a 76-year old cohort (75 for colorectal cancer) with average comorbidity. We then calculated the NNS/LYG for all cohorts by comorbidity level. The optimal age of final screening based on comorbidity level was determined by selecting the age of the cohort with the highest NNS/LYG that was at or below the threshold for the entire population at age 76.

Sensitivity analysis

We performed sensitivity analyses to assess the robustness of results to our choice of metric for the balance between benefits and harms by considering other metrics for harms (false-positive tests, overdiagnosed cancers) and benefits (cancer deaths prevented). We also varied our method for extrapolating comorbidity-specific life tables beyond the 13 years of available data by assuming that the hazard ratio between the average life table and the comorbidity-specific life table at the 13th year of observation was maintained until death.

RESULTS

Screening based on "average" comorbidity level

The currently recommended strategy to continue breast cancer screening until age 74 in individuals with average comorbidity levels requires screening 132-173 women (range across models) to gain one year of life compared to stopping at age 72 (Table 9.3). Continuing screening until age 76 required an additional 146-198 women to be screened to gain one year of life. The NNS/LYG for prostate and colorectal cancers were comparable, but the rates of false-positive and overdiagnosed cases are orders of magnitude higher for prostate cancer vs. breast or colorectal cancer screening.

Table 9.3. Incremental benefits and harms of screening in populations with average comorbidity, by cancer site, model, and age

Cancer site / model / age of screening cessation (years)	Incremental harms*			Incremental benefits*		Balance
	Screening tests†	False-positive tests‡	Over-diagnosed cases	Life-years gained	Cancer deaths prevented	NNS/LYG
Breast cancer						
MISCAN-Fadia						
74 years (vs. 72)	1000	79	0.8	7.6	0.9	132
76 years (vs. 74)	1000	77	1.0	6.9	0.9	146
SPECTRUM						
74 years (vs. 72)	1000	96	0.5	5.8	0.7	173
76 years (vs. 74)	1000	96	0.6	5.1	0.7	198
Prostate cancer						
MISCAN-prostate						
74 years (vs. 72)	1000	117	19.7	9.2	1.6	108
76 years (vs. 74)	1000	135	24.4	8.7	1.7	116
FHCRC prostate cancer model						
74 years (vs. 72)	1000	238	14.6	7.5	0.9	136
76 years (vs. 74)	1000	262	16.6	5.8	0.8	177
Colorectal cancer						
MISCAN-Colon						
74 years (vs. 73)	1000	41	0.3	5.0	0.7	201
75 years (vs. 74)	1000	41	0.3	4.6	0.7	218
CRC-SPIN						
74 years (vs. 73)	1000	38	0.0	3.9	0.7	256
75 years (vs. 74)	1000	38	0.0	3.9	0.6	254
SimCRC						
74 years (vs. 73)	1000	38	0.1	4.9	0.8	227
75 years (vs. 74)	1000	38	0.1	4.3	0.7	258

NNS/LYG: Number needed to screen to gain 1 life-year

* Results are per 1,000 individuals screened according to guidelines since age 50

† For colorectal cancer, the number of screening tests includes annual fecal immunochemical tests (FITs). Colonoscopies performed for follow-up of a positive FIT and for surveillance of persons who have had an adenoma removed are not included.

‡ Model assumptions for false-positive test rates were not standardized across the prostate cancer models due to differences in model approach for false-positive test results

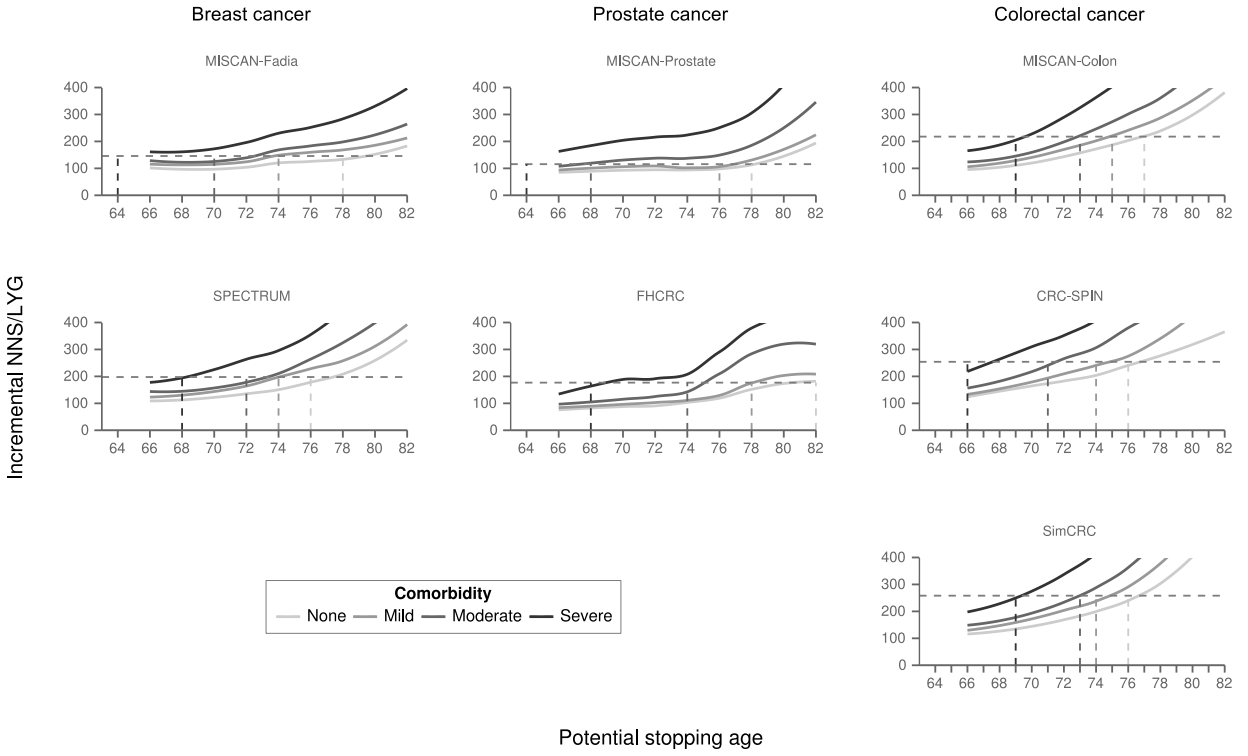


Figure 9.1. Number needed to screen per life-year gained (NNS/LYG) for final screening at the plotted age vs. at the prior screening age, by comorbidity level. The horizontal dotted line represents the NNS/LYG for screening the entire population until age 76 (75 for colorectal cancer). The vertical dotted lines indicate the age until which screening can be continued for each comorbidity group (oldest age for which the NNS/LYG falls under the vertical dotted line). Panels represent different models: MISCAN-Fadia and SPECTRUM for breast cancer; FHCRC prostate model and MISCAN-Prostate for prostate cancer; and CRC-SPIN, MISCAN-Colon, and SimCRC for colorectal cancer.

Screening by comorbidity level

In people with no comorbidities, breast cancer screening through age 74 resulted in NNS/LYG of 117–149, which is lower than that for the entire population (Figure 9.1). As a result, cessation of screening at age 76-78 among women with no comorbidities would yield the same NNS/LYG as seen with cessation of screening at age 74 in the entire population. The same result is obtained for prostate and colorectal cancer screening (Table 9.4).

For the mild comorbidity group, cessation of breast cancer screening at age 74 yields a higher NNS/LYG (146–195 range across models) than in the entire population (Figure 9.1), so that cessation at age 74 would yield a similar NNS/LYG. For prostate and colorectal cancer, screening should be continued until age 74 to 78 to yield a NNS/LYG comparable to the entire population (Table 9.4).

In individuals with moderate comorbidity, screening for all three cancers should cease at a median age of 72 (range 68-74) to result in a similar NNS/LYG as for the entire population screened until age 74. People with severe comorbidities had even higher NNS/LYG at age 74

Table 9.4. Recommended age of screening cessation where NNS/LYG is comparable to that of screening until age 74 in the entire population, among individuals regularly screened from age 50 by comorbidity level, cancer site, and model

Cancer site / model	Age of screening cessation by comorbidity level (years)[1]			
	No comorbidity	Mild comorbidity[2]	Moderate comorbidity[3]	Severe comorbidity[4]
Breast cancer				
MISCAN-Fadia	78 years	74 years	70 years	64 years
SPECTRUM	76	74	72	68
Prostate cancer				
MISCAN-prostate	78	76	68	64
FHCRC prostate model	78	78	74	68
Colorectal cancer				
MISCAN-Colon	77	75	73	69
CRC-SPIN	76	74	71	66
SimCRC	76	74	73	69

[1] Age of screening cessation is the oldest age at which the number needed to screen to gain one life-year (NNS/LYG) is still less than the NNS/LYG for a final screening age of 76 (75 for colorectal cancer) vs. age 74 in the entire population (from Table 3).

[2] Mild comorbidity includes having a life expectancy associated with having a history of MI, acute MI, ulcer, or rheumatologic disease

[3] Moderate comorbidity includes having a life-expectancy associated with having vascular disease, cardiovascular disease, paralysis or diabetes, or combinations of diabetes with MI, ulcer, or rheumatologic disease

[4] Severe comorbidity includes having a life expectancy associated with having AIDS, COPD, mild liver disease, severe liver disease, chronic renal failure, dementia, or congestive heart failure, or combinations of aforementioned diseases not categorized under moderate comorbidity

than the entire population (Figure 9.1), and screening could be discontinued for this group after median age of 68 (range 64–69) (Table 9.4).

Sensitivity analysis

Using false-positive tests or overdiagnosed cancers per LYG as measures of the balance of harms and benefit resulted in similar cessation ages as in the base case analysis (Appendix 9.A, tables 9.A1-9.A2). Using deaths prevented as a measure of benefit resulted in a wider range in recommended final screening ages across the comorbidity groups (Appendix 9.A, tables 9.A3-9.A5). The assumption that life expectancy does not converge after the 13 years of observed data resulted in even earlier stopping ages for those with moderate and severe comorbidity than base results assuming no convergence (Appendix 9.B).

DISCUSSION

This is the first study to employ collaborative modeling to evaluate screening across three cancer sites. It systematically quantifies the balance of benefits and harms of screening older individuals for breast, prostate, and colorectal cancers by comorbidity level. The results are robust across models and cancer sites and indicate that the ages of screening cessation based on comorbidity level differ from recommendations for the entire population. Around 70% of the current US population have no comorbidities at age 74. Our results suggest that this group could continue to be screened until age 77 and still have the same balance of benefits and harms expected in the entire population where screening is recommended from ages 50 to 74. However, screening for these three common cancers should have stopped at age 68 among the 13% of the US population aged 65 to 74 who have severe comorbidity levels.

Our findings are consistent with and extend prior research addressing the upper age limits for cancer screening. For instance, Walter and Covinsky [6] found that screening until around age 60 for those in the lower quartile of population life expectancy has the same number needed to screen to prevent one cancer death as those with the median life expectancy at age 75. They also estimated that screening could be continued to age 85 for those in the upper quartile of life expectancy.[6] This range is consistent with but wider than our model projections of 68 to 77 based on number needed to screen per life-year gained. When we use cancer deaths prevented rather than life-years gained, our ranges were closer to those of Walter and Covinsky, although our maximal upper age limit was still lower since we considered those without comorbidities (70% of the population) and they used the upper quartile (25%) of life-expectancy.

Walter and Covinsky [6] and other previous analyses [5,23-28] provide little guidance on applying this framework in clinical practice, leaving it to clinical judgment to estimate life-expectancy and individualize screening decisions. Several studies have investigated the relationship

between comorbidity level and life expectancy [29-31] but do not address the question of how this relationship influences cancer screening. To date, only two analyses have directly related comorbidity level to cancer screening recommendations. One focused on diabetes-related comorbidities and colorectal cancer screening [32] and one examined cardiovascular disease and breast screening.[5] The current analysis is a multi-model collaborative analysis of the three major screenable cancers.

Although our results are inspired by the USPSTF recommendations, they are certainly just as applicable for alternative screening recommendations. For example, the American Cancer Society also recommends that the decision to stop screening should be individualized based on the potential benefits and risks of screening in the context of overall health status.[33] Recent modelling studies indicate that stopping screening at age 69, raising the PSA threshold for biopsy referral for men above this age, or restricting further screening to men with low comorbidity would reduce overdiagnoses and unnecessary biopsies.[34,35] Our results confirm that the overdiagnosed cases for PSA screening are orders of magnitude higher than for breast or colorectal cancer screening. If PSA screening strategies that minimize harms achieve a sufficiently favorable balance of outcomes, our results advocate tailoring of screening cessation recommendations according to comorbidity.

Overall, our results provide clinicians with data for use in shared decision-making discussions about who might consider continuing screening and for how long. For example, if a clinician is meeting with a regularly-screened 70-year old patient with COPD, our results indicate that this individual falls in the severe comorbidity level and might consider not having any further cancer screening. However, a 76-year old individual with no comorbidities might consider continuing screening until age 77.

The fact that comorbidity-specific conclusions about ages of screening cessation differ meaningfully from those included in clinical guidelines highlights the tension between the need to provide broad public health recommendations for the entire population and the pressure to use a more personalized approach. Our suggested approach of continuing screening in the healthy and earlier cessation in the sickest does not increase the number of screens required in the population, but rather leads to a more efficient allocation of resources, maximizing the benefit and minimizing the harms to the growing older population.[36] The age-, gender-, and comorbidity-specific life expectancies used for these analyses also provide clinicians and the general screening-eligible population with a foundation for discussing preferences for benefits and harms, facilitating individual decision-making.

Despite the innovation and strengths of our approach, there are several caveats that should be considered in evaluating our results. First, we chose the number needed to screen to gain one life-year as our primary metric to quantify the balance of harms and benefits since this was the preferred metric in our previous work for the USPSTF.[18,19] Conclusions were similar when we examined other metrics. We did not explicitly consider complications from screening and diagnostic follow-up as harms, but these should be proportional to the

number of (false-positive) screening tests. Costs per quality-adjusted life-year gained are a common metric used in many countries, but this is not widely accepted in the US.[37-41] Second, we assumed that comorbidity level only influenced life expectancy and not cancer risk, biology or harms. Health conditions such as diabetes are known to be associated with obesity and other lifestyle factors [42] which, in turn, can be associated with improved mammography performance [43] and elevated cancer risk.[44-46] Conversely, adverse events of screening, such as perforations with colonoscopy, are also associated with comorbidities.[47] In the future it will be important to extend our work to capture the known impact of specific comorbidities on other model parameters. For now, competing non-cancer mortality is the single most germane parameter in screening decisions for the oldest age groups, so that our conclusions should be robust.[26]

The next limitation is that we chose to model 100% adherence starting at age 50 to demonstrate how current screening recommendations could be adapted based on comorbidity level. In general, optimal stopping ages are higher in individuals that are unscreened or have skipped previous screening rounds, because they have a higher risk of prevalent cancer. Furthermore, the models used life tables based on non-cancer cases and, therefore, do not include cancer-specific mortality for cancers other than the one targeted by screening. This underestimates the true rate of competing other-cause mortality and therefore the harm to benefit ratios, but does not affect our internal comparisons of comorbidity level groups to the entire population. Finally, we did not consider situations in which the comorbidity level decreased (e.g., from severe to moderate level) after the age of screening cessation. Given the chronic nature of the comorbid conditions in older age, this is a reasonable assumption.

Overall, the results across models and cancer sites were very robust and strongly suggest that the age of screening cessation based on comorbidity levels varies by nearly a 10-year interval around the age cut-point of 74 included in current breast and colorectal cancer screening recommendations. Our data on common chronic health conditions and their associated comorbidity level, together with model projections of screening benefits and harms at each of these comorbidity levels can inform discussions between providers and their older patients about personalizing decisions about when to stop cancer screening.

REFERENCES

1. U. S. Preventive Services Task Force. Screening for colorectal cancer: U.S. Preventive Services Task Force recommendation statement. *Ann Intern Med* 2008;149:627-37.
2. U. S. Preventive Services Task Force. Screening for breast cancer: U.S. Preventive Services Task Force recommendation statement. *Ann Intern Med* 2009;151:716-26, W-236.
3. U.S. Preventive Services Task Force. Screening for prostate cancer: U.S. Preventive Services Task Force recommendation statement. *Ann Intern Med* 2008;149:185-91.
4. United States Census Bureau. Projections of the United States resident population by age, sex, race, and Hispanic origin. In: US Department of Commerce; 2012.
5. Mandelblatt JS, Wheat ME, Monane M, Moshief RD, Hollenberg JP, Tang J. Breast cancer screening for elderly women with and without comorbid conditions. A decision analysis model. *Ann Intern Med* 1992;116:722-30.
6. Walter LC, Covinsky KE. Cancer screening in elderly patients: a framework for individualized decision making. *JAMA* 2001;285:2750-6.
7. Moyer V. Screening for prostate cancer: U.S. Preventive Services Task Force recommendation statement. *Ann Intern Med* 2012;157:120-34.
8. Wolf AM, Wender RC, Etzioni RB, et al. American Cancer Society guideline for the early detection of prostate cancer: update 2010. *CA Cancer J Clin* 2010;60:70-98.
9. Kawachi MH, Bahnson RR, Barry M, et al. Prostate cancer early detection. Clinical practice guidelines in oncology. *J Natl Compr Canc Netw* 2007;5:714-36.
10. Greene KL, Albertsen PC, Babaian RJ, et al. Prostate specific antigen best practice statement: 2009 update. *J Urol* 2009;182:2232-41.
11. Tan SY, van Oortmarssen GJ, de Koning HJ, Boer R, Habbema JD. The MISCAN-Fadia continuous tumor growth model for breast cancer. *J Natl Cancer Inst Monogr* 2006:56-65.
12. Mandelblatt J, Schechter CB, Lawrence W, Yi B, Cullen J. The SPECTRUM population model of the impact of screening and treatment on U.S. breast cancer trends from 1975 to 2000: principles and practice of the model methods. *J Natl Cancer Inst Monogr* 2006:47-55.
13. Draisma G, Boer R, Otto SJ, et al. Lead times and overdetection due to prostate-specific antigen screening: estimates from the European Randomized Study of Screening for Prostate Cancer. *J Natl Cancer Inst* 2003;95:868-78.
14. Gulati R, Inoue L, Katcher J, Hazelton W, Etzioni R. Calibrating disease progression models using population data: a critical precursor to policy development in cancer control. *Biostatistics* 2010;11:707-19.
15. Frazier AL, Colditz GA, Fuchs CS, Kuntz KM. Cost-effectiveness of screening for colorectal cancer in the general population. *JAMA* 2000;284:1954-61.
16. Rutter CM, Savarino JE. An evidence-based microsimulation model for colorectal cancer: validation and application. *Cancer Epidemiol Biomarkers Prev* 2010;19:1992-2002.
17. Loeve F, Boer R, van Oortmarssen GJ, van Ballegooijen M, Habbema JD. The MISCAN-COLON simulation model for the evaluation of colorectal cancer screening. *Comput Biomed Res* 1999;32:13-33.
18. Mandelblatt JS, Cronin KA, Bailey S, et al. Effects of mammography screening under different screening schedules: model estimates of potential benefits and harms. *Ann Intern Med* 2009;151:738-47.
19. Zauber AG, Lansdorp-Vogelaar I, Knudsen AB, Wilschut J, van Ballegooijen M, Kuntz KM. Evaluating test strategies for colorectal cancer screening: a decision analysis for the U.S. Preventive Services Task Force. *Ann Intern Med* 2008;149:659-69.
20. Gulati R, Wever EM, Tsodikov A, et al. What if i don't treat my PSA-detected prostate cancer? Answers from three natural history models. *Cancer Epidemiol Biomarkers Prev* 2011;20:740-50.
21. Klabunde CN, Legler JM, Warren JL, Baldwin LM, Schrag D. A refined comorbidity measurement algorithm for claims-based studies of breast, prostate, colorectal, and lung cancer patients. *Ann Epidemiol* 2007;17:584-90.
22. Klabunde CN, Potosky AL, Legler JM, Warren JL. Development of a comorbidity index using physician claims data. *J Clin Epidemiol* 2000;53:1258-67.
23. Harewood GC, Lawlor GO, Larson MV. Incident rates of colonic neoplasia in older patients: when should we stop screening? *J Gastroenterol Hepatol* 2006;21:1021-5.
24. Ko CW, Sonnenberg A. Comparing risks and benefits of colorectal cancer screening in elderly patients. *Gastroenterology* 2005;129:1163-70.
25. Maheshwari S, Patel T, Patel P. Screening for colorectal cancer in elderly persons: who should we screen and when can we stop? *J Aging Health* 2008;20:126-39.

26. Mandelblatt JS, Schechter CB, Yabroff KR, et al. Toward optimal screening strategies for older women. Costs, benefits, and harms of breast cancer screening by age, biology, and health status. *J Gen Intern Med* 2005;20:487-96.
27. Rich JS, Black WC. When should we stop screening? *Eff Clin Pract* 2000;3:78-84.
28. Stevens T, Burke CA. Colonoscopy screening in the elderly: when to stop? *Am J Gastroenterol* 2003;98:1881-5.
29. Schonberg MA, Davis RB, McCarthy EP, Marcantonio ER. External validation of an index to predict up to 9-year mortality of community-dwelling adults aged 65 and older. *J Am Geriatr Soc* 2011;59:1444-51.
30. Schonberg MA, Davis RB, McCarthy EP, Marcantonio ER. Index to predict 5-year mortality of community-dwelling adults aged 65 and older using data from the National Health Interview Survey. *J Gen Intern Med* 2009;24:1115-22.
31. Yourman LC, Lee SJ, Schonberg MA, Widera EW, Smith AK. Prognostic indices for older adults: a systematic review. *JAMA* 2012;307:182-92.
32. Dinh TA, Alperin P, Walter LC, Smith R. Impact of comorbidity on colorectal cancer screening cost-effectiveness study in diabetic populations. *J Gen Intern Med* 2012;27:730-8.
33. Smith RA, Cokkinides V, Brawley OW. Cancer screening in the United States, 2012: A review of current American Cancer Society guidelines and current issues in cancer screening. *CA Cancer J Clin* 2012;62:129-42.
34. Gulati R, Gore JL, Etzioni R. Comparative effectiveness of alternative prostate-specific antigen-based prostate cancer screening strategies: Model estimates of potential benefits and harms. *Ann Intern Med* 2013;in press.
35. Heijnsdijk EAM, Wever EM, Auvinen A, Hugosson J, Ciatto S, Nelen V, Kwiatkowski M, Villers A, Páez A, Moss SM, Zappa M, Tammela TLJ, Mäkinen T, Carlsson S, Korfage IJ, Essink-Bot M-L, Otto SJ, Draisma G, Bangma CH, Roobol MJ, Schröder FH, de Koning HJ. Quality-of-life effects of prostate-specific antigen screening. *New Engl J Med* 2012;367:595-605.
36. Mandelblatt J, Tosteston A, van Ravesteyn NT. Costs, evidence and value in the Medicare program: the challenges of technology innovation in breast cancer prevention and control. *Arch Intern Med* 2013 2013;173:227-8.
37. Begg C. Comments and response on the USPSTF recommendation on screening for breast cancer. *Ann Intern Med* 2010;152:540-1; author reply 3-4.
38. Braithwaite R. Comments and response on the USPSTF recommendation on screening for breast cancer. *Ann Intern Med* 2010;152:539-40; author reply 43-4.
39. Col N, Hansen MH, Fischhoff B, Pauker SG. Comments and response on the USPSTF recommendation on screening for breast cancer. *Ann Intern Med* 2010;152:542; author reply 3-4.
40. Harris R, Sawaya GF, Moyer VA, Calonge N. Reconsidering the criteria for evaluating proposed screening programs: reflections from 4 current and former members of the U.S. Preventive Services Task Force. *Epidemiol Rev* 2011;33:20-35.
41. Ho A. Comments and response on the USPSTF recommendation on screening for breast cancer. *Ann Intern Med* 2010;152:542-3; author reply 3-4.
42. Klijs B, Mackenbach JP, Kunst AE. Obesity, smoking, alcohol consumption and years lived with disability: a Sullivan life table approach. *BMC Public Health* 2011;11:378.
43. Chang Y, Schechter CB, van Ravesteyn NT, Near A, Heijnsdijk EA, Adams-Campbell L, Levy D, de Koning HJ, Mandelblatt JS. Collaborative Modeling of the Impact of Obesity on Race–specific Breast Cancer Incidence and Mortality. *Breast Cancer Res Treat* 2012;136:823-35.
44. Huxley RR, Ansary-Moghaddam A, Clifton P, Czernichow S, Parr CL, Woodward M. The impact of dietary and lifestyle risk factors on risk of colorectal cancer: a quantitative overview of the epidemiological evidence. *Int J Cancer* 2009;125:171-80.
45. Nelson NJ. Studies on how lifestyle factors may affect breast cancer risk and recurrence. *J Natl Cancer Inst* 2012;104:574-6.
46. Wolk A. Diet, lifestyle and risk of prostate cancer. *Acta Oncol* 2005;44:277-81.
47. Warren JL, Klabunde CN, Mariotto AB, et al. Adverse events after outpatient colonoscopy in the Medicare population. *Ann Intern Med* 2009;150:849-57, W152.

APPENDIX

Appendix 9.A

Recommended age of screening cessation – results of sensitivity analysis on the choice of metrics to capture the balance between harms and benefits

Table 9.A1. Age of screening cessation (difference from base case estimate) among regularly screened individuals, by cancer site, comorbidity level and model – based on numbers of false-positive tests per life-year gained

Cancer site / model	Age of screening cessation by comorbidity level[1]			
	No comorbidity	Mild comorbidity[2]	Moderate comorbidity[3]	Severe comorbidity[4]
Breast cancer				
MISCAN-Fadia	80 (+2)	72 (-2)	70 (0)	64 (0)
SPECTRUM	76 (0)	74 (0)	72 (0)	64 (-4)
Prostate cancer				
MISCAN-Prostate	76 (-2)	76 (0)	74 (+6)	68 (+4)
FHCRC prostate model	78 (0)	76 (-2)	74 (0)	74 (+6)
Colorectal cancer				
MISCAN-Colon	77 (0)	75 (0)	73 (0)	69 (0)
CRC-SPIN	76 (0)	74 (0)	71 (0)	66 (0)
SimCRC	76 (0)	74 (0)	73 (0)	69 (0)

[1] Age of screening cessation is the oldest age at which the numbers of false-positive tests per life-year gained is still less than that for a final screening age of 76 (75 for colorectal cancer) vs. age 74 in the entire population

[2] Mild comorbidity includes having a life expectancy associated with having a history of MI, acute MI, ulcer, or rheumatologic disease

[3] Moderate comorbidity includes having a life-expectancy associated with having vascular disease, cardiovascular disease, paralysis or diabetes, or combinations of diabetes with MI, ulcer, or rheumatologic disease

[4] Severe comorbidity includes having a life expectancy associated with having AIDS, COPD, mild liver disease, severe liver disease, chronic renal failure, dementia, or congestive heart failure, or combinations of aforementioned diseases not categorized under moderate comorbidity

Table 9.A2. Recommended age of screening cessation (difference from base case estimate) among regularly screened individuals, by cancer site, comorbidity level and model – based on numbers of overdiagnosed cases per life-year gained

Cancer site / model	Age of screening cessation by comorbidity level[1]			
	No comorbidity	Mild comorbidity[2]	Moderate comorbidity[3]	Severe comorbidity[4]
Breast cancer				
MISCAN-Fadia	78 (0)	74 (0)	72 (+2)	68 (+4)
SPECTRUM	78 (+2)	74 (0)	74 (+2)	66 (-2)
Prostate cancer				
MISCAN-Prostate	76 (-2)	76 (0)	74 (+6)	68 (+4)
FHCRC prostate model	78 (-4)	78 (0)	74 (0)	70 (+2)
Colorectal cancer				
MISCAN-Colon	77 (0)	75 (0)	72 (-1)	67 (-2)
CRC-SPIN	77 (+1)	75 (+1)	72 (+1)	66 (0)
SimCRC	77 (+1)	75 (+1)	73 (0)	67 (-2)

[1] Age of final screening is the oldest age at which the overdiagnosed cases per life-year gained are still less than those for a final screening age of 76 (75 for colorectal cancer) vs. age 74 in the entire population

[2] Mild comorbidity includes having a life expectancy associated with having a history of MI, acute MI, ulcer, or rheumatologic disease

[3] Moderate comorbidity includes having a life-expectancy associated with having vascular disease, cardiovascular disease, paralysis or diabetes, or combinations of diabetes with MI, ulcer, or rheumatologic disease

[4] Severe comorbidity includes having a life expectancy associated with having AIDS, COPD, mild liver disease, severe liver disease, chronic renal failure, dementia, or congestive heart failure, or combinations of aforementioned diseases not categorized under moderate comorbidity

Table 9.A3. Recommended age of screening cessation among regularly screened individuals (difference from base case estimate), by cancer site, comorbidity level and model – based on number needed to screen per cancer death prevented

Cancer site / model	Age of screening cessation by comorbidity level[1]			
	No comorbidity	Mild comorbidity[2]	Moderate comorbidity[3]	Severe comorbidity[4]
Breast cancer				
MISCAN-Fadia	82 (+4)	80 (+6)	70 (0)	64 (0)
SPECTRUM	76 (0)	72 (-2)	64 (-8)	64 (-4)
Prostate cancer				
MISCAN-Prostate	78 (0)	64 (-12)	64 (-4)	64 (0)
FHCRC prostate model	82 (+4)	80 (+2)	74 (0)	64 (-4)
Colorectal cancer				
MISCAN-Colon	77 (0)	75 (0)	72 (-1)	67 (-2)
CRC-SPIN	77 (+1)	76 (+2)	71 (0)	65 (-1)
SimCRC	76 (0)	74 (0)	72 (-1)	65 (-4)

[1] Age of final screening is the oldest age at which the number needed to screen to prevent one cancer death is still less than that for a final screening age of 76 (75 for colorectal cancer) vs. age 74 in the entire population

[2] Mild comorbidity includes having a life expectancy associated with having a history of MI, acute MI, ulcer, or rheumatologic disease

[3] Moderate comorbidity includes having a life-expectancy associated with having vascular disease, cardiovascular disease, paralysis or diabetes, or combinations of diabetes with MI, ulcer, or rheumatologic disease

[4] Severe comorbidity includes having a life expectancy associated with having AIDS, COPD, mild liver disease, severe liver disease, chronic renal failure, dementia, or congestive heart failure, or combinations of aforementioned diseases not categorized under moderate comorbidity

Table 9.A4. Recommended age of screening cessation among regularly screened individuals (difference from base case estimate), by cancer site, comorbidity level and model – based on numbers of false-positive tests per cancer death prevented

Cancer site / model	Age of screening cessation by comorbidity level[1]			
	No comorbidity	Mild comorbidity[2]	Moderate comorbidity[3]	Severe comorbidity[4]
Breast cancer				
MISCAN-Fadia	84 (+6)	80 (+6)	70 (0)	64 (0)
SPECTRUM	76 (0)	72 (-2)	64 (-8)	64 (-4)
Prostate cancer				
MISCAN-Prostate	76 (-2)	76 (0)	72 (+4)	64 (0)
FHCRC prostate model	78 (0)	78 (0)	74 (0)	72 (+4)
Colorectal cancer				
MISCAN-Colon	77 (0)	75 (0)	72 (-1)	67 (-2)
CRC-SPIN	77 (+1)	76 (+2)	71 (0)	65 (-1)
SimCRC	77 (+1)	74 (0)	71 (-2)	65 (-4)

[1] Age of final screening is the oldest age at which the numbers of false-positive tests per cancer death prevented are still less than those for a final screening age of 76 (75 for colorectal cancer) vs. age 74 in the entire population

[2] Mild comorbidity includes having a life expectancy associated with having a history of MI, acute MI, ulcer, or rheumatologic disease

[3] Moderate comorbidity includes having a life-expectancy associated with having vascular disease, cardiovascular disease, paralysis or diabetes, or combinations of diabetes with MI, ulcer, or rheumatologic disease

[4] Severe comorbidity includes having a life expectancy associated with having AIDS, COPD, mild liver disease, severe liver disease, chronic renal failure, dementia, or congestive heart failure, or combinations of aforementioned diseases not categorized under moderate comorbidity

Table 9.A5. Recommended age of screening cessation among regularly screened individuals (difference from base case estimate), by cancer site, comorbidity level and model – based on numbers of overdiagnosed cases per cancer death prevented

Cancer site / model	Age of screening cessation by comorbidity level[1]			
	No comorbidity	Mild comorbidity[2]	Moderate comorbidity[3]	Severe comorbidity[4]
Breast cancer				
MISCAN-Fadia	78 (0)	74 (0)	72 (+2)	66 (+2)
SPECTRUM	78 (+2)	74 (0)	74 (+2)	64 (-4)
Prostate cancer				
MISCAN-Prostate	76 (-2)	76 (0)	72 (+4)	64 (0)
FHCRC prostate model	80 (+2)	78 (0)	74 (0)	70 (+2)
Colorectal cancer				
MISCAN-Colon	77 (0)	75 (0)	71 (-2)	65 (-4)
CRC-SPIN	77 (+1)	76 (+2)	73 (+2)	65 (-1)
SimCRC	77 (+1)	75 (+1)	73 (0)	65 (-4)

[1] Age of final screening is the oldest age at which the numbers of overdiagnosed cases per cancer death prevented is still less than those for a final screening age of 76 (75 for colorectal cancer) vs. age 74 in the entire population

[2] Mild comorbidity includes having a life expectancy associated with having a history of MI, acute MI, ulcer, or rheumatologic disease

[3] Moderate comorbidity includes having a life-expectancy associated with having vascular disease, cardiovascular disease, paralysis or diabetes, or combinations of diabetes with MI, ulcer, or rheumatologic disease

[4] Severe comorbidity includes having a life expectancy associated with having AIDS, COPD, mild liver disease, severe liver disease, chronic renal failure, dementia, or congestive heart failure, or combinations of aforementioned diseases not categorized under moderate comorbidity

Appendix 9.B

Recommended age of screening cessation– results of sensitivity analysis on the method for extrapolating the comorbidity-specific life tables

Table 9.B1. Recommended age of screening cessation among regularly screened individuals, by cancer site, comorbidity level and model – based on extrapolation of lifetables assuming that the hazard ratio between the average life table and the comorbidity-specific life table at the 13 years of follow-up was maintained throughout the rest of life – based on number of screen tests per life-year gained

Cancer site / model	Age of screening cessation by comorbidity level[1]			
	No comorbidity	Mild comorbidity[2]	Moderate comorbidity[3]	Severe comorbidity[4]
Breast cancer				
MISCAN-Fadia	78 (0)	72 (-2)	70 (0)	64 (0)
SPECTRUM	76 (0)	74 (0)	70 (-2)	64 (-4)
Prostate cancer				
MISCAN-Prostate	78 (0)	76 (0)	64 (-4)	64 (0)
FHCRC prostate model	76 (-2)	74 (-4)	64 (-10)	64 (-4)
Colorectal cancer				
MISCAN-Colon	77 (0)	74 (-1)	70 (-3)	65 (-4)
CRC-SPIN	78 (+2)	76 (+2)	71 (0)	65 (-1)
SimCRC	76 (0)	74 (0)	69 (-4)	65 (-4)

[1] Age of final screening is the oldest age at which the number needed to screen to gain one life-year (NNS/LYG) is still less than the NNS/LYG for a final screening age of 76 (75 for colorectal cancer) vs. age 74 in the entire population

[2] Mild comorbidity includes having a life expectancy associated with having a history of MI, acute MI, ulcer, or rheumatologic disease

[3] Moderate comorbidity includes having a life-expectancy associated with having vascular disease, cardiovascular disease, paralysis or diabetes, or combinations of diabetes with MI, ulcer, or rheumatologic disease

[4] Severe comorbidity includes having a life expectancy associated with having AIDS, COPD, mild liver disease, severe liver disease, chronic renal failure, dementia, or congestive heart failure, or combinations of aforementioned diseases not categorized under moderate comorbidity

Chapter 10

Prediction of higher mortality reduction for the UK Breast Screening Frequency Trial: a model-based approach on screening intervals

Nicolien T. van Ravesteyn, Eveline A.M. Heijnsdijk, Gerrit Draisma, Harry J. de Koning

Br J Cancer. 2011 Sep 27;105(7):1082-8

ABSTRACT

Background: The optimal interval between two consecutive mammograms is uncertain. The UK Frequency Trial did not show a significant difference in breast cancer mortality between screening every year (study group) and screening every 3 years (control group). In this study, the trial is simulated in order to gain insight into the results of the trial and to predict the effect of different screening intervals on breast cancer mortality.

Methods: UK incidence, life tables and information from the trial were used in the micro-simulation model MISCAN-Fadia to simulate the trial and predict the number of breast cancer deaths in each group. To be able to replicate the trial, a relatively low sensitivity had to be assumed.

Results: The model simulated a larger difference in tumour size distribution between the two groups than observed and a relative risk (RR) of 0.83 of dying from breast cancer in the study group compared with the control group. The predicted RR is lower than that reported from the trial (RR 0.93), but within its 95% confidence interval (0.63 - 1.37).

Conclusion: The present study suggests that there is benefit of shortening the screening interval, although the benefit is probably not large enough to start annual screening.

INTRODUCTION

In randomised controlled trials, mammography screening has been shown to reduce breast cancer mortality.[1-3] The more frequently a woman has screening exams, the larger the probability of having an early diagnosis, and the larger the mortality reduction might be. However, with more frequent exams, the potential of false positive exams and overdiagnosis will also increase.[4-5] There is no consensus on the optimal screening interval (i.e., the time between two consecutive mammograms), as is illustrated by the variety of screening intervals used throughout the world. Most European screening programs use an interval of 2 years (e.g., The Netherlands, Sweden), whereas other countries use a 3-year interval (United Kingdom, Malta). Even within the same country screening recommendations vary: in the United States, the American Cancer Society recommends annual screening starting at the age of 40 years,[6] while the US Preventive Services Task Force recently changed their recommendation to biennial screening from age 50 to 74 years.[7]

Two randomised trials compared a 1-year screening interval with a 3-year screening interval, one in women between age 40 and 49 years [8] and one in women between age 50 and 62 years.[9] The latter, the UK Breast Screening Frequency Trial, was conducted from 1989 to 1996, in order to evaluate the difference in (predicted) breast cancer mortality between screening annually and screening once every 3 years. [9] The tumours in the trial group, offered annual screening, were significantly smaller than those diagnosed in the control group, offered screening every 3 years. For node status and histological grade, no significant difference between the two groups was found. The initially reported relative risk (RR) predicted on the basis of two prognostic indices showed a nonsignificant reduction in predicted breast cancer mortality. [9] The results were later updated with results on the actual observed number of breast cancer deaths in both groups again showing a nonsignificant reduction in breast cancer mortality. Women in the study group had an RR of 0.93 (95% confidence interval (CI) 0.63-1.37) of dying from breast cancer compared to women in the control group. [10]

This finding was (slightly) surprising and raised the question why no significant difference was found between the two groups. It might be that there is truly only a very small mortality benefit of more frequent screening, or there might be other reasons why no difference is found between the two groups, for example, a lack of power or low sensitivity of mammography. Most policy predictions are based on the assumption that increasing the screening frequency will lead to more early diagnoses and consequently in a reduction in breast cancer mortality, hence it is crucial to get more insight in the results of this trial. A simulation model is ideally suited to evaluate the effect of different screening intervals on mortality, because the effect of different screening test sensitivities can be assessed and the model guarantees that trial populations are identical, except for the factors investigated.

In the present study, the UK Breast Screening Frequency Trial was simulated using the microsimulation model MIcrosimulation of SCreening ANalysis - Fatal diameter (MISCAN-

Fadia), in order to gain insight into the results of the trial and estimate the effect of different screening intervals on breast cancer mortality.

MATERIALS AND METHODS

Model overview

MISCAN-Fadia is a microsimulation model developed within the Cancer Intervention and Surveillance Modeling Network (CISNET).[11] Briefly, the model simulates life histories in the absence of screening and then assesses how these life histories change as a consequence of screening programmes. MISCAN-Fadia explicitly models invasive tumour growth in combination with the concept of a fatal diameter. The model has been described in detail elsewhere [11] and information about the model can be found on the CISNET website (http://cisnet.cancer.gov/). A detailed description of the model components and model quantification for the present study is presented in the Appendix.

In brief, for the present study, the model simulates a population of women between the ages of 50 and 62 years in the year 1992 (i.e., the middle year of the trial) using the life tables of the UK female population. Among those who develop breast cancer, the natural history is modelled as a continuously growing tumour. Each tumour has a size (the fatal diameter, which differs between tumours) at which diagnosis and treatment will no longer result in cure given available treatment options. If the tumour is diagnosed (either on the basis of clinical presentation with symptoms or by screening) and treated before the tumour has reached its fatal diameter, the woman will be cured and will die of non-breast cancer causes. Variation between tumours is modelled by probability distributions of tumour growth, threshold diameter of screen detection, clinical diagnosis diameter, and fatal disease diameter.

When a screening program is applied, the pre-clinical tumour may be detected by screening. Each simulated tumour has a diameter at which it will be clinically diagnosed as well as a screen-detection threshold diameter. For the latter, screening test sensitivity is 0% below and 100% above this diameter. The threshold diameter is assumed to decrease with age and calendar year. Screening benefits result from detection of more tumours at a non-fatal size.[11]

Model calibration and validation

Several approaches have been used to assess the internal reliability of MISCAN-Fadia and the validity of the results against external data, as previously reported.[11] For the present study, age-specific breast cancer incidence rates for the years 1975-1988, that is, before the implementation of the National Health Service Breast Screening Programme (NHSBSP) were used to estimate age-specific parameters for disease onset. The age-specific breast cancer incidence rates for the year 1988 as simulated by MISCAN-Fadia were compared to the observed incidence rates for the year 1988 in the United Kingdom.

UK breast screening frequency trial

Five screening units participated in the trial between 1989 and 1996 [9] (see Figure 10.1 for an overview of the trial design). A total of 99 389 women aged 50-62 years who had been invited to a prevalence screen in the NHSBSP were randomised to a conventional screen after an interval of 3 years (control group, *n* = 50 216), or to three annual screenings (study group, *n* = 49 173). For the primary analysis, only women who attended the prevalence screen and in whom no cancer was found at the prevalence screen were included (*n* = 38 492 in the control group and *n* = 37 530 in the study group). The attendance rate in the control group, among women who had attended the prevalence screen, was 85%. In the study group, attendance rates at the three yearly screens were 78%, 78%, and 81%, respectively.[9]

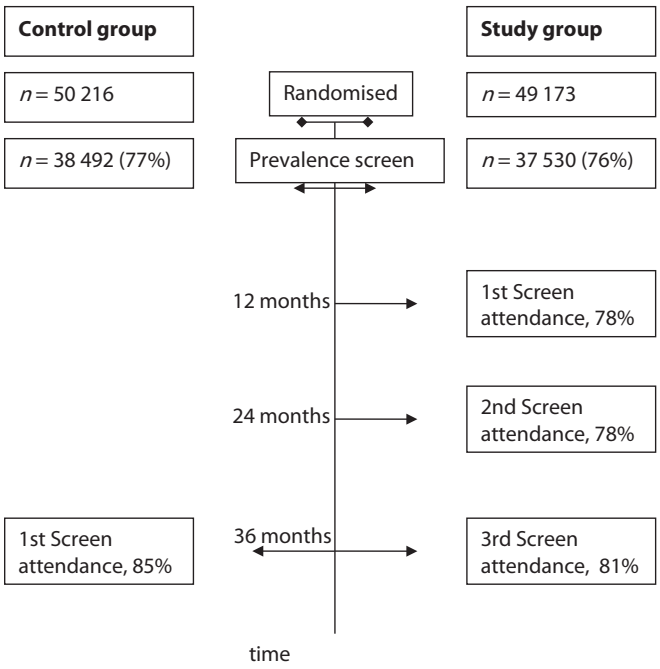


Figure 10.1. A schematic overview of the UK Breast Screening Frequency Trial

Trial replication and mortality prediction

Initially, the model based on data from randomised screening trials (extrapolated to the current period) and US data simulated a more favourable tumour size distribution than observed in the trial for both groups. Therefore, the threshold diameter and diameter of clinical detection were estimated using data from the Frequency Trial on the numbers of invasive breast cancers in both groups of the trial split out by tumour size and detection mode (see Appendix). Compared with the initially used values, the estimated values were somewhat higher for the diameter of clinical detection and the threshold diameter, corresponding to a lower screening sensitivity.

Subsequently, this fitted model was used to predict the number of breast cancer deaths from cancers diagnosed during the trial period in each group with a follow-up period up to 2006. From these numbers, a predicted RR of dying from breast cancer in the study group compared with the control group was calculated.

In addition, we investigated the effect of a longer follow-up period (i.e., until all women have died), a higher sensitivity (using the initial value for the threshold diameter), and full compliance (i.e., 100% attendance rates) on the predicted RR.

RESULTS

Model calibration and validation

The observed age-specific incidence rates in the year 1988 as reported by the NHSBSP were accurately reproduced by MISCAN-Fadia (Figure 10.2). For each 5-year age group (35-79 year), the difference between the observed and simulated incidence rates was <10%.

Trial replication and mortality prediction

The model with the threshold diameter and diameter of clinical detection estimated based on the trial data, simulated a total of 523 (445 invasive) breast cancers in women who attended the prevalence screen compared with a total of 535 (443 invasive) cancers observed in the trial. The numbers of detected breast cancers and percentages screen detected, and clinically detected cancers are close to the observed numbers and percentages in both the groups (Table 10.1).

For the trial period, the cumulative incidence (number of invasive breast cancers detected) in both groups over time since prevalence screen, as observed in the trial and simulated by the model is shown in Figure 10.3.

The model simulated a more favourable tumour size distribution in the study group than in the control group, in line with what was observed in the trial (Table 10.2). For the control group, the simulated size distribution was somewhat less favourable (61% small tumours simulated *vs* 66% observed) and for the study group, the simulated size distribution was

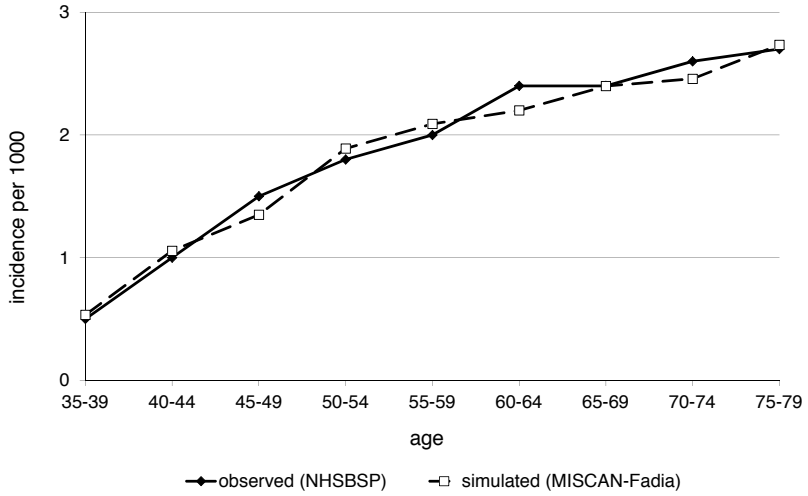


Figure 10.2. Age-specific breast cancer incidence rates in the UK for the year 1988 as observed and simulated by MISCAN-Fadia

Table 10.1. Cumulative number of breast cancers by detection mode in the control group (3-year screening interval) and screen group (1-year screening interval) as observed in the Frequency Trial and simulated by MISCAN-Fadia (prevalence screen attenders only)

	Observed				**Simulated**			
	Control		Study		Control		Study	
	Number	(%)	Number	(%)	Number	(%)	Number	(%)
Screen detected	135	(54)	206	(72)	129	(51)	195	(72)
Clinically detected	113	(46)	81	(28)	123	(49)	77	(28)
Total	248		287		251		272	

somewhat too favourable (77% small tumours simulated *vs* 73% observed). Thus, the model simulated a larger difference in size distribution between the control and study group than observed.

In the control group 55 breast cancer deaths from cancers diagnosed in the trial were observed during the median follow-up of 162 months, [10] compared with 54 deaths predicted. In the study group 50 breast cancer deaths were observed, [10] whereas the model predicted 45 breast cancer deaths. The predicted difference between the number of deaths in the control group and the study group was larger than the observed difference, corresponding to a predicted RR of 0.83 of dying from breast cancer in the study group compared with the control group.

A longer follow-up period (life-time) had no effect on the predicted RR (Table 10.3). Increasing the sensitivity led to a higher percentage of screen-detected cancers in both groups (78% in the study and 58% in the control group) and to a lower predicted RR of 0.81, as did

increasing the attendance to 100%. The combination of full compliance and higher sensitivity led to a predicted RR of 0.77.

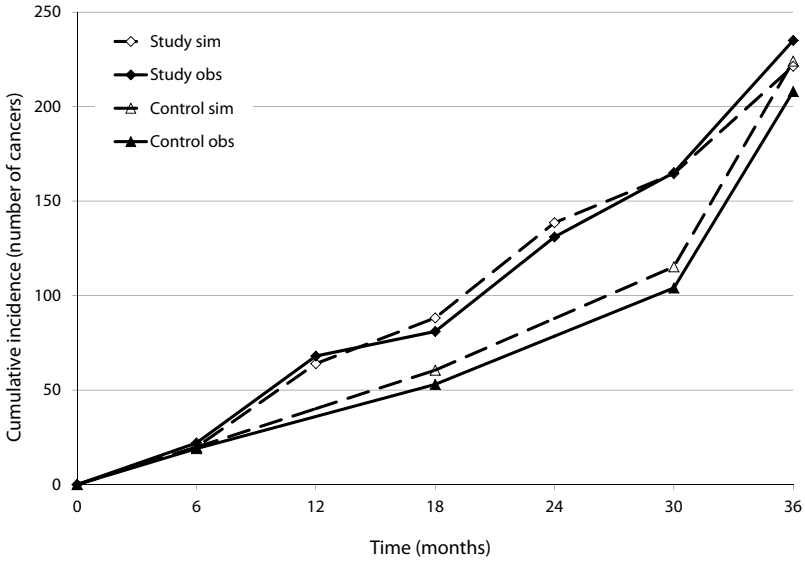


Figure 10.3. Cumulative incidence (number of invasive breast cancers) over time since prevalence screen in the control and study group as observed in the Frequency Trial (obs) and simulated by MISCAN-Fadia (sim)

Table 10.2. Number (%) of detected cancers (invasive by tumour size) in the control and study group as observed in the Frequency Trial and simulated by MISCAN-Fadia

	Observed				**Simulated**			
	Control		Study		Control		Study	
	Number	(%)	Number	(%)	Number	(%)	Number	(%)
DCIS	40	(16)	52	(18)	27	(11)	50	(19)
Invasive cancers	208	(84)	235	(82)	224	(89)	227	(81)
By tumour size								
1–20 mm	134	(66)	170	(73)	136	(61)	170	(77)
21–50 mm	64	(32)	59	(25)	78	(35)	48	(22)
50+ mm	5	(2)	4	(2)	10	(5)	3	(2)
Not known	5		2		0		0	
Total	248		287		251		272	

Abbreviation: DCIS = ductal carcinoma *in situ*

Table 10.3. Predicted relative risks using different assumptions

	Predicted RR
Base run	0.83
Life-time follow-up	0.83
Higher sensitivity	0.81
100% attendance	0.81
Higher sensitivity and 100% attendance	0.77

Abbreviation: RR = relative risk

DISCUSSION

The present study suggests that there is benefit in terms of a reduction in breast cancer mortality associated with shortening the screening interval from 3 years to 1 year. The results show that if the available information from the UK Breast Screening Frequency Trial is used in a microsimulation model that is based on the results of randomised screening trials including a large(r) number of women, a larger effect of shortening the screening interval is predicted. The predicted RR of breast cancer death for the study group (offered three annual screens) compared with the control group (offered one screen after 3 years) was 0.83. The effect of more frequent screening is predicted to be larger when the attendance rate or screening test sensitivity is increased.

The microsimulation model used in the present study fitted the data better when a somewhat higher threshold diameter, corresponding to a lower screening test sensitivity, and diameter of clinical detection is used compared with a model based on data from randomised screening trials (extrapolated to the current period) and from the US. Thus, it seems that the screening test sensitivity in the trial was relatively low, which is in line with previously reported results showing that screen-detected as well as interval cancers could benefit from improved sensitivity.[12] The results of the present study indicate that when the screening test sensitivity is higher, the effect of shortening the screening interval will be somewhat larger. This finding is important when considering shorter screening intervals for certain risk groups. For example, it has been hypothesized that women with high breast density might benefit more from more frequent screening, because they have a higher risk of breast cancer.[13] However, screening test sensitivity has been found to be lower in women with high breast density.[14] The present study shows that the lower sensitivity in this group might offset some of the potential benefit of more frequent screening in this group.

The most important limitation of the current study is the relative paucity of data available to simulate the trial. More information (e.g., on the age distribution of participants and tumour size distribution of screen *vs* clinically detected cancers in both groups) might further improve the model, and consequently the model predictions. In addition, detailed information on the attendance rates was not available. For example, the non-attendance in the study group was somewhat higher than in the control group (approximately 20% *vs* 15%).

It is unknown which proportion of the non-attenders in the study group missed multiple rounds.[15] Including more detailed information in the model will lead to better estimates of the effect of shortening the screening interval.

In addition, only one simulation model was used to estimate the effect of 1-year *vs* 3-year screening intervals. Having multiple models that come to similar findings might have strengthened the conclusion of the present study.

Despite these limitations, the microsimulation model, used in this study, adequately simulated the number of screen detected and interval cancers in both arms of the trial. Moreover, the predicted numbers of breast cancer deaths from cancers diagnosed in the trial were of the same magnitude as that of the reported numbers in both groups [10], and the predicted RR is within the 95% CI of the estimate reported from the trial.

The UK Breast Screening Frequency Trial showed a non-significant 7% reduction in breast cancer deaths in the study *vs* the control group, [10] whereas the present study finds a substantially larger effect of shortening the screening interval (17%). The question arises why the model outcomes differ from the trial results. Several factors might contribute.

Firstly, the RR predicted by the model is within the 95% CI of the trial-reported RR, indicating that the predicted RR is not statistically different from the trial-reported RR. The Frequency Trial invited 99 389 women, based on an expected 25% difference in breast cancer mortality between the study and control group.[16] The current study shows that this estimated difference of 25% was too optimistic, suggesting that the trial was underpowered to find a significant difference between the two groups. On the basis of the results of the current study (i.e., an RR of 0.83), approximately 945 000 women needed to have been invited for a power of 80% to demonstrate a significant (P=0.05) difference in breast cancer mortality between the two groups.[17] However, the trial was designed to show a difference in predicted mortality, based on surrogate end points; in this case the tumour size of the detected cancers. It was estimated that the sample size can be 2.74 times smaller without losing precision when surrogate end points are used.[16] This means that when surrogate end points (such as prognostic indices) are used, at least 345 000 women needed to have been invited in order to have 80% power. Our findings indicate that the number needed to invite can also be reduced by increasing compliance to screening tests or increasing screening test sensitivity.

Furthermore, in the trial more invasive breast cancers were detected in the study group than in the control group. Thus, more diagnoses have been moved forward in time in the study group than in the control group and then, more breast cancer deaths from cancers diagnosed in trial can be expected in the study group. An alternative would be to compare the number of breast cancer deaths from all breast cancers (during a certain follow up period). However, after the trial, everyone receives usual care (triennial screening), resulting in a dilution of the effect on mortality. Both comparisons will lead to an underestimation of the effect of more frequent screening (i.e., a bias towards an RR of 1).

The results of the available observational studies are somewhat contradictory on the effect of shortening the screening interval. For example, two retrospective studies showed similar prognostic factors for women screened annually *vs* biennially.[18,19] However, two other studies found that women who were screened annually had breast tumours that were smaller and less advanced than those who were screened every other year.[20,21] Furthermore, six independent models showed that there is some benefit when moving from biennial to annual screening, although the benefit diminished (i.e., the benefit of moving from biennial to annual screening is smaller than that of moving from no screening to biennial screening). For example, 68% - 90% of the benefit is maintained when moving from annual to biennial screening scenarios for women aged 50-69 years.[22] Thus, the benefit of screening every year is not three times as large as screening once every 3 years. The associated harms and costs also have to be taken into account when determining the optimal screening frequency and increases more steeply with more frequent screening than the benefits.

In conclusion, the present study suggests that there is benefit in terms of a mortality reduction of shortening the screening interval from 3 years to 1 year. However, the benefit is probably not large enough to start annual screening.[23] At the same time, there seems to be no reason to abolish the 2-year interval currently used in most European screening programmes. For these programmes benefits in terms of mortality reductions have been shown.[24,25]

Acknowledgements

This study was supported, in part, by grants 1U01CA152958 and 2U01CA088283 from the National Cancer Institute and by the National Institute for Public Health and the Environment (RIVM).

REFERENCES

1. Nystrom L, Andersson I, Bjurstam N, Frisell J, Nordenskjold B, Rutqvist LE. Long-term effects of mammography screening: updated overview of the Swedish randomised trials. *Lancet* 2002;359:909-19.
2. Tabar L, Fagerberg G, Duffy SW, Day NE, Gad A, Grontoft O. Update of the Swedish two-county program of mammographic screening for breast cancer. *Radiol Clin North Am* 1992;30:187-210.
3. Nystrom L, Rutqvist LE, Wall S, Lindgren A, Lindqvist M, Ryden S, et al. Breast cancer screening with mammography: overview of Swedish randomised trials. *Lancet* 1993;341:973-8.
4. Christiansen CL, Wang F, Barton MB, Kreuter W, Elmore JG, Gelfand AE, et al. Predicting the cumulative risk of false-positive mammograms. *J Natl Cancer Inst* 2000;92:1657-66.
5. Jansen JT, Zoetelief J. Optimisation of mammographic breast cancer screening using a computer simulation model. *Eur J Radiol* 1997;24:137-44.
6. Smith RA, Cokkinides V, Brooks D, Saslow D, Brawley OW. Cancer screening in the United States, 2010: a review of current American Cancer Society guidelines and issues in cancer screening. *CA Cancer J Clin* 2010;60:99-119.
7. U.S. Preventive Services Task Force. Screening for breast cancer: U.S. Preventive Services Task Force recommendation statement. *Ann Intern Med* 2009;151:716-26, W-236.
8. Klemi PJ, Toikkanen S, Rasanen O, Parvinen I, Joensuu H. Mammography screening interval and the frequency of interval cancers in a population-based screening. *Br J Cancer* 1997;75:762-6.
9. Breast Screening Frequency Trial Group. The frequency of breast cancer screening: results from the UKCCCR Randomised Trial. United Kingdom Co-ordinating Committee on Cancer Research. *Eur J Cancer* 2002;38:1458-64.
10. Duffy SW, Blamey R. Long-term mortality results from the UK Screening Frequency Trial. *6th European Breast Cancer Conference*. Berlin, Germany, 2008.
11. Tan SY, van Oortmarssen GJ, de Koning HJ, Boer R, Habbema JD. The MISCAN-Fadia continuous tumor growth model for breast cancer. *J Natl Cancer Inst Monogr* 2006:56-65.
12. Warren RM, Young JR, McLean L, Lyons K, Wilson AR, Evans A, et al. Radiology review of the UKCCCR Breast Screening Frequency Trial: potential improvements in sensitivity and lead time of radiological signs. *Clin Radiol* 2003;58:128-32.
13. Kerlikowske K, Cook AJ, Buist DS, Cummings SR, Vachon C, Vacek P, et al. Breast cancer risk by breast density, menopause, and postmenopausal hormone therapy use. *J Clin Oncol* 2010;28:3830-7.
14. Carney PA, Miglioretti DL, Yankaskas BC, Kerlikowske K, Rosenberg R, Rutter CM, et al. Individual and combined effects of age, breast density, and hormone replacement therapy use on the accuracy of screening mammography. *Ann Intern Med* 2003;138:168-75.
15. Andersson I. Comment on "The frequency of breast cancer screening: results from the UKCCCR Randomised Trial". *Eur J Cancer* 2002;38:1427-8; discussion 65.
16. Day NE, Duffy SW. Trial Design Based on Surrogate End Points--Application to Comparison of Different Breast Screening Frequencies. *J R Stat Soc A Stat* 1996;159:49-60.
17. Lwanga SK, Lemeshow S. Sample size determination in health studies: a practical manual. Geneva: World Health Organization, 1991.
18. Wai ES, D'Yachkova Y, Olivotto IA, Tyldesley S, Phillips N, Warren LJ, et al. Comparison of 1- and 2-year screening intervals for women undergoing screening mammography. *Br J Cancer* 2005;92:961-6.
19. White E, Miglioretti DL, Yankaskas BC, Geller BM, Rosenberg RD, Kerlikowske K, et al. Biennial versus annual mammography and the risk of late-stage breast cancer. *J Natl Cancer Inst* 2004;96:1832-9.
20. Field LR, Wilson TE, Strawderman M, Gabriel H, Helvie MA. Mammographic screening in women more than 64 years old: a comparison of 1- and 2-year intervals. *AJR Am J Roentgenol* 1998;170:961-5.
21. Hunt KA, Rosen EL, Sickles EA. Outcome analysis for women undergoing annual versus biennial screening mammography: a review of 24,211 examinations. *AJR Am J Roentgenol* 1999;173:285-9.
22. Mandelblatt JS, Cronin KA, Bailey S, Berry DA, de Koning HJ, Draisma G, et al. Effects of mammography screening under different screening schedules: model estimates of potential benefits and harms. *Ann Intern Med* 2009;151:738-47.
23. Boer R, de Koning H, Threlfall A, Warmerdam P, Street A, Friedman E, et al. Cost effectiveness of shortening screening interval or extending age range of NHS breast screening programme: computer simulation study. *BMJ* 1998;317:376-9.

24. Otto SJ, Fracheboud J, Looman CW, Broeders MJ, Boer R, Hendriks JH, et al. Initiation of population-based mammography screening in Dutch municipalities and effect on breast-cancer mortality: a systematic review. *Lancet* 2003;361:1411-7.
25. Tabar L, Yen MF, Vitak B, Chen HH, Smith RA, Duffy SW. Mammography service screening and mortality in breast cancer patients: 20-year follow-up before and after introduction of screening. *Lancet* 2003;361:1405-10.

APPENDIX

This appendix consists of three parts:

(1) a model overview containing a description of MISCAN-Fadia,
(2) a description of the model components, and
(3) a description of the model quantifications of each model component

Model overview

MISCAN-Fadia (MIcrosimulation of SCreening ANalysis – Fatal diameter) is a microsimulation model generating independent life histories. It uses the so-called parallel universe approach and first simulates the individual life histories for women in the absence of screening and then assesses how these histories would change as a consequence of a screening program. A certain percentage of the modelled population develops pre-clinical disease. The natural history of breast cancer is modelled as a continuously growing tumour. MISCAN-Fadia includes a sub model for ductal carcinoma *in situ* (DCIS), with three different types of preclinical DCIS: regressive DCIS, DCIS that will be diagnosed clinically, and DCIS that will progress to invasive disease. When a screening program is applied, the pre-clinical tumour may be detected by screening if it is larger than a screen-detection threshold diameter.

Model components

Demographics

The demography part of the model simulates individual life histories without breast cancer to form a population. For each person, a date of birth and a date of death of other causes than breast cancer are simulated. The distribution of births and deaths can be adjusted to represent the population simulated.

Incidence

A certain percentage of the modelled population develops preclinical disease. This percentage varies between birth cohorts, while the cohorts have the same age distribution of onset of breast cancer.

Natural history

Among women who develop disease, the natural history of breast cancer is modelled as a continuously growing tumour. Each tumour has a size (the fatal diameter, which differs between tumours) at which diagnosis and treatment will no longer result in cure given available treatment options. If the tumour is diagnosed (either on the basis of clinical presentation with symptoms or by screening) and treated before the tumour reaches the fatal diameter, the woman will be cured and will die of non-breast cancer causes (Appendix Figure 10.1). Variation between tumours is modelled by probability distributions of tumour growth rate, threshold diameter of screen detection, clinical diagnosis diameter, fatal disease diameter, and survival duration since fatal diameter.

Screening

When a screening program is applied, the preclinical tumour may be detected by screening. Each simulated tumour has a diameter at which it will be clinically diagnosed and a screen-detection threshold diameter. For the latter, screening test sensitivity is 0% below and 100% above this diameter. The threshold diameter is dependent on the calendar year and age of the woman (decreasing with calendar year and older age). Screening benefits result from detection of more tumours at a non-fatal size. The characteristics of organized screening programs, such as screening ages, screening interval and attendance can be specified, and the type of screening (e.g. 'organized' or 'opportunistic') can be defined in the model.

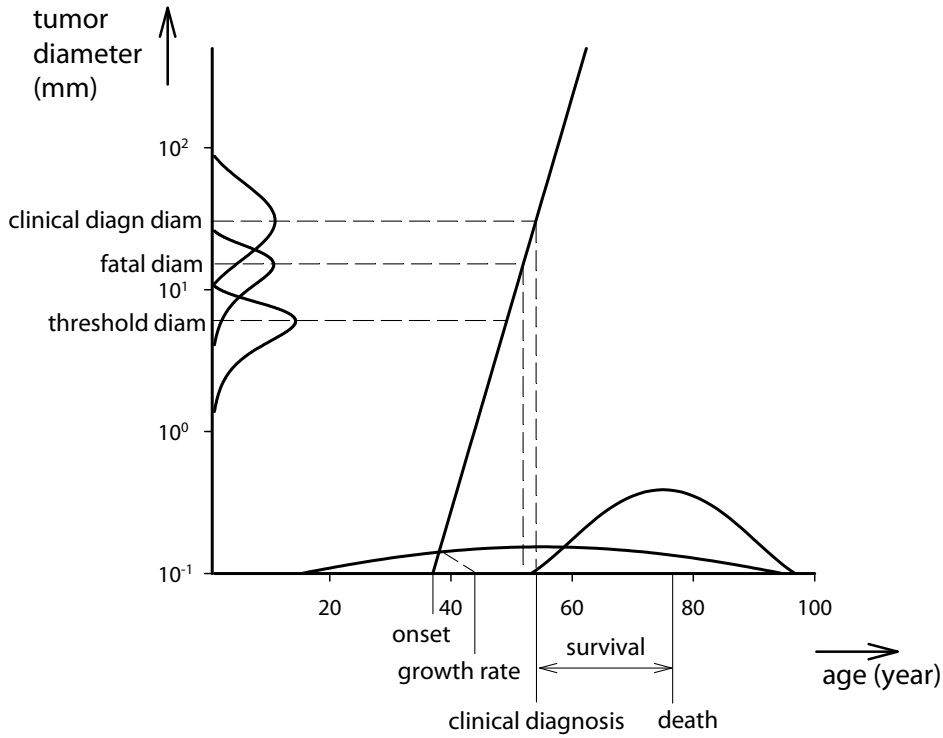
Treatment

The benefit of adjuvant treatment is modelled as a shift in the fatal diameter for treated women. For each adjuvant treatment a cure proportion is estimated (depending on age). These cure proportions are then translated into corresponding fatal diameters (i.e., a more effective treatment can cure a larger tumour). The dissemination of adjuvant treatment is modelled as the probability of being treated with a certain type of treatment (e.g. chemotherapy, tamoxifen).

Model quantification

Demographics

A female population born between 1930 and 1942 (thus, age 50-62 in 1992) was simulated. We assumed that all cohorts were represented in equal proportions. Each woman is assigned a date of death due to non-breast cancer causes based on the UK female population cohort life tables. These life tables were available from the Human Mortality Database.[1] The simulated woman dies because of breast cancer or of other causes, whichever comes first.



Appendix Figure 10.1. The MISCAN-Fadia natural history model.
The model is illustrated by a woman who is diagnosed with incurable breast cancer and for whom screening could have been beneficial. The natural history of breast cancer is simulated through the random selection of six variables from probability distributions, denoted by the various curves: onset = age at tumour onset, growth rate = tumour growth rate, survival = duration between the moment at which the tumour reaches the fatal diameter and the moment of death from breast cancer (not shown), clinical diagn diam = tumour diameter at which the tumour will be diagnosed clinically because of the primary tumour, fatal diam = tumour diameter at which available treatment options will no longer result in cure, threshold diam = tumour diameter at which the tumour becomes screen detectable. After onset the tumour starts growing exponentially according to the tumour growth rate. The diagnosis results from the clinical diagnosis diameter combined with the tumour growth rate. If the tumour is diagnosed after it has reached the fatal diameter, the woman will die from breast cancer. Survival is modelled since fatal diameter. For observed survival (shown), the time between clinical diagnosis and the moment the tumour has reached its fatal diameter has to be subtracted. Screening can change this natural history: After the tumour has reached the threshold diameter, the tumour can be screen detected. If the tumour has not reached the fatal diameter yet at the moment of screen detection, the woman will be cured. Otherwise, screening will not affect the woman's age of death. Reprinted from Tan *et al*., 2006 with permission from Oxford University Press.

Incidence

We used age-specific breast cancer incidence rates for the years 1975-1988, i.e. before the implementation of the National Health Service Breast Screening Programme to estimate age-specific onset parameters for the onset of breast cancer.

Natural history

All parameters were previously estimated using detailed data from the Two County Study .[2,3] Subsequently, the fatal diameter was calibrated to U.S. data concerning 1975 stage distribution and 1975 survival (SEER data).[3]

For the present study, we re-estimated two parameters: the diameter at clinical diagnosis and the screening threshold diameters. We re-estimated these parameters, because these parameters can vary across countries and over time, whereas other parameters (e.g. the tumour growth rate) are assumed to be more or less universal. To estimate these two parameters we used the following data from the UK Frequency Trial:

- The number of detected invasive cancers over time in both groups (control & study group)
- The total number of invasive cancers by tumour size and detection mode in both groups

The estimated values of the diameter at clinical diagnosis and the screening threshold diameters are shown in Appendix Table 10.1.

The estimated values for these parameters are somewhat higher than the values previously found based on data from the Two County Study (Lognormal (0.8, 0.6) and Weibull (1.0, 3.0) for the diameter at clinical diagnosis and the screening threshold diameter, respectively).[3] If these lower values are used to simulate the Frequency Trial, a more favourable tumour size is simulated than the one that is observed in the trial.

Screening

In the present study we simulated the screening in the UK Breast Screening Frequency Trial using previously published attendance rates. The attendance rate in the control group, among women who had attended the prevalence screen, was 85%. In the study group, attendance rates at the three yearly screens were 78, 78, and 81%, respectively.[4]

Treatment

We used treatment effectiveness data based on meta-analyses of the Early Breast Cancer Trialists' Collaborative Group.[5,6] The dissemination of adjuvant treatment (the probability of being treated with a certain type of treatment) was based on data from the U.S.[7]

Appendix Table 10.1. Maximum likelihood estimates for MISCAN-Fadia natural history parameters and the data sources used to estimate the parameter distribution

a. Parameter estimates and their distribution

Variable		Distribution	mean	sd	Data used [ref]
Growth rate (1/year)		Lognormal (0.062, 0.87)	1.6	1.7	TCS [2,3]
Fatal diameter (cm)		Weibull (4.0, 0.95)	4.1	4.3	SEER [3]
Survival duration since fatal diameter (years)		Lognormal (2.4, 1.1)	22	35	TCS [2,3]
Diameter at clinical diagnosis because of primary tumour (cm)		Lognormal (0.88, 0.6)	2.8	1.8	Freq [4]
Screening threshold diameters (cm)	Age 50-59	Weibull (1.33, 2.95)	1.2	0.4	Freq [4]
	Age 60-65	(1.05, 2.95)	0.9	0.3	Freq [4]

b. Correlation between variables

Variable	r	
growth rate – survival	-0.9	TCS [2,3]
growth rate – clinical diagnosis diameter because of the primary tumour	0.41	TCS [2,3]
clinical diagnosis diameter because of the primary tumour – survival	-0.43	TCS [2,3]

c. Time since start of fatal disease at which metastases lead to clinical diagnosis of the tumour (fraction of the total survival time after reaching the fatal diameter): 0.9 — TCS [2,3]

Abbreviations
TCS: Two County Study
sd: standard deviation
SEER: Surveillance, Epidemiology and End Results
Freq: UK Breast Screening Frequency Trial

Appendix references

1. Human Mortality Database, University of California, Berkeley (USA), and Max Planck Institute for Demographic Research (Germany).
2. Tabar L, Fagerberg G, Duffy SW, et al. Update of the Swedish two-county program of mammographic screening for breast cancer. *Radiol Clin North Am* 1992;30:187-210.
3. Tan SY, van Oortmarssen GJ, de Koning HJ, et al. The MISCAN-Fadia continuous tumor growth model for breast cancer. *J Natl Cancer Inst Monogr* 2006;56-65.
4. Breast Screening Frequency Trial Group. The frequency of breast cancer screening: results from the UKCCCR Randomised Trial. United Kingdom Co-ordinating Committee on Cancer Research. *Eur J Cancer* 2002;38:1458-64.
5. Early Breast Cancer Trialists' Collaborative Group. Polychemotherapy for early breast cancer: an overview of the randomised trials. *Lancet* 1998;352:930-42.
6. Early Breast Cancer Trialists' Collaborative Group. Tamoxifen for early breast cancer: an overview of the randomised trials. *Lancet* 1998;351:1451-67.
7. Mariotto AB, Feuer EJ, Harlan LC, et al. Dissemination of adjuvant multiagent chemotherapy and tamoxifen for breast cancer in the United States using estrogen receptor information: 1975-1999. *J Natl Cancer Inst Monogr* 2006;7-15.

Chapter 11

Consequences of the transition from film to digital mammography for breast cancer screening through the U.S. National Breast and Cervical Cancer Early Detection Program

Nicolien T. van Ravesteyn, Lisanne van Lier, Clyde B. Schechter, Donatus U. Ekwueme, Janet Royalty, Jacqueline W. Miller, Aimee M. Near, Kathleen A. Cronin, Eveline A.M. Heijnsdijk, Jeanne S. Mandelblatt*, Harry J. de Koning*

** dual senior authors*

Submitted

ABSTRACT

Background: The National Breast and Cervical Cancer Early Detection Program (NBCCEDP) funds mammograms and diagnostic services for qualifying low-income, uninsured women aged 40-64 in the U.S.

Purpose: To assess the implications of replacing film with digital mammography for the NBCCEDP.

Methods: Data from the NBCCEDP and national data representative of the program's target population were used in two established simulation models. Cost-effectiveness was examined for 10 screening strategies varying by screening interval (annual, biennial, once), starting age (40, 50) and modality (film, digital) for White, Black, and Hispanic women. We compared the impact of replacing film with digital mammography in the NBCCEDP on health effects, costs and number of women reached.

Results: Cost-effectiveness ratios varied by race/ethnicity and were higher for annual than for biennial screening. Digital mammography gained 2-4% (range across models) more life-years than film mammography, but had higher costs for screening and diagnostics (34-35%). With a fixed budget, 25-26% fewer women could be served and 22-24% fewer life-years would be gained if all plain-film services were converted to digital services. The loss in life-years could be reversed to a 8-13% increase by only including biennial screening.

Conclusions: Digital mammography could save slightly more lives than film mammography. However, with a fixed budget, fewer women can be served and fewer life-years can be gained. Price reduction of digital mammography, budget increases, or changes in the program, such as prolonging the screening interval, are options for achieving comparable health effects when replacing film with digital mammography within the NBCCEDP.

INTRODUCTION

In the United States (US) breast cancer mortality rates decreased steadily from 1990 to 2009.[1] The decrease in mortality has been found to be in part attributable to mammography screening.[2] However, not all women have benefited equally from screening mammography, with low-income, uninsured women having substantially lower screening rates than their more advantaged, insured counterparts.[3]

To reduce disparities in screening rates, the Centers for Disease Control and Prevention (CDC) established the National Breast and Cervical Cancer Early Detection Program (NBCCEDP). Started in 1991, the program offers free or low-cost mammograms to low-income, uninsured and underinsured women aged 40 to 64 years. Specific eligibility criteria vary by state but are generally limited to women with incomes ≤250% of the federal poverty level, which applies to approximately 10% of all US women age 40-64.[4] In 2011, NBCCEDP screened 332,788 women for breast cancer with mammography [4] reaching a variety of race/ethnicity groups (47% White, 24% Hispanic, 18% Black, 5% Asian/Pacific Islanders, and 6% other, multiracial, unknown).[5]

Following publication of the Digital Mammographic Imaging Screening Trial (DMIST) in 2005,[6] program facilities gradually started shifting from plain-film to digital mammography and by 2010, 47% of screening examinations within the NBCCEDP were performed using digital mammography. Digital mammography has been found to have higher test sensitivity than plain-film for women under the age of 50 years and women with dense breasts.[6] However, overall the diagnostic accuracy of digital and film mammography is similar, but digital screening is more expensive than plain-film.[6,7] This poses a potential dilemma for the NBCCEDP which operates within a fixed budget determined by the US Congress. If the program were to cover the higher costs of digital screening, it would not be able to reach as many women as it could by paying for the less expensive plain-film modality. However, if digital were to perform better in the NBCCEDP target population, then more lives might be saved. Thus, it is unclear what the consequences of the transition from film to digital mammography would mean for the number of averted breast-cancer deaths and life-years gained (LYG) through the program.

We used two established microsimulation models to estimate the impact of the transition from film to digital mammography among the target population of the NBCCEDP in terms of program budget, numbers of women served, numbers of averted breast cancer deaths, and LYG. In addition, we evaluated ways to increase the efficiency of the program. To do so, we evaluated the cost-effectiveness (cost per LYG) of mammography by race/ethnicity (White, Black and Hispanic women) for different screening strategies, reflecting the screening behaviors observed in the NBCCEDP. Then, we assessed what the implications would be if only the most cost-effective digital screening scenarios (with regard to screening interval and starting age) were included in the NBCCEDP.

METHODS

Models

We used two models (MISCAN-Fadia and SPECTRUM) developed independently within the Cancer Intervention and Surveillance Modeling Network (CISNET) to assess the implications of shifting from film to digital mammography for the NBCCEDP. The models have previously been described in detail [8,9] and information about the models can also be found online (cisnet.cancer.gov/breast/profiles.html). Briefly, the models simulated life histories for individual women. After estimating breast-cancer incidence and mortality in the absence of screening and adjuvant therapy, the models overlaid screening use and improvements in survival associated with treatment advances.[2,10]

MISCAN-Fadia models tumor growth, where tumors can be detected once they are beyond a detection threshold and cured if the tumor is below a fatal diameter.[8] In SPECTRUM, tumors progress through stages, with screening effects due to age and stage shifts and adjuvant treatment reducing the hazard of death.[9] In both models ductal carcinoma in situ (DCIS) is represented as a state that can regress, remain and be diagnosed or progress to invasive cancer.

Model parameters

We used a combination of primary data from the NBCCEDP and national data representative of the program's target population to develop common input parameters. In 2008 the majority of women served by the NBCCEDP consisted of low-income White, Black, and Hispanic women.[11] We modeled these three groups separately since the three groups have different risk of disease and disease characteristics.[12] The following set of race/ethnicity-specific data inputs was used to model breast cancer incidence and mortality by race/ethnicity.

Demographics

We simulated a multi-cohort population matching the demographics of the three racial/ethnic groups of women born between 1945 and 1970 (i.e., age 40-64 in 2010). Each woman was assigned a date of death due to causes other than breast cancer based on race-specific data from the National Center for Health Statistics.[13] We used the non-breast cancer mortality of White women for Hispanic women, because their life expectancy has been found to be very similar.[14] The simulated women die because of breast cancer or of other causes, whichever comes first.

Incidence

The breast cancer incidence in the absence of screening was based on an age-period-cohort model for the U.S. population.[15] These data were applied for incidence rates for White women. Data for age-specific relative risks for Black vs. White and Hispanic vs. White women from the

Surveillance, Epidemiology and End Results (SEER) database were used to create age-period-cohort data for each race/ethnic group.

Natural history

SEER data for stage distribution and breast cancer-specific survival for White and Black women from 1975 to 1979 were used to model the natural history of breast cancer in the absence of mammography screening and adjuvant therapy, since these cancer control interventions did not begin to disseminate into the population in a substantial manner until after 1980.[10] SEER data for Hispanic women specifically was only available from 1990 onwards. Therefore, we used the stage distribution for clinical and interval cancers from the Breast Cancer Surveillance Consortium (BCSC) to estimate the stage distribution in the absence of screening for Hispanic women (unpublished data). Breast cancer specific survival of Hispanic women was assumed to be equal to that of Whites, when corrected for multiple factors, based on prior reports.[16]

Screening test characteristics

We used data from BCSC on recent screening performance in terms of sensitivity and specificity of film and digital mammography by age and screening interval from the years 1996 to 2007 and is assumed to be equal for the three race/ethnicity groups (unpublished data).

Screening scenarios

We evaluated the costs and effects of 10 screening scenarios, reflecting the screening behaviors observed in the NBCCEDP: 1) annual screening starting at age 40; 2) annual screening starting at age 50; 3) biennial screening starting at age 40; 4) biennial screening starting at age 50; and 5) irregular screening; performed using film or digital mammography.

Treatment

Treatment effectiveness is age- and ER/HER2-specific and based on synthesis of recent clinical trials.[17-19] Treatment effects are modeled as a proportionate reduction in mortality risk or the proportion cured. Treatment use and treatment effects were modified for Blacks vs. Whites based on prior research.[10] Treatment impact and use for Hispanic women was assumed to be equal to that for Whites.[16] Specific treatment data were available until the year 2005; we assumed the same rates thereafter.

Costs

We used most recent cost estimates for breast-cancer screening, diagnosis and treatment from the literature (Table 11.1). For the cost of screening we used the weighted average NBCCEDP reimbursement rates in the period 2009-2010 for film and digital mammography. For the cost of diagnostics, we used mean work up costs within 12 months of initial screening of

Table 11.1. Overview of cost estimates

		Cost in 2010 $	95% CI	Source
Screening				
Film		84		D. Ekwueme & J. Royalty; personal communication
Digital		115		
Diagnostics				
	Total cost with film			
	TP	1,842	(1,714-1,970)	[7]
	FP	394	(376-411)	[7]
	Total cost with digital			
	TP	1,896	(1,767-2,026)	[7]
	FP	443	(425-460)	[7]
Treatment				
Initial	Stage at diagnosis			
	In situ	9,186	(5,054-13,318)	[20]
	Local	12,156	(11,861-12,451)	[21]
	Regional	22,983	(22,315-23,649)	[21]
	Distant	35,495	(33,013-37,976)	[21]
Terminal	Cancer death	94,284		[22]
	Other cause	748		[22]

TP: True-Positive mammography interpretation
FP: False-Positive mammography interpretation

women with a true-positive and false-positive mammography interpretation for digital and plain-film evaluations.[7] Total treatment costs were calculated for three commonly reported phases of care, also known as the initial, continuing and the last year of life phases. The initial phase include all care provided in the first 12 months following diagnosis. The terminal phase includes all care provided within 12 months prior to death, and the continuing phase includes all care provided in the months between these two phases. All costs were updated to 2010 US dollars, using the medical care component of the Consumer Price Index available at http://www.bls.gov/cpi/data.htm.

Model validation

Both models have previously been validated for Black and White women.[10] For Hispanic women, we compared model projected age-adjusted incidence and mortality with actual SEER rates available for the years 1992-2007 and 1990-2007, respectively.

Cost-effectiveness analysis

We evaluated the cost-effectiveness of 10 screening scenarios varying by starting age (40 vs. 50), screening interval (annual, biennial, once), and screening modality (film vs. digital) for White, Black, and Hispanic women. For each screening strategy the total costs were calculated and the effects consisted of the total number of life-years gained in each scenario. We calculated the cost-effectiveness ratio (CER) as the difference in total costs between a

situation with screening (film or digital) and a situation without screening, divided by the difference in life-years gained between a situation with screening and no screening. In addition, the incremental cost-effectiveness was assessed by ranking the strategies (within each race/ethnicity group and modality) based on the total costs and comparing each screening scenario to its less intensive counterpart. We also assessed the incremental cost-effectiveness of digital vs. film mammography for each screening schedule.

Implications for the NBCCEDP

We estimated the costs and effects of the NBCCEDP by using the percentages of screens currently performed annually (with an interval of <18 months), biennially (18-30 months) and irregularly or once (> 30 months) in each age and race group for film mammography. These screening rates are based on data from CDC showing that in 2010 around 41% of women had a screen in the preceding 18 months and an additional 10% a screen in the preceding 30 months. The remaining 49% did not have a screen in the NBCCEDP in the preceding 30 months (Table 11.2). Subsequently, we evaluated the costs and health effects (LYG and breast cancer deaths averted) of shifting to digital mammography assuming the same screening distributions. In addition, we estimated the health effects of the program using digital mammography assuming a fixed budget for the cost of screening and diagnostics. Finally, we assessed what the implications would be if only the most cost-effective digital screening scenarios were included in the NBCCEDP.

Table 11.2. Distribution of mammogram screening interval in the NBCCEDP by age and race

Screening interval	Age	White	Black	Hispanic	Total
'annual' (<18 months)	40-49	6.2%	1.5%	3.1%	10.9%
	50-64	18.7%	6.3%	5.6%	30.6%
'biennial' (18-30 months)	40-49	2.0%	0.6%	1.3%	3.9%
	50-64	2.9%	1.4%	1.4%	5.7%
Irregular/once	40-64	26.9%	9.3%	12.7%	49.0%
total		56.8%	19.0%	24.1%	100%

Sensitivity analysis

We performed multiple sensitivity analyses to assess the uncertainty around the costs and performance of digital mammography. We used the upper and lower limit for the screening test performance (sensitivity, specificity) and varied the cost of digital screening and diagnostics.

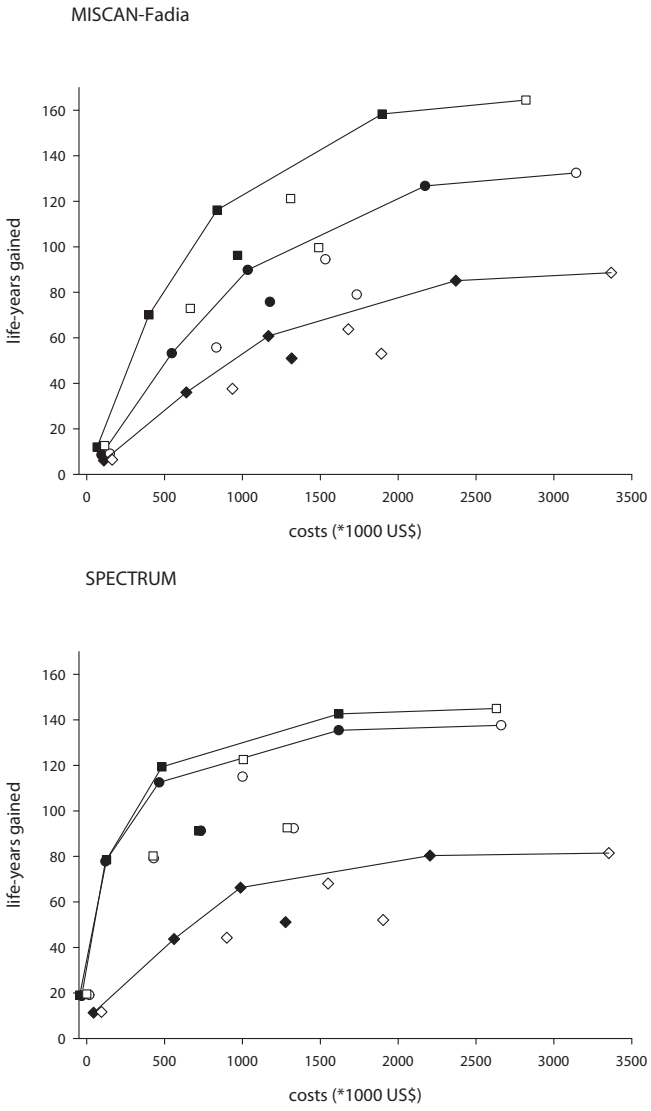


Figure 11.1. Efficiency frontiers for 10 screening scenarios for three race/ethnicity groups (White, Black and Hispanic women)

Figure legend
Squares: Black women
Circles: White women
Diamonds: Hispanic women
On frontier/line=not dominated (i.e., efficient strategies)
Closed symbols represent film mammography and open symbols represent digital mammography

RESULTS

Model validation

The predicted values of both models for age-adjusted incidence and mortality were close to the observed rates (Appendix Figure 11.1&11.2).

Cost-effectiveness

The cost-effectiveness of screening varied by race/ethnicity; screening was most cost-effective for Black and least cost-effective for Hispanic women. The ranking of the strategies within each race/ethnicity group was very similar and for each group the same screening scenarios were on (or close to) the efficiency frontier in both models (Figure 11.1).

For film compared to no screening, biennial screening scenarios were more cost-effective than annual screening scenarios, with CERs ranging from $2,000-$19,000/LYG for biennial and $8,000-$28,000/LYG for annual screening depending on race/ethnicity and model. Similarly, for digital strategies compared to no screening, biennial screening scenarios were more cost-effective for all race/ethnicity groups (with CERs ranging from $5,000-$26,000/LYG) than annual screening scenarios (CERs ranging from $14,000-$41,000/LYG) (Appendix Table 11.1).

Screening with film mammography was more cost-effective than screening with digital mammography. The incremental cost-effectiveness ratios of digital vs. film were higher than $78,000/LYG in both models for all race/ethnicity groups.

Implications for the NBCCEDP

The models estimated that within the NBCCEDP using film mammography, the 329,721 mammograms that were performed in 2010 would avert 107-136 (range across models) breast cancer deaths and gain 1,948-2,305 life-years. If the same number of mammograms were performed digitally, there would be 111-138 breast cancer deaths averted and 2,034-2,345 LYG; a 2-4% increase (Table 11.3). However, to perform the same number of screens with digital as with plain film, the budget would need to increase substantially (34-35%). If the budget is assumed to be fixed, then fewer women can be served (25-26%) with digital than with film mammography. This would result in fewer breast cancer deaths averted (23-24%) and LYG (22-24%) (Table 11.3).

As biennial screening was found to be more cost-effective than annual screening, we evaluated the effect of only including biennial screening, which was defined as an interval of 18 to 30 months. Although fewer women are served, the benefit per screen is higher for biennial than for annual screening. Restricting the screening interval to biennial, thus, leads to a higher number of LYG per screen. If a fixed budget is assumed, then the change in LYG is inversed from a loss of 22-24% to an increase of 8-13% (Table 11.3). Moreover, if screening is restricted to women aged 50 and older, the program would become even more efficient and the number of LYG would increase by 16-17%.

Table 11.3. Implications of switching from film to digital mammography in the NBCCEDP

modality	screening	assumptions	number of mmg	*(% change)^*	LYG	*(% change)^*	bc deaths averted	*(% change)^*	Costs screen & diagnostics (million $)	*(% change)^*
MISCAN-Fadia										
film	as observed within the NBCCEDP*	fixed no. of mmg	329,721		1948		107		$45.2	
digital	as observed within the NBCCEDP**	fixed no. of mmg	329,721		2034	*(4.4%)*	111	*(4.2%)*	$60.9	*(34.7%)*
digital	as observed within the NBCCEDP*	fixed budget	244,723	*(-25.8%)*	1510	*(-22.5%)*	83	*(-22.6%)*	$45.2	
digital	biennial#	fixed budget	277,257	*(-15.9%)*	2196	*(12.8%)*	116	*(8.9%)*	$45.2	
digital	biennial# age 50+	fixed budget	282,474	*(-14.3%)*	2257	*(15.9%)*	135	*(26.1%)*	$45.2	
SPECTRUM										
film	as observed within the NBCCEDP*	fixed no. of mmg	329,721		2305		136		$48.8	
digital	as observed within the NBCCEDP**	fixed no. of mmg	329,721		2345	*(1.7%)*	138	*(1.7%)*	$65.4	*(34.0%)*
digital	as observed within the NBCCEDP*	fixed budget	246,020	*(-25.4%)*	1749	*(-24.1%)*	103	*(-24.1%)*	$48.8	
digital	biennial#	fixed budget	275,113	*(-16.6%)*	2481	*(7.7%)*	141	*(3.6%)*	$48.8	
digital	biennial# age 50+	fixed budget	273,420	*(-17.1%)*	2693	*(16.8%)*	169	*(24.0%)*	$48.8	

* see Table 2

^ compared to film mammography

screening interval 18-30 months

Abbreviations

bc: breast cancer

lyg: life-years gained

mmg: mammograms

Sensitivity analysis

Varying the screening test performance of digital did not change the results substantially. When the same cost of diagnostics is assumed for digital as for film, then the budget needs to be increased by 29-30% instead of 34-35%, and by 26% (both models) assuming a 10% lower cost for digital screening.

DISCUSSION

Our results indicate that with the fixed budget of the NBCCEDP there are fewer breast cancer deaths averted and fewer LYG when film mammography is replaced by digital mammography, because a lower number of screening tests can be funded. Although digital mammography performs better than film mammography, it is more costly. In order to keep the same number of mammograms, the budget of the NBCCEDP needs to be increased by 34-35%. We also found that changes to the program, in particular restricting the screening interval to biennial will lead to a more efficient program and can offset the decrease in benefits of shifting to digital mammography.

There are several options to achieve comparable health effects (LYG and breast cancer deaths averted) of film and digital mammography and maintain the reach of the NBCCEDP. The first option is to increase the federal budget available for breast cancer by 34-35%. A second option would be a price reduction of digital mammography. These two options are, however, not easily achievable for CDC. A third way to achieve an equal number of LYG for the two modalities is to make some program changes by only funding screening with an biennial interval, then the number of LYG can be increased. Another option is to only include women age 50 years and older. However, the effect of this change is quite small as only 15% of federally-funded mammograms are performed on women younger than 50. In addition to the options we investigated, there might be other program changes that would potentially be even more cost-effective, for example, screening with even longer intervals or focusing more on the highest risk groups. However, since the NBCCEDP is concerned with equity and delivering comparable services as are offered to insured women in the US, we focused on scenarios that were in line with current breast cancer screening recommendations.[23]

The results from both models indicated that biennial screening was more cost-effective than annual screening and both models found that film mammography was more cost-effective than digital mammography, which is line with previous work.[7] Screening Hispanic women was less cost-effective than screening White women due to their lower incidence; screening Black women was slightly more cost-effective due to a more advanced stage of disease in the absence of screening, resulting in a larger window of opportunity for screening. However, for all race/ethnicity groups the CERs compared to no screening are below the commonly used willingness-to-pay threshold of $50,000 per LYG.

The results across the two models were consistent on the ranking of the strategies, although the shape of the efficiency curves differed somewhat between the models, indicating that in one model (SPECTRUM) the incremental benefits of adding additional mammograms is smaller than in the other model. This reflects the uncertainty around the incremental effect of more frequent screening and is in line with previous findings.[24,25]

An important strength of our study is that we specifically included an ethnically diverse population by modeling White, Black, and Hispanic women separately. As these racial/ethnic groups have been found to have different risks of getting breast cancer and dying from it,[26,27] it is essential to specifically incorporate these differences. In addition, we used two models to provide a range of plausible effects and illustrate the effects of differences in model structure. We also used recent data on screening performance from the BCSC as previous work found that results from BCSC and NBCCEDP were broadly similar.[28] An advantage is that data from BCSC reflects screening performance in community practice, which might be different than the performance observed in screening trials.

Despite these strengths and our consistent results, our study has some limitations. First, we assumed that women in the NBCCEDP received the same treatment (once diagnosed with breast cancer) as women in the overall US population. Although low-income, uninsured women might not always receive optimal treatment [29] women diagnosed through the NBCCEDP are enrolled in Medicaid; thus, in theory, access to care is not an issue. In addition, we were not able to include other minority groups, such as Asian, Pacific Islanders, and multiracial women, because of the limited available data for these groups. Together these latter groups comprise 11% of NBCCEDP's population. However, adding these other minority groups would probably not have influenced our findings with regard to the implications of switching from film to digital mammography screening, although Asian women have more dense breasts [30] and for them, digital mammography might be more cost-effective than for the other race/ethnicity groups. Finally, it has been found that over time the performance of digital may improve and costs may decrease. For example, a recent study found the increase in downstream breast related cost associated with digital mammography diminished over time.[31] In our sensitivity analysis, when assuming the same cost of diagnostics for digital as for film, we found that the budget needs to be increased by 29-30% instead of 34-35%. Thus, even when the diagnostics costs diminish over time, the budget still needs to be increased substantially.

In conclusion, although digital mammography saved slightly more lives than film mammography, fewer women can be served by the NBCCEDP with shifting from film to digital mammography. However, changes in the program services, such as increasing the screening interval, can help ensure that the same health benefits are achieved for program-eligible women.

Acknowledgements

We thank the Breast Cancer Surveillance Consortium (BCSC) investigators for the data they have provided for this study. BCSC data collection was supported by the NCI-funded Breast Cancer Surveillance Consortium (U01CA63740, U01CA86076, U01CA86082, U01CA63736, U01CA70013, U01CA69976, U01CA63731, U01CA70040, and HHSN261201100031C). A list of the BCSC investigators and procedures for requesting BCSC data for research purposes are provided at: http://breastscreening.cancer.gov/.

Disclaimer

The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the Centers for Disease Control and Prevention. Model results are the sole responsibility of the authors.

REFERENCES

1. Jemal A, Simard EP, Dorell C, Noone AM, Markowitz LE, Kohler B, et al. Annual Report to the Nation on the Status of Cancer, 1975-2009, Featuring the Burden and Trends in Human Papillomavirus (HPV)-Associated Cancers and HPV Vaccination Coverage Levels. *J Natl Cancer Inst* 2013;105:175-201.
2. Berry DA, Cronin KA, Plevritis SK, Fryback DG, Clarke L, Zelen M, et al. Effect of screening and adjuvant therapy on mortality from breast cancer. *N Engl J Med* 2005;353:1784-92.
3. Sabatino SA, Coates RJ, Uhler RJ, Breen N, Tangka F, Shaw KM. Disparities in mammography use among US women aged 40-64 years, by race, ethnicity, income, and health insurance status, 1993 and 2005. *Med Care* 2008;46:692-700.
4. Centers for Disease Control and Prevention. National Breast and Cervical Cancer Early Detection Program (NBCCEDP). Available from: http://www.cdc.gov/cancer/nbccedp/about.htm. Accessed: Oct. 15, 2012.
5. Centers for Disease Control and Prevention. National Breast and Cervical Cancer Early Detection Program (NBCCEDP). Available from: http://www.cdc.gov/cancer/nbccedp/data/summaries/national_aggregate.htm. Accessed: Feb. 20, 2013.
6. Pisano ED, Gatsonis C, Hendrick E, Yaffe M, Baum JK, Acharyya S, et al. Diagnostic performance of digital versus film mammography for breast-cancer screening. *N Engl J Med* 2005;353:1773-83.
7. Tosteson AN, Stout NK, Fryback DG, Acharyya S, Herman BA, Hannah LG, et al. Cost-effectiveness of digital mammography breast cancer screening. *Ann Intern Med* 2008;148:1-10.
8. Tan SY, van Oortmarssen GJ, de Koning HJ, Boer R, Habbema JD. The MISCAN-Fadia continuous tumor growth model for breast cancer. *J Natl Cancer Inst Monogr* 2006:56-65.
9. Mandelblatt J, Schechter CB, Lawrence W, Yi B, Cullen J. The SPECTRUM population model of the impact of screening and treatment on U.S. breast cancer trends from 1975 to 2000: principles and practice of the model methods. *J Natl Cancer Inst Monogr* 2006:47-55.
10. van Ravesteyn NT, Schechter CB, Near AM, Heijnsdijk EA, Stoto MA, Draisma G, et al. Race-specific impact of natural history, mammography screening, and adjuvant treatment on breast cancer mortality rates in the United States. *Cancer Epidemiol Biomarkers Prev* 2011;20:112-22.
11. Eheman CR, Benard VB, Blackman D, Lawson HW, Anderson C, Helsel W, et al. Breast cancer screening among low-income or uninsured women: results from the National Breast and Cervical Cancer Early Detection Program, July 1995 to March 2002 (United States). *Cancer Causes Control* 2006;17:29-38.
12. Desantis C, Siegel R, Bandi P, Jemal A. Breast cancer statistics, 2011. *CA Cancer J Clin* 2011;61:409-18.
13. Carter S, Gartner S, Haines M, Olmstead A, Stuch R, Wright G. Historical statistics of the United States. New York: Cambridge Universiy Press, 2006.
14. Smith DP, Bradshaw BS. Rethinking the Hispanic paradox: death rates and life expectancy for US non-Hispanic White and Hispanic populations. *Am J Public Health* 2006;96:1686-92.
15. Holford TR, Cronin KA, Mariotto AB, Feuer EJ. Changing patterns in breast cancer incidence trends. *J Natl Cancer Inst Monogr* 2006:19-25.
16. Li CI, Malone KE, Daling JR. Differences in breast cancer stage, treatment, and survival by race and ethnicity. *Arch Intern Med* 2003;163:49-56.
17. Early Breast Cancer Trialists' Collaborative Group. Polychemotherapy for early breast cancer: an overview of the randomised trials. *Lancet* 1998;352:930-42.
18. Early Breast Cancer Trialists' Collaborative Group. Tamoxifen for early breast cancer: an overview of the randomised trials. *Lancet* 1998;351:1451-67.
19. Early Breast Cancer Trialists' Collaborative Group. Effects of chemotherapy and hormonal therapy for early breast cancer on recurrence and 15-year survival: an overview of the randomised trials. *Lancet* 2005;365:1687-717.
20. Subramanian S, Trogdon J, Ekwueme DU, Gardner JG, Whitmire JT, Rao C. Cost of Breast Cancer Treatment in Medicaid: Implications for State Programs Providing Coverage for Low-Income Women. *Med Care* 2011;49:89-95.
21. Yabroff KR, Lamont EB, Mariotto A, Warren JL, Topor M, Meekins A, et al. Cost of care for elderly cancer patients in the United States. *J Natl Cancer Inst* 2008;100:630-41.
22. Mariotto AB, Robin Yabroff K, Shao Y, Feuer EJ, Brown ML. Projections of the Cost of Cancer Care in the United States: 2010-2020. *J Natl Cancer Inst* 2011;103:117-28.
23. U. S. Preventive Services Task Force. Screening for breast cancer: U.S. Preventive Services Task Force recommendation statement. *Ann Intern Med* 2009;151:716-26, W-236.

24. Mandelblatt JS, Cronin KA, Bailey S, Berry DA, de Koning HJ, Draisma G, et al. Effects of mammography screening under different screening schedules: model estimates of potential benefits and harms. *Ann Intern Med* 2009;151:738-47.
25. van Ravesteyn NT, Miglioretti DL, Stout NK, Lee SJ, Schechter CB, Buist DS, et al. Tipping the balance of benefits and harms to favor screening mammography starting at age 40 years: a comparative modeling study of risk. *Ann Intern Med* 2012;156:609-17.
26. SEER Cancer Statistics Review - Lifetime Risk (Percent) of Being Diagnosed with Cancer by Site and Race/Ethnicity. Females, 17 SEER Areas, 2005-2007, table 1.16.
27. SEER Cancer Statistics Review - Lifetime Risk (Percent) of Dying from Cancer by Site and Race/Ethnicity. Females, Total U.S., 2005-2007. Table 1.19.
28. Smith-Bindman R, Chu PW, Miglioretti DL, Sickles EA, Blanks R, Ballard-Barbash R, et al. Comparison of screening mammography in the United States and the United kingdom. *JAMA* 2003;290:2129-37.
29. Coburn N, Fulton J, Pearlman DN, Law C, DiPaolo B, Cady B. Treatment variation by insurance status for breast cancer patients. *Breast J* 2008;14:128-34.
30. El-Bastawissi AY, White E, Mandelson MT, Taplin S. Variation in mammographic breast density by race. *Ann Epidemiol* 2001;11:257-63.
31. Henderson LM, Hubbard RA, Onega TL, Zhu W, Buist DS, Fishman P, et al. Assessing health care use and cost consequences of a new screening modality: the case of digital mammography. *Med Care* 2012;50:1045-52.

APPENDIX

Appendix Table 11.1. Effects, costs, and cost-effectiveness by race, modality, interval, and starting age per 1000 women age 40

MISCAN-Fadia

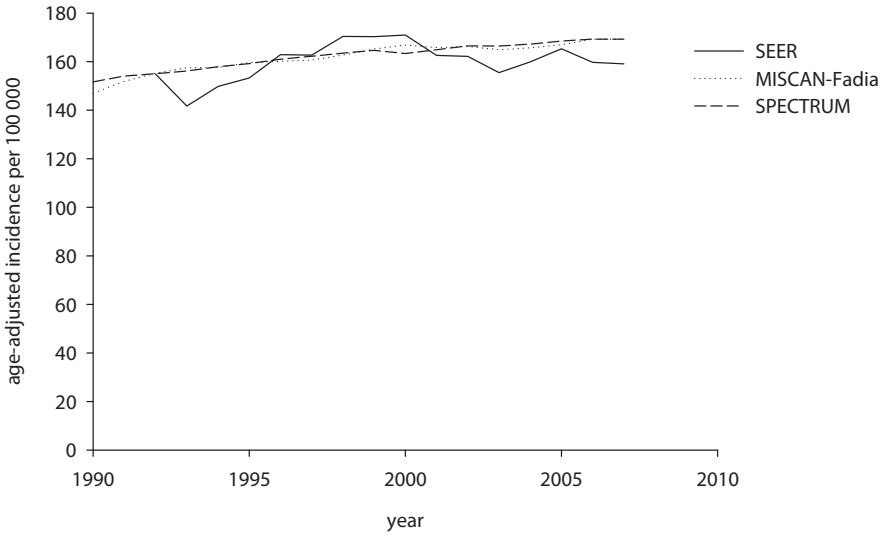
race	modality	screening scenario				effects	total costs	cost-effectiveness*	incremental cost-effectiveness*	
			number of mmg	bc deaths averted	mortality reduction	LYG	(US$2010) x 1000	$/lyg	cf. previous	digital vs. film
White	film	once	929	0.5	1.6%	9	94	11,000	-	
		biennial_50	6872	3.1	9.7%	53	546	10,000	10,000	
		biennial_40	11746	4.5	13.9%	90	1034	12,000	13,000	
		annual_50	13623	4.4	13.4%	76	1176	16,000	*dominated*	
		annual_40	23322	6.2	19.1%	127	2173	17,000	31,000	
	digital	once	929	0.5	1.7%	9	148	16,000		106,000
		biennial_50	6871	3.3	10.1%	56	832	15,000	15,000	112,000
		biennial_40	11745	4.7	14.5%	95	1532	16,000	18,000	107,000
		annual_50	13619	4.5	14.0%	79	1734	22,000	*dominated*	173,000
		annual_40	23318	6.5	19.9%	132	3142	24,000	42,000	168,000
Black	film	once	893	0.8	2.3%	12	67	6,000	-	
		biennial_50	6560	4.3	13.3%	70	400	6,000	6,000	
		biennial_40	11401	6.1	18.8%	116	837	7,000	9,000	
		annual_50	13017	5.8	17.8%	96	969	10,000	*dominated*	
		annual_40	22652	8.1	25.1%	158	1897	12,000	25,000	
	digital	once	893	0.8	2.4%	13	115	9,000		78,000
		biennial_50	6559	4.5	13.8%	73	665	9,000	9,000	94,000
		biennial_40	11399	6.4	19.6%	121	1309	11,000	13,000	93,000
		annual_50	13014	6.0	18.4%	100	1490	15,000	*dominated*	153,000
		annual_40	22648	8.4	26.0%	165	2820	17,000	35,000	150,000
Hispanic	film	once	946	0.3	1.6%	6	110	18,000	-	
		biennial_50	6990	2.1	9.9%	36	640	18,000	18,000	
		biennial_40	11900	3.0	14.2%	61	1167	19,000	21,000	
		annual_50	13886	2.9	13.7%	51	1316	26,000	*dominated*	
		annual_40	23674	4.2	19.5%	85	2370	28,000	49,000	
	digital	once	946	0.4	1.7%	6	163	25,000	-	163,000
		biennial_50	6990	2.2	10.4%	38	936	25,000	25,000	186,000
		biennial_40	11899	3.2	14.9%	64	1680	26,000	28,000	173,000
		annual_50	13884	3.1	14.3%	53	1892	36,000	*dominated*	280,000
		annual_40	23671	4.3	20.3%	89	3367	38,000	68,000	286,000

* rounded to the nearest 1,000

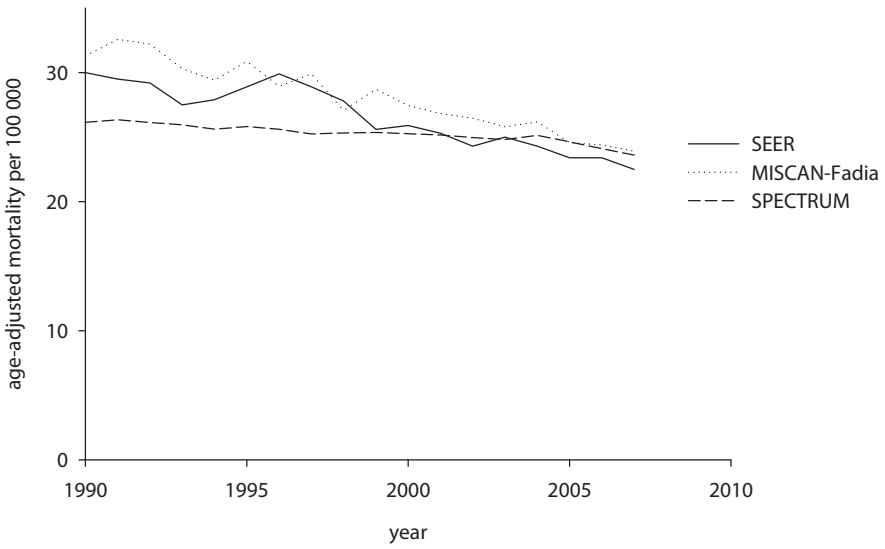
SPECTRUM

race	modality	screening scenario	number of mmg	bc deaths averted	mortality reduction	effects LYG	total costs (US$2010) x 1000	cost-effectiveness* $/lyg	incremental cost-effectiveness* cf. previous	digital vs. film
White	film	once	923	1.2	3.4%	19	-33	-2,000	-	
		biennial_50	7242	4.9	14.0%	78	120	2,000	3,000	
		biennial_40	12143	6.2	17.8%	113	465	4,000	10,000	
		annual_50	13542	5.7	16.3%	91	733	8,000	*dominated*	
		annual_40	23309	7.3	21.1%	135	1618	12,000	50,000	
	digital	once	923	1.2	3.5%	19	17	1,000		114,000
		biennial_50	7241	5.0	14.3%	79	431	5,000	7,000	222,000
		biennial_40	12141	6.3	18.1%	115	1000	9,000	16,000	210,000
		annual_50	13540	5.7	16.5%	92	1329	14,000	*dominated*	572,000
		annual_40	23306	7.4	21.4%	138	2661	19,000	74,000	472,000
Black	film	once	884	1.2	3.6%	19	-44	-2,000	-	
		biennial_50	6891	4.9	15.1%	79	127	2,000	3,000	
		biennial_40	11751	6.5	19.9%	119	483	4,000	9,000	
		annual_50	12891	5.7	17.4%	91	721	8,000	*dominated*	
		annual_40	22567	7.6	23.4%	143	1619	11,000	49,000	
	digital	once	884	1.2	3.7%	20	0	0,000		93,000
		biennial_50	6890	5.0	15.4%	80	426	5,000	7,000	176,000
		biennial_40	11749	6.6	20.4%	123	1005	8,000	14,000	161,000
		annual_50	12890	5.8	17.7%	93	1288	14,000	*dominated*	401,000
		annual_40	22564	7.8	23.8%	145	2631	18,000	72,000	421,000
Hispanic	film	once	942	0.7	3.3%	11	44	4,000	-	
		biennial_50	7375	2.7	13.4%	44	561	13,000	16,000	
		biennial_40	12297	3.6	17.5%	66	988	15,000	19,000	
		annual_50	13806	3.2	15.6%	51	1277	25,000	*dominated*	
		annual_40	23625	4.3	20.9%	80	2204	27,000	86,000	
	digital	once	942	0.7	3.5%	12	94	8,000	-	157,000
		biennial_50	7374	2.8	13.6%	44	899	20,000	25,000	611,000
		biennial_40	12295	3.7	17.9%	68	1550	23,000	27,000	314,000
		annual_50	13805	3.2	15.8%	52	1903	37,000	*dominated*	694,000
		annual_40	23623	4.3	21.1%	81	3353	41,000	135,000	1048,000

* rounded to the nearest 1,000



Appendix Figure 11.1. Age-adjusted incidence rates over time as observed (SEER) and predicted by MISCAN-Fadia and SPECTRUM for Hispanic U.S. women aged 25 years and older



Appendix Figure 11.2. Age-adjusted breast cancer mortality rates over time as observed and predicted by MISCAN-Fadia and SPECTRUM for Hispanic U.S. women aged 25 years and older

Chapter 12

General discussion

12.1 MAIN FINDINGS

Impact of mammography screening in the U.S. population

Measuring the effect of breast cancer screening in the population using observational data is not easy and has to take into account several potential biases, as well as specific issues relating to screening.[1,2] We found that an adequate follow-up time and adequate time after implementation of a screening program are crucial when evaluating the effects of screening. Measuring effects too early can result in an underestimation of the effect of breast cancer screening on breast cancer mortality. A measured effect of 9.9% in Norway after a short follow-up time is likely to increase to 16.2% if the follow-up time is prolonged by 5 years.

Breast cancer screening, followed by early treatment, can reduce breast cancer mortality substantially. We found that breast cancer screening has reduced breast cancer mortality in a recent period (2004-2006) by 18-22% in U.S. women age 25 years and older. Equal use of screening in Black and White women can, however, not eliminate the racial disparity in breast cancer mortality. When Black and White women have the same screening rates, the disparity is decreased only to a small extent (7-8%). This reflects the finding that screening rates are currently already quite similar in Black and White women.[3] We found that the majority of the racial disparity in breast cancer mortality could be attributed to variations in the natural history (26-44%) of the disease and yet unknown factors (38-46%).

When looking at the potential of screening to reduce the number of breast cancer deaths in the future compared to other interventions, we found that, compared to continuation of current trends, maximal reductions in breast cancer deaths could be achieved by 2025 through optimizing treatment use, followed by increasing screening use and then prevention of obesity. Thus, increasing adjuvant therapy use has the largest effect, which might indicate that current use diverges more from optimal for adjuvant treatment than for mammography screening in the U.S. The effect of obesity elimination is smaller, although it is the only intervention that can also reduce the number of breast cancer cases (incidence) and has important effects on the risks for other diseases.

We found that obesity accounts for 3-4% of breast cancer cases and 3-9% of breast cancer deaths in U.S. Black and White women. Even though Black women have a higher obesity prevalence in most age groups, elimination of obesity will not reduce the disparity in breast cancer mortality, because obesity has a protective effect in pre-menopausal women and obesity prevalence is higher in White women beyond age 74 years.

In sum, breast cancer screening has had a substantial effect in the U.S. female population and is expected to continue to do so in the future. The racial disparity in breast cancer mortality cannot be explained by racial differences in obesity prevalence, and only a small part (7-8%) can be explained by differences in screening use. Future research aimed at racial variation in breast cancer biology and racial differences in actual treatment utilization is needed.

Effects of different mammography screening scenarios

There are multiple factors that influence the balance between benefits and harms of mammography screening. We investigated the benefits and harms of different starting ages, stopping ages, screening intervals, and screening modalities (film and digital mammography). We compared the effects to the screening scenario recommended by the U.S. Preventive Services Task Force (USPSTF) since 2009 (partly based on our analysis) which is biennial screening between age 50 and 74 years.

With regard to the starting age of screening, important factors are that younger women have a lower risk of developing disease (within the next 10-years) and that the sensitivity of the screening test is lower,[4] due to more dense breast tissue.[5] In addition, the specificity of the screening test is lower,[4] while the tumor growth rate is higher.[6] Taking these factors into account, we found that women age 40 with a two-fold risk (compared to average) can expect the same balance of benefits and harms as average-risk women age 50 from starting biennial screening. Factors that lead to a two-fold risk are having extremely dense breasts (density category Breast Imaging-Reporting and Data System (BI-RADS) 4) and having a first degree relative with breast cancer.[7] In the U.S. 13% of women in their forties have extremely dense breast and 9% of women in this age group have a first degree relative with breast cancer.[7]

When assessing a possible upper age limit of screening, there are several factors that influence the balance between benefits and harms. Older women have a higher risk of developing disease, and the sensitivity of the screening test has also been found to be higher.[8] On the other hand, their remaining life expectancy is shorter, limiting the potential to save a life and leading to more overdiagnosis (i.e., the detection of tumors that would not have been detected in a woman's lifetime in the absence of screening). An additional complicating factor is that there is limited information about the benefits and harms of screening beyond age 74 years, because none of the randomized controlled trials included women older than 74 years. We found that despite the limited knowledge, which was reflected in model differences, the age at which harms began to outweigh benefits was the same across models; beyond age 90 years harms began to outweigh benefits because of the increasing amount of overdiagnosis. The balance between benefits and harms was less favorable after age 74 and depended on a woman's comorbidity level. Healthy women (with no or mild comorbidity) can expect more benefit from a screen at age 74 than women with (moderate or severe) comorbidities. Women with comorbidities might therefore choose a younger stopping age than women without comorbidities.

With regard to screening interval, the more frequent a woman has screening exams, the larger the probability of having an earlier diagnosis, and the larger the benefit from screening might be. However, the harms (false positive exam, overdiagnosis) will also increase with more frequent screening. A randomized trial, the UK Breast Screening Frequency Trial, did not show a significant difference in breast cancer mortality between screening every year and screening once every three year.[9,10] To be able to replicate the trial in our microsimulation

model, we had to assume a relatively low screening test sensitivity. In addition, we found evidence that the trial was underpowered to find a statistically significant difference between the two groups. Our results indicated that there was benefit of shortening the screening interval, although there seemed to be a 'diminishing return on investments'. This is in line with what we found for annual vs. biennial screening intervals. For example, biennial screening between age 50 and 74 years maintained most of the benefit (72-95%) of annual screening between age 50 and 74 years, while almost halving the harms (e.g., number of false positives) of screening.

When comparing screening modalities, digital mammography has been found to be more sensitive than film mammography, especially in younger women, but is also more expensive.[11] We found that when performing the same number of screens, digital mammography saved slightly more lives than film mammography, but also increased the number of false positive exams and costs of screening and diagnostics. Therefore, the harm-benefit ratio and cost-effectiveness was found to be more favorable for film than for digital mammography. In addition, with a fixed budget, such as in a budget-fixed program for low-income, uninsured women in the U.S., replacing film by digital mammography would result in fewer women that can be served (25-26%), and fewer lives that are saved (23-24%). The loss in life-years could be reversed to a 8-13% increase by restricting the screening interval to biennial. Thus, price reduction of digital mammography, budget increases, or changes in the program, such as prolonging the screening interval, are warranted to achieve comparable health effects of film and digital mammography within the program.

12.2 METHODOLOGICAL CONSIDERATIONS

The results presented in this thesis are based on microsimulation models. This is one of the main strengths as well as a possible weakness of this thesis.

On the one hand, we used a microsimulation model that had previously been developed and validated. This model had previously been shown to be able to reproduce breast cancer incidence and mortality in the U.S. over time (1975-2000).[12,13] An advantage of using a well-established simulation model is that a model can be used to translate the results from randomized trials to a current situation. Simulation models can be seen as a way to synthesize the best available data from different sources. For example, models can incorporate the evidence from randomized controlled trials and extrapolate the effects from trials to other situations (e.g., other attendance rates, background incidence, treatment effects and test sensitivities) and compare these predictions with observations. In addition, in microsimulation models individual life histories can be simulated with and without screening. In this way, two identical groups, only differing in whether they are screened or not, can be compared, an mimicking an ideal trial.

Furthermore, a model can be used to estimate the lifetime benefits and harms, which is difficult in a trial; it requires a very long follow-up period and no screening in the control group, whereas in trials the control group is usually offered screening after a certain period. In addition, in most chapters we used multiple models, which enabled us to give a range of plausible outcomes. The use of multiple models, using similar inputs, also served as a sensitivity analysis on model structure.[14] Moreover, the use of these models made it possible to do an extensive evaluation of various screening scenarios taking into account benefits as well as harms as well as estimating the effect of screening in the population.

On the other hand, the use of models is also associated with several limitations. The most obvious one is that models are not perfect, as they are simplifications of reality. We used MISCAN-Fadia, which is a model that is based on continuous tumor growth. The modeled events (e.g., screen detection, clinical diagnosis, and curability) are defined in terms of tumor size.[13] Although the model is biologically plausible and intuitive, a large part of the tumor growth is unobservable, making it difficult to use direct model inputs. The biological model structure leads to a less direct relationship between empirical data and model parameters.[13] For example, breast-cancer specific survival by stage or stage-distributions cannot be used as direct model inputs, but are used rather to estimate the underlying model parameters.

Another obvious limitation is that the quality of the model outputs depends on the quality of the model inputs. In all studies we used the best available data from reliable sources, such as data from Surveillance, Epidemiology and End Results (SEER) Program and the Breast Cancer Surveillance Consortium (BCSC). However, when trying to include subgroups in the model data can be lacking. Even when using data from a large population, such as the U.S. the data can get sparse for certain subgroups. For example, the models use data on the use of different treatments by calendar year, age, tumor stage, estrogen receptor (ER) status and human epidermal growth factor receptor 2 (HER2) status, obesity, and race/ethnicity as input, but for some combinations the sample sizes were small. Including more subgroups in the future would even enlarge the difficulty of obtaining data for each combination of variables.

Another difficulty with regard to the use of data is that for a large part of the natural history of disease no direct data is available, because the disease progression is unobservable. This is particularly the case for data on ductal carcinoma *in situ* (DCIS) progression rates. Little is known on the natural history of DCIS, because DCIS is usually removed as soon as it is detected. There is evidence for progression of DCIS from studies in women with low grade DCIS which is initially mistakenly diagnosed as benign. These studies report that 14%–60% of those women develop invasive cancer within 10–20 years.[15-17] There is, however, also evidence that not all DCIS become invasive, for example from autopsy studies that found a prevalence of DCIS of 0-15% in women not known to have had breast cancer.[18] However, these lesions are not representative of the DCIS found by mammography. Our results do not provide additional information on the natural history of DCIS, since several models with different assumptions on the natural history of DCIS can adequately replicate incidence trends

over time. Future studies focusing on molecular markers or clinical factors that can be used to identify patients at risk of future events of DCIS and invasive breast cancer are needed.

12.3 CURRENT STATUS OF BREAST CANCER SCREENING

Controversy around breast cancer screening

The value of breast cancer screening has been discussed extensively.[19-21] The discussion focuses on two issues: the magnitude of the benefit in terms of a breast cancer mortality reduction and the magnitude of harm, in particular overdiagnosis. The debate arises mainly because typically there is no appropriate control group anymore; once a screening program is implemented it is unknown what would have happened in the absence of screening. In this section we discuss some of the recently published studies on breast cancer screening with regard to the mortality reduction and overdiagnosis.

Recently, the independent UK panel performed an extensive evaluation of the evidence on breast cancer screening for the benefits and the harms and chose to rely mainly on findings from randomized trials, because they felt the observational studies might be influenced by several biases. The panel concluded that for the UK setting the evidence suggest a 20% breast cancer mortality reduction accompanied by 11% overdiagnosis (i.e., in women invited for screening 11% of the of the cancers diagnoses in their lifetime constitute overdiagnosis).[2]

With regard to the benefit of screening, it has been argued that the findings of randomized trials, which were performed a few decades ago, are no longer valid, because they do not accurately reflect current practice due to improvements in breast cancer treatment and increased awareness of breast cancer. Therefore, several study designs have been used to estimate the effects of screening on breast cancer mortality using observational data. Study designs that have been used include trend studies (or ecological studies), incidence-based-mortality approaches, and case-control studies.

In trend studies, it is difficult to assess the causative impact of screening, because multiple other factors besides screening influence breast cancer mortality trends, such as changes in treatment and risk factors. In addition, the estimated effect might be diluted because it usually takes a number of years before a screening program is fully implemented. Therefore, analyzing breast cancer mortality rates over time is of limited value for the assessment of the impact of screening impact. Despite these limitations, many trend studies have been published recently trying to estimate the effect of screening.[22,23] Many of these studies question the benefit of screening, while ignoring the limitations of their study design. Others have argued that other methods and individual data are necessary to properly quantify the screening effect [24] and the independent UK panel highlighted the inappropriateness of trend studies.[2]

Studies that use an incidence-based-mortality approach and case-control studies usually use individual data and most studies try to control for differences between screened and unscreened women (selection bias), although residual bias might still influence their estimates.[1,25,26] Recently, the effects of screening estimated by studies using different types of study designs were compared. It was found that the estimates from well-executed observational studies, using different well-executed study designs, indicated a breast cancer mortality reduction of 25–31% for women invited for population-based screening and 38-48% for women actually screened.[1]

There are many factors that influence the measured benefit of screening. Even when evaluating the evidence of randomized trials several factors need to be taken into account, such as the attendance rate, contamination, number of screens, screening ages, screening interval, and screening quality.[27,28] Most studies compare the total number of breast cancer deaths after a certain follow up period, but most reported results do not take into account the time-specific effect of screening.[29,30] The measured effects of screening are a weighted average of several cohorts and ages. When starting a screening program it takes a considerable amount of time before the program is fully implemented and additionally many years before the first cohort has undergone all screens and has had the potential to fully benefit from the program.

To illustrate the time-dependent effect of screening on breast cancer mortality we modeled the implementation of a hypothetical screening program of biennial screening starting at age 50 years and ending at age 74 years with an attendance rate of 80%. From 1975 onwards all women are invited to screening. The measured screening effect clearly varies by time, cohort and age (Figure 12.1). In certain age groups, the reduction in breast cancer mortality can be as high as 38%, which reflects 48% for women actually attending. However, when measured as a total effect for all ages and cohorts combined, the effect of screening is considerably lower (reaching a maximum of 24% in 2020). It is also apparent that it takes a long time before the full effect becomes visible and measurable on the population level. Even after 25 years of screening, the total effect of screening is still increasing. Thus, large breast cancer mortality reductions can be found in subgroups (certain age groups) of women having undergone regular screening, whereas the effect will be smaller when based on a larger age range or in women having undergone only part of the screening program, especially if the effect is measured shortly after the implementation of the screening program.

With regard to estimating the amount of overdiagnosis associated with breast cancer screening also many issues have to be taken into account. A wide range of estimates of up to 54% has been published in recent years.[31-33] There are several reasons for the wide range of published estimates. It has been shown that overdiagnosis is overestimated when calculations are derived from the implementation period of a screening program and when there is insufficient follow-up to observe a reduction in breast cancer incidence in women above

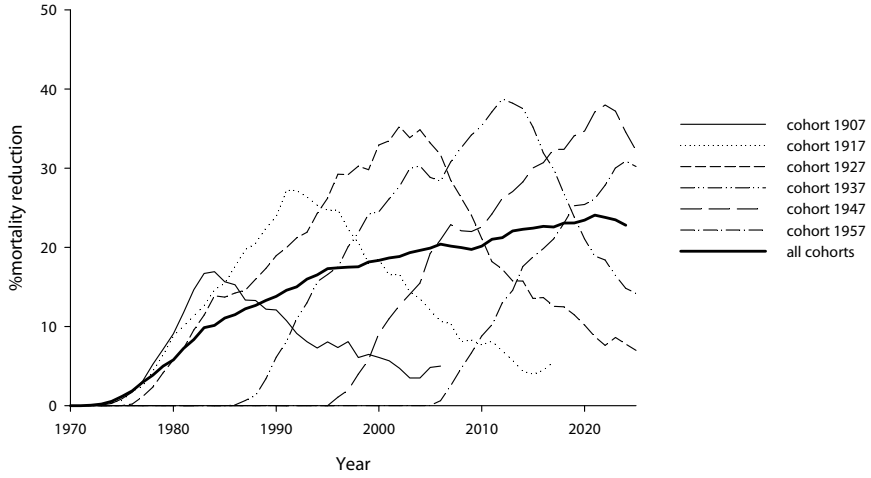


Figure 12.1. Effect of screening by cohort and time.
We modeled the implementation of a hypothetical screening program of biennial screening starting at age 50 years and ending at age 74 years with an attendance rate of 80%. From 1975 onwards all women are invited to screening.
The effect of screening is defined as the percent reduction in breast cancer mortality (breast cancer mortality in a situation without screening – breast cancer mortality in a situation with screening divided by breast cancer mortality in a situation without screening). Women born in 1907 (cohort 1907) are age 68 years in 1975 when screening starts and are thus only screened from age 68 until 74 years and do not experience the full benefit of screening. In contrast, women born in 1927 (cohort 1927) are age 48 years in 1975 and are thus screened from age 50 years (in 1977) until age 74 years (in 2001) and can expect the full benefit of screening. For this cohort the largest effect is visible from 2000 up to 2005 when these women are in their early seventies. The effect of screening is slightly higher for cohort 1937, because of the increasing screening sensitivity over time (1975-2000). For cohorts of women born after 1937, the effect of screening is roughly the same as for cohort 1937.
On the population level (all cohorts) the effect of screening is a weighted average of several cohorts and ages and is expected to increase over time. It takes a long time before the screening effect reaches its maximum on the population level.

the upper age for screening.[34] Similarly, the range of overdiagnosis estimates is considerably smaller and the estimates are lower (1% to 10%) when only studies that adequately adjust for lead time and changes in breast cancer risk are included.[35]

A recent study on overdiagnosis in the U.S. estimated that 31% of breast cancers diagnosed in 2008 were overdiagnosed.[36] This study highlighted that the estimated overdiagnosis depends on the extent to which background breast cancer incidence increased over time. The estimate of 31% diagnosis was based on an increase of 0.25% per year in background incidence. If an 'extreme' increase of 0.50% per year was assumed, the estimated overdiagnosis reduced to 26%. However, based on data from a longer time period, the increase in breast cancer incidence seems to be considerable higher than the 'extreme' assumption (Figure

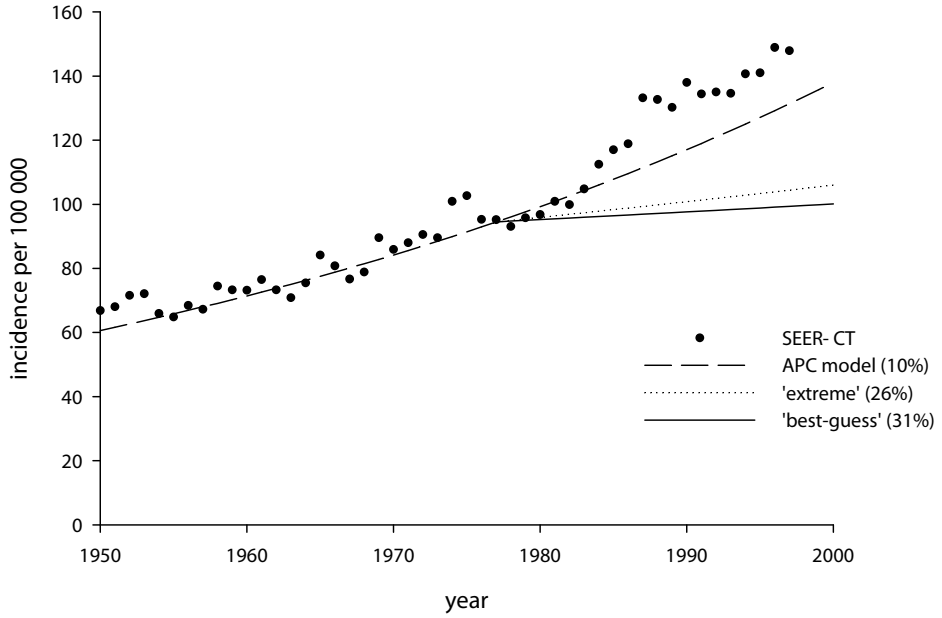


Figure 12.2. Breast cancer incidence per 100 000 women over time. The estimate of overdiagnosis is shown for each scenario (%) as the proportion of breast cancers detected in 2008 that are overdiagnosed.
SEER-CT: Connecticut registry
APC model: increase in incidence through 2000 based on an age-period-cohort model [37]
'best guess': Bleyer and Welch's best guess of an increase of 0.25% per year
'extreme': Bleyer and Welch's 'extreme' increase of 0.50% per year

12.2). Bleyer and Welch's assumed breast-cancer incidence in the absence of screening is based on changes in rates for women under age 40 and ignores longstanding cohort-specific trends. If, however, the background incidence is based on an extensive age-period-cohort model, taken into account a longer period, incidence has been found to be rising about 1.67% annually.[37] Applying a 1.67% annual increase through 2000 in Bleyer and Welch's calculation reduces their estimate of over-diagnosis to about 10%. Thus, when data from a longer time period is considered, a stronger increase in breast cancer incidence is seen in the U.S. Taking into account this stronger increase would lead to a decrease of the estimated overdiagnosis from 31% to 10% of all breast cancers diagnosed in 2008 in the U.S. being overdiagnosed.

Breast cancer screening in Europe and the United States

There are many differences between Europe and the U.S. in how breast cancer screening is organized, funded, and viewed by the public. In Europe, many countries have implemented nationwide screening programs, inviting all eligible women within a certain age group (usually 50-69) every other year to mammography. In the U.S, however, no centrally organized breast cancer screening program exists, leading to a variety of screening behaviors. Some

women are screened every year starting from a young age, whereas others start later and are screened irregularly, depending for an important part on a woman's health insurance status.

With regard to radiological reading procedures, European guidelines recommend independent double reading with consensus and recommend that radiologists read 5000 screening mammograms per year,[38] while in the U.S. single reading is the usual practice and the Mammography Quality Standards Act requires radiologists to read only 960 mammograms every two years. It has been found that single vs. double reading influences screening performance.[39] An additional factor that might influence performance is that the U.S. has a high rate of malpractice lawsuits for missing breast cancer,[40] while in Europe those are rare.[41]

In addition, most European countries have population-based cancer registries which can often be linked to the screening program. Moreover, the implementation of a screening programs is often accompanied by changes in the organization of breast cancer care, including quality assurance activities such as training and audits, together with the setting up of specialist breast units for management of breast lesions.[38] In the U.S., there are multiple insitutions that collect data. For example, the SEER program began collecting data in 1973 with a limited amount of registries, is still expanding, and currently covers approximately 28% of the U.S. population. In addition, the BCSC collects very detailed information on patient factors, detected breast cancers, screening behavior, and treatment.

Previously in the U.S., many people viewed cancer screening as a simple and safe way to save lives and public health campaigns were aimed at maximizing the uptake of screening, using persuasive messages, without mentioning the harms of screening.[42] Consequently, the public enthusiasm for cancer screening was high.[43] Recently, more attention has emerged for the harms of screening leading to the acknowledgement for the need to increase the awareness of the benefits and harms to encourage informed personal decisions rather than to persuade women to participate in screening. The most recent USPSTF recommendation for breast cancer screening [44] reflected this awareness of harms leading to a less intense recommended screening scenario more in line with Europe's policy.

All these factors might influence the impact of a screening program on breast cancer mortality and on harms, such as false positives and overdiagnosis. For example, it has been found that the sensitivity and specificity were higher in Europe (Norway) than in the U.S., [39] indicating that screening in the U.S. might be accompanied by more harm (false positives) without additional benefit. A comprehensive international comparison of screening programs with different characteristics might point to additional areas for improvements in each country's program.

12.4 FUTURE DIRECTIONS

Future developments and their influence on the effects of screening

There are several future developments that might influence the benefits and harms of screening. For illustrative purposes the predicted effects of potential future developments are presented in terms of benefits (lives saved), harms (overdiagnosed women) and the balance between the benefits and harms (their ratio) (Table 12.1). The results presented here should not be interpreted as predictions of the future, but rather are meant to illustrate the direction of change (increase or decrease) of potential future developments on the benefits and harms of screening.

First of all, the implementation of other screening modalities (such as ultrasound or tomosynthesis) and the improvement of currently used modalities might lead to an increase in screening test sensitivity.[45,46] A higher screening test sensitivity will increase the number of lives saved per 10 000 screens, but will also increase the number of overdiagnosed women. The balance between lives saved and overdiagnosis is roughly unaffected (slightly more favorable).

Further, it is expected that the effectiveness of adjuvant therapies will improve in the future. Usually it is assumed that when treatment is more effective, screening might be less effective.[47] When including a larger effect for treatment in the model, the benefit of screening indeed becomes smaller, while the amount of overdiagnosis remains stable. The extent to which the screening effect is reduced presumably depends on the differential improvement of treatment by stage. If treatment for late stage disease is improved so much that late stage disease has a comparable survival as early stage disease, then there is less to add for screening. However, if mainly treatments for early stage are improved, then the effect of screening might become even larger.

Additionally, it is assumed that awareness about the importance of seeking care for breast abnormalities has been increasing and might increase further in the future. This means that women report symptoms earlier than before and that even in the absence of screening there

Table 12.1. The influence of possible future developments on the number of lives saved, overdiagnosed women, and their ratio (all outcomes per 10,000 screens)

	Lives saved	change	Overdiagnosis	change	Ratio (Overdx: life saved)	change
Reference: 50-74 biennial film	8.0		11.0		1.38	
Improved sensitivity	8.4	↑	11.1	↑	1.32	↓
Improved adjuvant therapy	7.4	↓	11.0	=	1.48	↑
Increased risk	9.4	↑	12.9	↑	1.37	=
Increased awareness	7.1	↓	10.6	↓	1.50	↑
Increased life expectancy	8.2	↑	10.3	↓	1.25	↓

Assumptions: a cohort of U.S. women born in 1960 with optimal (100%) screening and treatment use.

is a shift to the detection of more early stage disease. If this will continue in the future, the increased awareness will lead to a reduced benefit of screening and a somewhat less favorable balance between lives saved and overdiagnosis. However, in the U.S. and most other western countries breast cancer awareness is already quite high, so in practice the effects might be limited.

Also, breast cancer incidence might increase in the future due to changes in risk factor prevalence. For example, more women get children at a later age,[48] alcohol use among women is increasing,[49] and women are less physically active.[50] A higher risk of developing breast cancer will lead to an increased number of lives saved per 10 000 screens and an increase in number of overdiagnosed women with the same extent. The balance between lives saved and overdiagnosis is therefore largely unaffected by a higher risk of developing breast cancer.

Finally, life expectancy has increased and is predicted to increase further in the future, although less rapidly than in the past.[51] With a longer life expectancy, the benefits of screening will increase, while the amount of overdiagnosis will decrease, resulting in a more favorable balance between lives saved and overdiagnosis.

Thus, in the future, an increased awareness and improvements in adjuvant therapy will lead to a slightly less favorable balance of lives saved and overdiagnosed women, whereas an improved sensitivity and life expectancy will lead to a more favorable balance between lives saved and overdiagnosed women. The total impact of these potential future developments and how they might interact remains to be investigated.

Personalizing screening

We found that the benefits and harms of screening vary substantially by characteristics, such as risk, race/ethnicity, obesity, and health status (comorbidity). Targeting screening to subgroups with the highest potential benefit and the lowest potential harm can improve the overall balance between benefits and harms in the population.

The idea of personalizing screening is that women are screened in line with their preferences for specific benefits and harms (the weight they attach to specific benefits and harms), and that the information given to these women is specific to their personal probabilities of benefiting or being harmed by screening. The application of targeted screening will allow screening strategies to be tailored to different groups in the population. Eventually, it may be possible to move from one-size-fits-all in terms of screening policies to more individualized screening approaches.

An important next step towards an individualized approach includes incorporating available information on breast density. Breast density influences breast cancer risk and screening test sensitivity, and is therefore important to take into account, although a complicating factor in using breast density in screening stratification is that it requires a baseline mammogram to know a woman's breast density. Dense breast tissue has been found to be associated

with a higher breast cancer incidence and mortality,[52] indicating that women with dense breast tissue might benefit from shorter screening intervals. It is unclear whether the relation between breast density and breast cancer risk varies according to breast cancer subtype.[53,54] More information is needed on the biological mechanisms of how breast density influences risk.

In addition, it has been widely acknowledged that breast cancer is a heterogeneous disease and more knowledge is emerging on distinct molecular subtypes.[55,56] ER and HER2 status have been found to be important in targeting the treatment of disease, but might also be important in the screening context. Certain factors increase risk for the development of a specific molecular subtype of breast cancer. Information on the incidence, disease progression and survival in terms of risk factors for the development of specific molecular subtypes is needed to be able to estimate the effect of targeted screening based on risk for specific subtypes. Eventually, women with risk factors that increase risk of aggressive disease might choose more intense screening than women at low risk.

In general, risk is probably the most important predictor of the absolute magnitude of benefit of screening. The application of risk models may allow screening strategies to be tailored to different risk groups in the population. For example, women at high risk might be offered screening starting at a younger age, more frequent screening or MRI screening, whereas women at lower risk might consider less intense screening. This may, in turn, improve the balance between the benefits and harms of screening for breast cancer. Targeted screening might be based on an extensive risk prediction model, such as the Gail or the BCSC risk model.[57,58] However, the predictive accuracy of the current models is quite low, and individual risk prediction is very limited,[57-59] although on a population level risk might be used to stratify groups.

Another approach would be personalizing screening based on a genetic risk profile. Recent genome-wide association studies (GWAS) have identified various breast cancer susceptibility variants.[60] In contrast to BRCA1 and BRCA2, these susceptibility variants have weak effects and contribute to small increases in breast cancer risk individually, but combining these susceptibility loci in risk models might be used for risk stratification. The currently known risk alleles do not provide sufficient discrimination to warrant individualized prevention, but they might be used for risk stratification at the population level.[61,62] In order to be able to estimate the effects of targeted screening based on genetic risk, it is important to know whether and how the natural history of the disease varies by genetic risk profile.

Finally, it is crucial to factor in the values of individual women who are faced with making a screening decision. Only a small part of the individuals undergoing a screening test will benefit, while the majority will not benefit. Some might experience harm. In addition, individuals value benefits and harms differently. It is, therefore, crucial that the eligible population is informed about the benefits as well as the harms of screening. It has, however, been found that understanding cancer screening statistics is not easy [63] and not everyone is aware

of the potential harms of screening. For example, a recent study in Australia found that few women were aware of the possibility of overdiagnosis, but after having been informed, many considered it important to take overdiagnosis into account when making choices about whether to have screening or not.[64] In the Netherlands, where information on overdiagnosis is included in the leaflet enclosed with the invitation letter, 67% of women receiving a first invitation responded correctly on a question about overdiagnosis.[65] This might indicate that if information on the harms of screening is provided to women, their knowledge increases and it emphasizes the need to provide women with information on the benefits as well as harms of screening.

Moreover, information on the benefits and harms should be used to support the invited women to decide about participation in a screening program based on an informed decision. It is generally assumed that informed decision making leads to decisions that are in line with the patient's values, reduce uncertainty or decisional conflict, improve patient satisfaction and ultimately lead to an improved quality of life. To facilitate making an informed decision, decision aids can be used.[66] A decision aid provides structured evidence-based information on the benefits and harms of an intervention. Several studies on breast cancer screening found that a decision aid increased knowledge and reduced decisional conflict.[67,68] The effect of making an informed decision about participation in breast cancer screening on quality of life is, however, uncertain. A further complicating factors is that the current (generic) instruments to measure quality of life might not be sensitive enough to detect small changes in quality of life due to undergoing screening and/or having a false-positive exam.

Future research efforts should focus on what information is desired by women and their primary caregivers, how it should be provided to them in order to ensure that they can make an informed decision about whether or not to participate in screening, and, finally, whether making an informed decision increases their quality of life.

Recommendations for future research

Priorities for future research include:

- Estimating the combined impact on the efficacy of screening of ongoing developments such as changes in risk factors, improvements in adjuvant therapy, implementation of new screening modalities, and improvements in life expectancy
- Incorporating information on the risk for the development of specific molecular subtypes (ER and HER2) and how subtypes influence the natural history of disease in models, in order to assess the effects of breast cancer screening and adjuvant therapy by molecular subtype
- Evaluating the effects of targeted screening based on either 'classical' risk factors or genetic risk profiles versus untargeted screening based only on age and sex

- Obtaining more information on racial variation in breast cancer biology and racial differences in actual treatment utilization in order to better understand the racial disparity in breast cancer mortality
- Performing an international comparison of the characteristics of different screening programs and outcomes to determine the successful components and areas for improvement
- Assessing internationally how information on the benefits and harms of screening can be best communicated to women in order to enable them to make an informed decision about screening participation and whether making an informed-decision influences their quality of life

12.5 CONCLUSIONS & RECOMMENDATIONS

- Because of the time-dependent effect of screening, the use of a short follow-up time after implementation of the screening program underestimates the 'true' effect of screening on breast cancer mortality.
- Breast cancer mortality is currently 20% lower in U.S. women age 25 years and older than it probably would have been without mammography screening.
- The racial disparity in breast cancer mortality cannot be explained by racial differences in obesity prevalence, and only a small part (7-8%) can be explained by differences in screening use.
- Women with a 4-fold increased risk can expect the same balance of benefits and harms from starting annual screening at age 40 years as women with a 2-fold risk from biennial screening starting at age 40 years, and as average-risk women from biennial screening between age 50 and 74 years.
- The balance between benefits and harms of screening is less favorable after age 74 years and varies by comorbidity level. Healthy older women can expect more benefit and less harm from screening than women of the same age who have comorbidities
- Doubling the screening frequency (annual vs. biennial), doubles the number of screens, almost doubles the number of false-positives, but does not double the number of lives saved or life-years gained by screening; they roughly increase by a factor 1.4
- Digital mammography can save more lives than film mammography, but it increases the number of false-positives and is more expensive. The harm-benefit ratio and cost-effectiveness ratio of screening are, therefore, not improved by the implementation of digital mammography.
- Targeting screening to subgroups with the highest chance for benefit and/or the lowest chance for harm has the potential to improve the balance between benefits and harms at the population level.

REFERENCES

1. Broeders M, Moss S, Nystrom L, Njor S, Jonsson H, Paap E, et al. The impact of mammographic screening on breast cancer mortality in Europe: a review of observational studies. *J Med Screen* 2012;19 Suppl 1:14-25.
2. Independent U. K. Panel on Breast Cancer Screening. The benefits and harms of breast cancer screening: an independent review. *Lancet* 2012;380:1778-86.
3. Breen N, Wagener DK, Brown ML, Davis WW, Ballard-Barbash R. Progress in cancer screening over a decade: results of cancer screening from the 1987, 1992, and 1998 National Health Interview Surveys. *J Natl Cancer Inst* 2001;93:1704-13.
4. Carney PA, Miglioretti DL, Yankaskas BC, Kerlikowske K, Rosenberg R, Rutter CM, et al. Individual and combined effects of age, breast density, and hormone replacement therapy use on the accuracy of screening mammography. *Ann Intern Med* 2003;138:168-75.
5. Buist DS, Porter PL, Lehman C, Taplin SH, White E. Factors contributing to mammography failure in women aged 40-49 years. *J Natl Cancer Inst* 2004;96:1432-40.
6. Gilliland FD, Joste N, Stauber PM, Hunt WC, Rosenberg R, Redlich G, et al. Biologic characteristics of interval and screen-detected breast cancers. *J Natl Cancer Inst* 2000;92:743-9.
7. Nelson HD, Zakher B, Cantor A, Fu R, Griffin J, O'Meara ES, et al. Risk factors for breast cancer for women aged 40 to 49 years: a systematic review and meta-analysis. *Ann Intern Med* 2012;156:635-48.
8. Sinclair N, Littenberg B, Geller B, Muss H. Accuracy of screening mammography in older women. *AJR Am J Roentgenol* 2011;197:1268-73.
9. Breast Screening Frequency Trial Group. The frequency of breast cancer screening: results from the UKCCCR Randomised Trial. United Kingdom Co-ordinating Committee on Cancer Research. *Eur J Cancer* 2002;38:1458-64.
10. Duffy SW, Blamey R. Long-term mortality results from the UK Screening Frequency Trial. *6th European Breast Cancer Conference*. Berlin, Germany, 2008.
11. Tosteson AN, Stout NK, Fryback DG, Acharyya S, Herman BA, Hannah LG, et al. Cost-effectiveness of digital mammography breast cancer screening. *Ann Intern Med* 2008;148:1-10.
12. Berry DA, Cronin KA, Plevritis SK, Fryback DG, Clarke L, Zelen M, et al. Effect of screening and adjuvant therapy on mortality from breast cancer. *N Engl J Med* 2005;353:1784-92.
13. Tan SY, van Oortmarssen GJ, de Koning HJ, Boer R, Habbema JD. The MISCAN-Fadia continuous tumor growth model for breast cancer. *J Natl Cancer Inst Monogr* 2006:56-65.
14. Gold MR, Siegel RE, Russell LB, Weinstein MC. Cost-effectiveness in health and medicine. New York: Oxford University Press, 1996.
15. Betsill WL, Jr., Rosen PP, Lieberman PH, Robbins GF. Intraductal carcinoma. Long-term follow-up after treatment by biopsy alone. *JAMA* 1978;239:1863-7.
16. Eusebi V, Feudale E, Foschini MP, Micheli A, Conti A, Riva C, et al. Long-term follow-up of in situ carcinoma of the breast. *Semin Diagn Pathol* 1994;11:223-35.
17. Feig SA. Ductal carcinoma in situ. Implications for screening mammography. *Radiol Clin North Am* 2000;38:653-68, vii.
18. Welch HG, Black WC. Using autopsy series to estimate the disease "reservoir" for ductal carcinoma in situ of the breast: how much more breast cancer can we find? *Ann Intern Med* 1997;127:1023-8.
19. Jorgensen KJ, Keen JD, Gotzsche PC. Is mammographic screening justifiable considering its substantial overdiagnosis rate and minor effect on mortality? *Radiology* 2011;260:621-7.
20. Gotzsche PC, Nielsen M. Screening for breast cancer with mammography. *Cochrane Database Syst Rev* 2011:CD001877.
21. Kopans DB, Smith RA, Duffy SW. Mammographic screening and "overdiagnosis". *Radiology* 2011;260:616-20.
22. Autier P, Boniol M, Gavin A, Vatten LJ. Breast cancer mortality in neighbouring European countries with different levels of screening but similar access to treatment: trend analysis of WHO mortality database. *BMJ* 2011;343:d4411.
23. Jorgensen KJ, Zahl PH, Gotzsche PC. Breast cancer mortality in organised mammography screening in Denmark: comparative study. *BMJ* 2010;340:c1241.
24. Moss SM, Nystrom L, Jonsson H, Paci E, Lynge E, Njor S, et al. The impact of mammographic screening on breast cancer mortality in Europe: a review of trend studies. *J Med Screen* 2012;19 Suppl 1:26-32.
25. Walter SD. Mammographic screening: case-control studies. *Ann Oncol* 2003;14:1190-2.

26. Connor RJ, Boer R, Prorok PC, Weed DL. Investigation of design and bias issues in case-control studies of cancer screening using microsimulation. *Am J Epidemiol* 2000;151:991-8.
27. de Koning HJ. Mammographic screening: evidence from randomised controlled trials. *Ann Oncol* 2003;14:1185-9.
28. Humphrey LL, Helfand M, Chan BK, Woolf SH. Breast cancer screening: a summary of the evidence for the U.S. Preventive Services Task Force. *Ann Intern Med* 2002;137:347-60.
29. Hanley JA. Measuring mortality reductions in cancer screening trials. *Epidemiol Rev* 2011;33:36-45.
30. Hanley JA. Analysis of mortality data from cancer screening studies: looking in the right window. *Epidemiology* 2005;16:786-90.
31. Biesheuvel C, Barratt A, Howard K, Houssami N, Irwig L. Effects of study methods and biases on estimates of invasive breast cancer overdetection with mammography screening: a systematic review. *Lancet Oncol* 2007;8:1129-38.
32. Duffy SW, Tabar L, Olsen AH, Vitak B, Allgood PC, Chen TH, et al. Absolute numbers of lives saved and overdiagnosis in breast cancer screening, from a randomized trial and from the Breast Screening Programme in England. *J Med Screen* 2010;17:25-30.
33. Jorgensen KJ, Gotzsche PC. Overdiagnosis in publicly organised mammography screening programmes: systematic review of incidence trends. *BMJ* 2009;339:b2587.
34. de Gelder R, Heijnsdijk EA, van Ravesteyn NT, Fracheboud J, Draisma G, de Koning HJ. Interpreting overdiagnosis estimates in population-based mammography screening. *Epidemiol Rev* 2011;33:111-21.
35. Puliti D, Duffy SW, Miccinesi G, de Koning H, Lynge E, Zappa M, et al. Overdiagnosis in mammographic screening for breast cancer in Europe: a literature review. *J Med Screen* 2012;19 Suppl 1:42-56.
36. Bleyer A, Welch HG. Effect of three decades of screening mammography on breast-cancer incidence. *N Engl J Med* 2012;367:1998-2005.
37. Holford TR, Cronin KA, Mariotto AB, Feuer EJ. Changing patterns in breast cancer incidence trends. *J Natl Cancer Inst Monogr* 2006:19-25.
38. Perry N, Broeders M, de Wolf C, Tornberg S, Holland R, von Karsa L. European guidelines for quality assurance in breast cancer screening and diagnosis. Fourth edition--summary document. *Ann Oncol* 2008;19:614-22.
39. Hofvind S, Geller BM, Skelly J, Vacek PM. Sensitivity and specificity of mammographic screening as practised in Vermont and Norway. *Br J Radiol* 2012;85:e1226-32.
40. Elmore JG, Taplin SH, Barlow WE, Cutter GR, D'Orsi CJ, Hendrick RE, et al. Does litigation influence medical practice? The influence of community radiologists' medical malpractice perceptions and experience on screening mammography. *Radiology* 2005;236:37-46.
41. van Breest Smallenburg V, Setz-Pels W, Groenewoud JH, Voogd AC, Jansen FH, Louwman MW, et al. Malpractice claims following screening mammography in The Netherlands. *Int J Cancer* 2012;131:1360-6.
42. Woloshin S, Schwartz LM, Black WC, Kramer BS. Cancer screening campaigns--getting past uninformative persuasion. *N Engl J Med* 2012;367:1677-9.
43. Schwartz LM, Woloshin S, Fowler FJ, Jr., Welch HG. Enthusiasm for cancer screening in the United States. *JAMA* 2004;291:71-8.
44. U. S. Preventive Services Task Force. Screening for breast cancer: U.S. Preventive Services Task Force recommendation statement. *Ann Intern Med* 2009;151:716-26, W-236.
45. Skaane P, Bandos AI, Gullien R, Eben EB, Ekseth U, Haakenaasen U, et al. Comparison of Digital Mammography Alone and Digital Mammography Plus Tomosynthesis in a Population-based Screening Program. *Radiology* 2013 Jan 7. [Epub ahead of print].
46. Rafferty EA, Park JM, Philpotts LE, Poplack SP, Sumkin JH, Halpern EF, et al. Assessing radiologist performance using combined digital mammography and breast tomosynthesis compared with digital mammography alone: results of a multicenter, multireader trial. *Radiology* 2013;266:104-13.
47. Welch HG. Screening mammography--a long run for a short slide? *N Engl J Med* 2010;363:1276-8.
48. Matthews TJ, Hamilton BE. Delayed childbearing: more women are having their first child later in life. *NCHS Data Brief* 2009;21:1-8.
49. Sieri S, Krogh V, Saieva C, Grobbee DE, Bergmann M, Rohrmann S, et al. Alcohol consumption patterns, diet and body weight in 10 European countries. *Eur J Clin Nutr* 2009;63 Suppl 4:S81-100.
50. Brownson RC, Boehmer TK, Luke DA. Declining rates of physical activity in the United States: what are the contributors? *Annu Rev Public Health* 2005;26:421-43.
51. Stewart ST, Cutler DM, Rosen AB. Forecasting the effects of obesity and smoking on U.S. life expectancy. *N Engl J Med* 2009;361:2252-60.

52. Chiu SY, Duffy S, Yen AM, Tabar L, Smith RA, Chen HH. Effect of baseline breast density on breast cancer incidence, stage, mortality, and screening parameters: 25-year follow-up of a Swedish mammographic screening. *Cancer Epidemiol Biomarkers Prev* 2010;19:1219-28.
53. Arora N, King TA, Jacks LM, Stempel MM, Patil S, Morris E, et al. Impact of breast density on the presenting features of malignancy. *Ann Surg Oncol* 2010;17 Suppl 3:211-8.
54. Yaghjyan L, Colditz GA, Collins LC, Schnitt SJ, Rosner B, Vachon C, et al. Mammographic breast density and subsequent risk of breast cancer in postmenopausal women according to tumor characteristics. *J Natl Cancer Inst* 2011;103:1179-89.
55. Sotiriou C, Pusztai L. Gene-expression signatures in breast cancer. *N Engl J Med* 2009;360:790-800.
56. Perou CM, Sorlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, et al. Molecular portraits of human breast tumours. *Nature* 2000;406:747-52.
57. Gail MH, Brinton LA, Byar DP, Corle DK, Green SB, Schairer C, et al. Projecting individualized probabilities of developing breast cancer for white females who are being examined annually. *J Natl Cancer Inst* 1989;81:1879-86.
58. Tice JA, Cummings SR, Smith-Bindman R, Ichikawa L, Barlow WE, Kerlikowske K. Using clinical factors and mammographic breast density to estimate breast cancer risk: development and validation of a new predictive model. *Ann Intern Med* 2008;148:337-47.
59. Gail MH. Value of adding single-nucleotide polymorphism genotypes to a breast cancer risk model. *J Natl Cancer Inst* 2009;101:959-63.
60. Easton DF, Pooley KA, Dunning AM, Pharoah PD, Thompson D, Ballinger DG, et al. Genome-wide association study identifies novel breast cancer susceptibility loci. *Nature* 2007;447:1087-93.
61. Pashayan N, Pharoah P. Population-based screening in the era of genomics. *Per Med* 2012;9:451-55.
62. Pharoah PD, Antoniou AC, Easton DF, Ponder BA. Polygenes, risk prediction, and targeted prevention of breast cancer. *N Engl J Med* 2008;358:2796-803.
63. Wegwarth O, Schwartz LM, Woloshin S, Gaissmaier W, Gigerenzer G. Do physicians understand cancer screening statistics? A national survey of primary care physicians in the United States. *Ann Intern Med* 2012;156:340-9.
64. Hersch J, Jansen J, Barratt A, Irwig L, Houssami N, Howard K, et al. Women's views on overdiagnosis in breast cancer screening: a qualitative study. *BMJ* 2013;346:f158.
65. van Agt H, Fracheboud J, van der Steen A, de Koning H. Do women make an informed choice about participating in breast cancer screening? A survey among women invited for a first mammography screening examination. *Patient Educ Couns* 2012;89:353-9.
66. O'Connor AM, Bennett CL, Stacey D, Barry M, Col NF, Eden KB, et al. Decision aids for people facing health treatment or screening decisions. *Cochrane Database Syst Rev* 2009:CD001431.
67. Mathieu E, Barratt A, Davey HM, McGeechan K, Howard K, Houssami N. Informed choice in mammography screening: a randomized trial of a decision aid for 70-year-old women. *Arch Intern Med* 2007;167:2039-46.
68. Rimer BK, Briss PA, Zeller PK, Chan EC, Woolf SH. Informed decision making: what is its role in cancer screening? *Cancer* 2004;101:1214-28.

Summary

SUMMARY

Breast cancer is an important public health problem with an estimated number of 1.38 million breast cancer cases and 458,000 deaths from the disease worldwide per year. In many countries it is the most common cancer in women and it is also the principal cause of death from cancer among women globally. The chance of surviving breast cancer depends on the stage at which it is detected. Early detection of disease might lead to a better prognosis, and, consequently, mortality from the disease might be decreased.

Randomized trials have shown that mammography screening significantly reduces breast cancer mortality. Besides the benefits in terms of lives saved, mammography screening is, however, also associated with harms. Harms include undergoing an uncomfortable and sometimes painful test, experiencing anxiety and undergoing biopsy from false-positive test results, undergoing radiation, and overdiagnosis, i.e., the detection of tumors that would not have been detected in a woman's lifetime in the absence of screening. The balance between the benefits and harms is generally found to be favorable for biennial screening between the ages of 50 and 69 (or 74) years.

In order to extrapolate trial results to different screening ages, screening intervals, and screening test, microsimulation models can be used. Furthermore, models can be used to assess the impact of mammography (and other interventions) on population health and for specific population groups. In this thesis, we used the microsimulation model MISCAN-Fadia, developed within the Cancer Intervention and Surveillance Modeling Network (CISNET), to gain insight into the benefits and harms of mammography screening. The first part of this thesis describes the effects of screening in the population as well as in specific population subgroups and compares these effects to the effects of other interventions. In the second part the benefits and harms of different screening scenarios are evaluated.

In **Chapter 2**, one of the main challenges in measuring the effects of breast cancer screening is described. We found that an adequate follow-up time and adequate time after implementation of a screening program are crucial when evaluating the effects of screening. Measuring effects too early can result in an underestimation of the effect of breast cancer screening on breast cancer mortality. A measured effect of 9.9% in Norway after a short follow-up time is likely to increase to 16.2% if the follow-up time is prolonged by 5 years.

Chapter 3 describes the effects of several interventions aimed at reducing breast cancer mortality. The hypothetical effects of optimal adjuvant therapy use, optimal screening use (defined as annual screening between age 40-54 and biennial between age 55-99 years) and elimination of obesity are estimated for the U.S. population in future years up to 2025. Compared to continuation of current trends, maximal reductions in breast cancer deaths could be achieved by 2025 through optimizing treatment use, followed by increasing screening use and then obesity prevention.

U.S. Black women have higher breast cancer mortality rates than White women despite lower incidence. In **Chapter 4**, we evaluated potential reasons for this disparity. Specifically, we estimated the impact of differences in the natural history of disease, screening use and adjuvant therapy use on the racial disparity in breast cancer mortality in the U.S using two simulation models. We found that the majority of the racial disparity in breast cancer mortality could be attributed to differences in natural history parameters (26-44%), use of adjuvant therapy (11-19%) and uptake of mammography screening (7-8%), leaving 38 to 46% unexplained (range across models).

In **Chapter 5**, we extended this work by evaluating the impact of obesity on breast cancer incidence and mortality rates for U.S. White and Black women and assess its effect on the racial disparity in breast cancer mortality. We found that obesity accounts for 3-4% of breast cancer cases and 3-9% of breast cancer deaths in U.S. Black and White women. Even though Black women have a higher obesity prevalence in most age groups, elimination of obesity will not reduce the disparity in breast cancer mortality, because obesity has a protective effect in pre-menopausal women and obesity prevalence is higher in White women beyond age 74 years.

In the second part, we addressed questions around the optimal screening policy. In **Chapter 6**, the effects of 20 screening strategies with varying starting and stopping ages applied annually or biennially were evaluated in 6 models. We found that screening biennially maintained an average of 81% (range across strategies and models, 67% to 99%) of the benefit of annual screening with almost half the number of false-positive results. Screening biennially from ages 50 to 69 years achieved a median 16.5% (range, 15% to 23%) reduction in breast cancer deaths versus no screening. Initiating biennial screening at age 40 years (vs. 50 years) reduced mortality by an additional 3% (range, 1% to 6%), consumed more resources, and yielded more false-positive results. Biennial screening after age 69 years yielded some additional mortality reduction in all models, but overdiagnosis increased most substantially at older ages. In the subsequent chapters, the benefits and harms of breast cancer screening were evaluated in more detail for different starting ages (Chapter 7), stopping ages (Chapter 8 and Chapter 9), and screening intervals (Chapter 10).

With regard to the starting age of screening, important factors are that younger women have a lower risk of developing disease (within the next 10-years) and that the sensitivity of the screening test is lower, due to more dense breast tissue. In addition, the specificity of the screening test is lower, while the tumor growth rate is higher. In **Chapter 7**, taking into account these factors, we found that women age 40 with a two-fold risk (compared to average) can expect the same balance of benefits and harms as average-risk women age 50 from starting biennial screening. Factors that lead to a two-fold risk are having extremely dense breasts (density category Breast Imaging-Reporting and Data System (BI-RADS) 4) and having a first degree relative with breast cancer. In the U.S. 13% of women in their forties have

extremely dense breast and 9% of women in this age group have a first degree relative with breast cancer.

Older women have a higher risk of developing disease, and the sensitivity of the screening test has also been found to be higher. On the other hand, their remaining life expectancy is shorter, limiting the potential to save a life and leading to more overdiagnosis. An additional complicating factor is that there is limited information about the benefits and harms of screening beyond age 74 years, because none of the randomized controlled trials included women older than 74 years. In **Chapter 8**, we found that despite the limited knowledge, which was reflected in model differences, the age at which harms began to outweigh benefits was the same across models. Screening benefits outweigh harms at age 74, but the balance becomes less favorable with advancing age. At age 90, harms outweigh benefits, largely as a consequence of overdiagnosis. This age was the same across the three models, despite important model differences in assumptions on ductal carcinoma *in situ*.

In **Chapter 9**, we assessed how heterogeneity in comorbidity might shift the balance of harms and benefits towards cessation at younger or older ages. We found that healthy women (with no or mild comorbidity) can expect more benefit from a screen at age 74 years than women with (moderate or severe) comorbidities. Women with comorbidities might therefore choose a younger stopping age than women without comorbidities. The age of screening cessation based on comorbidity levels varied by nearly a 10-year interval around the age cut-point of 74 years included in the current breast cancer screening recommendation.

In **Chapter 10** we modeled a trial comparing a 3-year with a 1-year screening interval. The more frequent a woman has screening exams, the larger the probability of having an earlier diagnosis, and the larger the benefit from screening might be. However, the harms (false positive exams, overdiagnosis) will also increase with more frequent screening. A randomized trial, the UK Breast Screening Frequency Trial, did not show a significant difference in breast cancer mortality between screening every year and screening once every 3 years. To be able to replicate the trial in our microsimulation model, we had to assume a relatively low screening test sensitivity. In addition, we found evidence that the trial was underpowered to find a statistically significant difference between the two groups. Our results indicated that there was benefit of shortening the screening interval, although the benefit is probably not large enough to start annual screening.

In **Chapter 11**, we investigated the costs and effects of screening using film vs. digital mammography and the implications of replacing film by digital mammography for a program for low-income, uninsured women in the U.S. Digital mammography has been found to be more sensitive than film mammography, especially in younger women, but is also more expensive. We found that when performing the same number of screens, digital mammography saved slightly more lives than film mammography, but also increased the number of false positive exams and costs of screening and diagnostics. Therefore, the cost-effectiveness

was found to be more favorable for film than for digital mammography. In addition, when operating with a fixed budget, replacing film by digital mammography would result in fewer women that can be served (25-26%), and fewer lives that are saved (23-24%). Thus, price reduction of digital mammography, budget increases, or changes in the program, such as prolonging the screening interval, are warranted to achieve comparable health effects of film and digital mammography.

In **Chapter 12** (the general discussion of this thesis), the research questions are answered and discussed. In addition, the current status of breast cancer screening is described, focusing on the current controversy around breast cancer screening. Furthermore, implications and directions for future research are provided.

We found that mammography screening has had a substantial impact on breast cancer mortality in the U.S. and is projected to continue to do so in the future. Screening women biennially from age 50 to 74 years leads to a favorable balance between benefits and harms. More intensive screening (either extending the age ranges or increasing the screening frequency) leads to more benefits (lives saved), but also leads to more harms (false-positives and overdiagnosis). Overall, starting earlier (at age 40 vs. 50), screening more often (annual vs. biennial or triennial) and continuing screening after age 74, led to a less favorable balance between benefits and harms than biennial screening between age 50 and 74 years. However, the balance between benefits and harms of extending screening is more favorable for specific subgroups, for example women at increased risk and women without comorbidities. Thus, women at higher risk might choose an earlier starting age and healthy women might continue screening until an older age than the average female population.

Future research should further evaluate the effects of targeted screening based on either 'classical' risk factors or genetic risk profiles versus untargeted screening based only on age and sex. In addition, in order to assess the effects of breast cancer screening by molecular subtype (ER and HER2), it is needed to incorporate information on the risk for the development of specific molecular subtypes and how subtypes influence the natural history of disease. Finally, informing all eligible women about the benefits and harms of screening and supporting them in making an informed decision about screening participation is crucial.

Samenvatting

SAMENVATTING

Borstkanker is een belangrijk probleem voor de volksgezondheid met een geschat aantal van 1,38 miljoen gevallen van borstkanker en 458.000 sterfgevallen door de ziekte per jaar wereldwijd. In veel landen is het de meest voorkomende vorm van kanker bij vrouwen en het is ook de belangrijkste oorzaak van overlijden door kanker bij vrouwen wereldwijd. De kans om borstkanker te overleven is afhankelijk van het stadium waarin het wordt ontdekt. Vroege opsporing van de ziekte kan leiden tot een eerdere behandeling, een betere prognose, en daardoor mogelijk tot vermindering van sterfte aan de ziekte.

Gerandomiseerde studies hebben aangetoond dat borstkankerscreening met mammografie borstkankersterfte kan voorkomen. Naast de voordelen (geredde levens) heeft screening echter ook nadelen. De nadelen bestaan uit het ondergaan van een onprettige en soms pijnlijke test, de angst en het ondergaan van een biopsie ten gevolge van een fout-positieve screeningstest, blootstelling aan (een lage dosis) straling en, ten slotte, overdiagnose. Er is sprake van overdiagnose als een tumor zonder screening niet ontdekt zou zijn gedurende het leven van een vrouw. De balans tussen de voor- en nadelen van borstkankerscreening wordt over het algemeen gunstig bevonden voor eens in de 2 jaar screenen in de leeftijdsgroep van 50 tot 69 (of 74) jaar..

Om de resultaten van gerandomiseerde studies te extrapoleren naar andere situaties (bijvoorbeeld om het effect van screening in andere leeftijdsgroepen of met andere screeningsintervallen te schatten) kunnen microsimulatiemodellen gebruikt worden. Deze modellen kunnen bovendien gebruikt worden om de effecten van mammografie en andere interventies op de borstkankerincidentie en -sterfte te voorspellen, zowel voor de algemene bevolking als voor specifieke bevolkingsgroepen. In dit proefschrift hebben we het microsimulatiemodel MISCAN-Fadia gebruikt om de voor- en nadelen van borstkankerscreening te evalueren. Het model is ontwikkeld binnen het Cancer Intervention and Surveillance Modeling Network (CISNET). Het eerste deel van dit proefschrift beschrijft de effecten van screening in de bevolking en in specifieke bevolkingsgroepen en vergelijkt deze effecten met de effecten van andere interventies. In het tweede deel worden de voor- en nadelen van verschillende screeningsscenario's, onder andere variërend in start- en stopleeftijd en screeningsinterval, geëvalueerd.

In **hoofdstuk 2** wordt één van de grootste uitdagingen bij het meten van de effecten van borstkankerscreening beschreven. Een adequate follow-up tijd is cruciaal bij het evalueren van de effecten van screening. Als de effecten te vroeg gemeten worden, bijvoorbeeld slechts enkele jaren na implementatie van een screeningsprogramma, kan dat leiden tot een onderschatting van het effect van borstkankerscreening op de borstkankersterfte. Naar verwachting zal een gemeten effect (sterftereductie) van 9,9% in Noorwegen na een korte follow-up tijd toenemen tot 16,2% als de follow-up tijd verlengd wordt met 5 jaar.

Hoofdstuk 3 beschrijft de effecten van verschillende interventies gericht op het verminderen van borstkankersterfte in de Verenigde Staten (VS). We hebben de hypothetische effecten van maximaal gebruik van adjuvante therapie (chemotherapie en hormonale therapie), maximale deelname aan screening (gedefinieerd als 100% deelname aan jaarlijkse screening van leeftijd 40 tot 54 jaar en 100% deelname aan screening eens in de 2 jaar van leeftijd 55 tot 99 jaar) en eliminatie van obesitas geschat voor de Amerikaanse vrouwelijke bevolking in de komende jaren tot en met 2025. Vergeleken met wat er zou gebeuren als de huidige trends zouden doorzetten, schatten we dat de grootste daling in borstkankersterfte gerealiseerd kan worden door het optimaliseren van therapiegebruik, gevolgd door het verhogen van screeningsdeelname en vervolgens door de eliminatie van obesitas.

Afro-Amerikaanse vrouwen hebben een lagere borstkankerincidentie, maar hogere sterfte aan borstkanker dan blanke Amerikaanse vrouwen. In **hoofdstuk 4** hebben we mogelijke verklaringen voor de raciale ongelijkheid in borstkankersterfte onderzocht. We hebben twee modellen gebruikt om het effect te schatten van verschillen in natuurlijk beloop van de ziekte, verschillen in screeningsdeelname en verschillen in gebruik van adjuvante therapie. Het grootste deel van de raciale ongelijkheid in borstkankersterfte kan worden toegeschreven aan verschillen in het natuurlijk beloop van de ziekte (26-44%, variërend tussen de twee modellen), verschillen in gebruik van adjuvante therapie (11-19%) en verschillen in screeningsdeelname (7-8%), waardoor 38 tot 46% van de raciale ongelijkheid in borstkankersterfte onverklaard bleef.

In **hoofdstuk 5** hebben we dezelfde twee modellen gebruikt om de impact van obesitas op de borstkankerincidentie en -sterfte voor zowel Afro-Amerikaanse als blanke Amerikaanse vrouwen te evalueren. Bovendien hebben we het effect van verschillen in de prevalentie van obesitas (percentage vrouwen met een BMI>30) op de raciale ongelijkheid in borstkankersterfte onderzocht. Zo'n 3-4% van de gevallen van borstkanker en 3-9% van alle borstkankersterfgevallen in beide groepen kan worden toegeschreven aan obesitas. Hoewel Afro-Amerikaanse vrouwen een hogere prevalentie van obesitas hebben in de meeste leeftijdsgroepen, zal de raciale ongelijkheid in borstkankersterfte niet verminderen door de eliminatie van obesitas, omdat obesitas een beschermend effect heeft bij premenopauzale vrouwen en omdat de prevalentie van obesitas bij vrouwen ouder dan 74 jaar hoger is bij blanke dan bij Afro-Amerikaanse vrouwen.

In het tweede deel worden de effecten van verschillende screeningstrategieën geëvalueerd. In **hoofdstuk 6** worden de effecten van 20 screeningstrategieën met verschillende start- en stopleeftijden en verschillende screeningsintervallen (1 en 2 jaar) geëvalueerd in 6 modellen. Screenen met een 2-jaars interval handhaaft zo'n 81% (variërend tussen strategieën en modellen van 67% tot 99%) van het voordeel (geredde levens) van een 1-jaars screeninginterval, terwijl het aantal fout-positieven bijna gehalveerd wordt. Eens in de twee jaar screenen van leeftijd 50 tot 69 jaar leidt tot een 16,5% (15-23%) lagere borstkankersterfte ten opzichte

van een situatie zonder screening. Eerder beginnen met screenen (op leeftijd 40 jaar) leidt tot een extra 3% (1-6%) sterftereductie, maar leidt ook tot meer fout-positieven. Doorgaan met screenen na leeftijd 69 jaar leidt ook tot een extra sterftereductie in alle modellen, maar de hoeveelheid overdiagnose neemt aanzienlijk toe bij screenen op hogere leeftijd. In de volgende hoofdstukken zijn de voor- en nadelen van borstkankerscreening in meer detail bekeken voor verschillende startleeftijden (hoofdstuk 7), stopleeftijden (hoofdstuk 8 en hoofdstuk 9), en screeningsintervallen (hoofdstuk 10).

Met betrekking tot de startleeftijd van screening zijn belangrijke factoren dat jongere vrouwen een lager risico hebben om borstkanker te krijgen en dat de gevoeligheid (sensitiviteit) van de test lager (vaker fout-positief) is vanwege dichter borstweefsel. Bovendien is de specificiteit van de test lager, terwijl de groeisnelheid van tumoren op jonge leeftijd hoger is. In **hoofdstuk 7** hebben we rekening gehouden met al deze factoren en vonden we dat vrouwen van leeftijd 40 jaar met een twee keer zo hoog risico als gemiddeld, dezelfde balans van voor- en nadelen kunnen verwachten van eens in de 2 jaar screenen, als vrouwen met een gemiddeld risico die op leeftijd 50 jaar beginnen met eens in de 2 jaar screenen. Factoren die leiden tot een twee keer zo hoog risico zijn het hebben van heel dicht borstweefsel (Breast Imaging-Reporting and Data System (BI-RADS) categorie 4) en het hebben van een eerstegraads familielid met borstkanker. In de VS heeft 13% van de vrouwen in de veertig heel dicht borstweefsel en 9% van de vrouwen in deze leeftijdsgroep heeft een eerstegraads familielid met borstkanker.

Bij de evaluatie van de bovenste leeftijdsgrens speelt een rol dat oudere vrouwen een hoger risico op borstkanker hebben dan jongere vrouwen en dat de sensitiviteit van de test hoger is op hogere leeftijd, wat zou kunnen beteken dat de balans van screenen voor oudere vrouwen gunstiger is dan voor jongere vrouwen. Echter, de resterende levensverwachting van oudere vrouwen is korter, wat de mogelijkheid om een leven te redden beperkt en bovendien tot meer overdiagnose leidt. Een extra complicerende factor is dat er slechts beperkte informatie over de voor- en nadelen van screening na leeftijd 74 jaar beschikbaar is, omdat geen van de gerandomiseerde studies vrouwen ouder dan 74 jaar heeft geïncludeerd. In **hoofdstuk 8** zagen we dat de beperkte kennis tot uiting kwam in verschillen tussen de drie modellen die gebruikt werden. Ondanks de verschillen tussen de modellen was de leeftijd waarop de nadelen begonnen op te wegen tegen de voordelen hetzelfde in de drie modellen. Op leeftijd 74 jaar waren de voordelen groter dan de nadelen, maar de balans werd minder gunstig met toenemende leeftijd. Vanaf leeftijd 90 hadden de nadelen de overhand, met name als gevolg van toenemende overdiagnose. Deze leeftijd was hetzelfde in de drie modellen, ondanks belangrijke verschillen tussen de modellen in aannames over het natuurlijk beloop van de ziekte (bijvoorbeeld over de progressie van ductaal carcinoma *in situ*).

In **hoofdstuk 9** hebben we onderzocht hoe verschillen in gezondheid op oudere leeftijd (het al dan niet hebben van comorbiditeit) de balans tussen de voor- en nadelen van screen-

ing beïnvloedt. De stopleeftijd van screening op basis van comorbiditeit, gedefinieerd als de leeftijd waarop de balans tussen de voor- en nadelen net zo gunstig was als voor screenen op leeftijd 74 voor de algemene bevolking, varieerde met bijna 10-jaar rond de leeftijd van 74 jaar (de leeftijd die op dit moment aanbevolen wordt als bovengrens in de huidige borstkankerscreeningsrichtlijnen). Gezonde vrouwen (zonder comorbiditeit) kunnen meer voordeel verwachten van een screen op leeftijd 74 jaar dan vrouwen met (matige of ernstige) comorbiditeit en zouden daarom voor een hogere stopleeftijd kunnen kiezen dan vrouwen met comorbiditeit, die er juist weer voor zouden kunnen kiezen eerder te stoppen met screenen.

In **hoofdstuk 10** hebben we een gerandomiseerde studie nagebootst waarin een 3-jaars screeningsinterval vergeleken wordt met een 1-jaars interval. Des te vaker een vrouw screeningstesten ondergaat, des te groter de kans is op een vroegere diagnose, en des te groter het voordeel van screening kan zijn. De nadelen (fout-positieve tests, overdiagnose) zullen echter ook toenemen naarmate vaker gescreend wordt. Een gerandomiseerde studie, de UK Breast Screening Frequency Trial, vond geen significant verschil in borstkankersterfte tussen elk jaar screenen en eens in de 3 jaar screenen. Om de resultaten van deze studie na te bootsen met ons model moesten we aannemen dat de sensitiviteit van de screeningstest in de studie relatief laag was. Bovendien vonden we aanwijzingen dat de studie een laag onderscheidend vermogen ('power') had om een statistisch significant verschil tussen de twee groepen te vinden. Onze resultaten laten zien dat er voordeel is van een verkorting van het screeningsinterval van 3 naar 1 jaar, maar dat het voordeel waarschijnlijk niet groot genoeg is om jaarlijkse screening in te voeren.

In **hoofdstuk 11** hebben we de kosten en effecten van screening met behulp van analoge en digitale mammografie onderzocht en bekeken wat de gevolgen zouden zijn van vervanging van analoge mammografie door digitale mammografie voor een programma voor onverzekerde vrouwen met een laag inkomen in de VS. Digitale mammografie heeft een hogere sensitiviteit dan analoge mammografie, vooral bij jongere vrouwen, maar is ook duurder. Als hetzelfde aantal testen wordt uitgevoerd, kan digitale mammografie iets meer levens redden dan analoge mammografie, maar tegelijkertijd nemen dan het aantal fout-positieve tests en de kosten van screening en diagnostiek ook toe. De kosteneffectiviteit van digitale mammografie was daardoor ongunstiger dan de kosteneffectiviteit van analoge mammografie. Bovendien zou het vervangen van analoge door digitale mammografie in een programma met een vast budget betekenen dat er minder vrouwen kunnen worden bereikt (25-26%), en minder levens worden gered (23-24%). Om vergelijkbare effecten op de gezondheid te bereiken met digitale als met analoge mammografie is het daarom noodzakelijk dat de prijs van digitale mammografie verlaagd wordt, het programmabudget verhoogd wordt, of dat er veranderingen in het programma gemaakt worden, zoals het verlengen van het screeningsinterval.

In **hoofdstuk 12** (de algemene discussie van dit proefschrift) worden de onderzoeksvragen beantwoord en besproken. Bovendien wordt de huidige status van borstkankerscreening beschreven, inclusief een beschrijving van de huidige controverse rondom de effecten van borstkankerscreening. Verder worden implicaties van het onderzoek en richtingen voor toekomstig onderzoek gegeven.

Eén van onze bevindingen is dat mammografie een aanzienlijke impact op de borstkankersterfte in de VS heeft gehad en naar verwachting blijft hebben in de toekomst. Het eens in de 2 jaar screenen van vrouwen vanaf leeftijd 50 tot 74 jaar leidt tot een gunstige balans tussen de voor- en nadelen van screening. Meer intensieve screening (het verruimen van de leeftijdsgrenzen of het verhogen van de screeningsfrequentie) leidt tot meer voordelen (levens gered), maar leidt ook tot meer nadelen (fout-positieven en overdiagnose). Daardoor leidt eerder beginnen met screenen (bijvoorbeeld op leeftijd 40 versus 50 jaar), vaker screenen (bijvoorbeeld elk jaar versus eens in de 2 of 3 jaar) en doorgaan met screening na de leeftijd van 74 jaar, tot een minder gunstige balans tussen de voor- en nadelen dan eens in de 2 jaar screenen van 50 tot 74 jaar. De balans tussen de voor- en nadelen van uitbreiding van screening is echter gunstiger voor specifieke subgroepen, zoals vrouwen met een verhoogd risico en vrouwen zonder comorbiditeit. Zo zouden vrouwen met een verhoogd risico kunnen kiezen voor een eerdere startleeftijd en zouden gezonde vrouwen ervoor kunnen kiezen langer door te gaan met screening dan de gemiddelde vrouwelijke bevolking.

Toekomstig onderzoek zou zich moeten richten op de effecten van gepersonaliseerde screening op basis van hetzij 'klassieke' risicofactoren, hetzij genetische risicoprofielen versus de effecten van de huidige algemene screening, uitsluitend gestratificeerd op basis van leeftijd en geslacht. Bovendien is het van belang om informatie over het risico voor de ontwikkeling van specifieke moleculaire subtypen (ER, HER2) en hoe deze subtypes het natuurlijk beloop van de ziekte beïnvloeden in modellen mee te nemen om te kunnen inschatten of de effecten van borstkankerscreening variëren naar subtype. Ten slotte is het cruciaal om alle vrouwen die in aanmerking komen voor borstkankerscreening te informeren over de voor- en nadelen van screening en hen te ondersteunen bij het maken van een weloverwogen beslissing (een geïnformeerde keuze) over deelname aan borstkankerscreening.

Dankwoord & acknowledgments
Curriculum Vitae
Publications
PhD portfolio

DANKWOORD

Het is eigenlijk een beetje raar om pas aan het eind van een promotietraject je dank te uiten en ik hoop dan ook dat ik ook tijdens het traject zo nu en dan mijn dankbaarheid heb getoond. Ik heb de afgelopen jaren met ontzettend veel plezier aan mijn proefschrift gewerkt en dat is voor een groot deel te danken aan onderstaande personen.

Als eerste mijn promotor prof. dr. Harry de Koning. Harry, vanaf het begin van onze samenwerking gaf je me het idee dat je vertrouwen in me had. Ik waardeer het zeer dat je me veel vrijheid en verantwoordelijkheid hebt gegeven, waardoor ik met heel veel plezier heb gewerkt en daarnaast enorm veel heb geleerd. Je wist altijd haarfijn de vinger te leggen op de sterke en zwakke punten van artikelen en reageerde altijd razendsnel als iets belangrijk was. Hartelijk dank voor je optimisme, de kansen die je me hebt geboden en het prettige contact. Ik kijk er erg naar uit om onze samenwerking voort te zetten in de toekomst!

Eveline, ik bewonder je efficiënte manier van werken. Bedankt dat je deur, letterlijk en figuurlijk, altijd voor me open stond. Ik vond het fijn dat ik altijd bij je terecht kon om dingen te bespreken en wil je hartelijk danken voor je input en ideeën in alle artikelen en voorstellen waar we samen aan gewerkt hebben. Gerrit, je geduldige manier van uitleggen van MISCAN-Fadia heeft me enorm op weg geholpen de eerste tijd. Bovendien vond ik het daarna ook altijd erg gezellig als je nog eens langs kwam op MGZ.

Many articles in this thesis are a result of the collaboration within CISNET. I would like to thank all breast CISNET-ers for the fruitful and pleasant collaboration. There are a few people who deserve a special word of thanks. First of all, Jeanne, I admire your enormous amount of energy. Thank you very much for all the hard work and time you committed to the articles in this thesis and to CISNET. You are a pleasure to work with, because of your instantaneous replies to emails, very detailed feedback on articles, and your energetic input at meetings, even while knitting... Clyde, it is a shame that we did not have any webcams during our numerous conference calls, because then you would have seen me nodding my head in agreement very often when you spoke. Your intelligent remarks and insightful comments have greatly improved the papers we worked on together. Aimee, thank you for your continued support in organizing conference calls, structuring projects, and summarizing results; you are an important contributor to the efficiency of the breast CISNET group. All other CISNET-ers, including Amy, Tasha, Brian, Sandra, An-Tsun, Diego, Gary, Sylvia, Harald, Oguz, Kathy, Yao, Hui, Don, Marvin, thank you all for working together, for many interesting discussions on breast cancer screening during meetings and dinners, and for taking me out to a baseball game.

I would like to thank the BCSC (Diana M, Diana, B, Karla, among others) not only for the high-quality data provided, but also for the high-quality discussions during meetings, at which you showed me the power of asking (the right) questions. It has also been a pleasure to work with the EPC (Heidi Nelson) and CDC (Donatus Ekwueme).

De leden van de kleine commissie, prof. dr. Hunink, prof. dr. den Heeten en prof. dr. Rutten-van Mölken, wil ik bedanken voor het beoordelen van het manuscript. Ook de overige commissieleden wil ik hartelijk bedanken voor hun interesse in dit onderzoek en voor hun bereidheid hierover met mij van gedachten te wisselen.

Van alle andere collega's verdienen de medebewoners van Ae-134 misschien nog wel de meeste dank. Noortje, Ilse, Jan en Rienke, jarenlang hebben we (gelukkig niet allemaal tegelijkertijd) als kamergenoten (veel) lief en (een beetje) leed gedeeld. Bedankt voor jullie praatjes, grapjes, praktische adviezen en het halen van vele kopjes koffie en thee. Jan, je bent altijd zeer betrokken en geïnteresseerd en weet op een mooie manier bescheidenheid aan zelfverzekerdheid te koppelen. Ik vind het daarom een eer dat je mijn paranimf wilt zijn.

Binnen de screensectie van MGZ heb ik erg veel gehad aan praktische MISCAN-tips en aangenaam gezelschap tijdens CISNET trips. Dank daarvoor aan Elisabeth, Rianne, Paula, Iris, Kevin, Tiago, Joost, Sonja, Janneke en Frank. Lisanne, ik vond het leuk dat je bij ons stage kwam lopen. Bedankt voor het onderzoek dat je hebt gedaan; het uitzoeken van alle kostendata en uitvoeren van vele modelruns. Iedereen van de helpdesk en het secretariaat, hartelijk dank voor de ondersteuning. En in het bijzonder Arry, bedankt dat ik altijd bij je kon binnenlopen voor een blik in Harry's agenda, praktische vragen en hulp bij allerhande zaken.

Dan is er nog een aantal (ex) MGZ collega's die het werken op de afdeling zo leuk maken en daarbuiten ook voor vertier zorgen. Ilke, Lenneke, Lidy, Mirjam 2x, Mariëlle, Marieke, Jitske, Helen, Hester, Else-Mariette, Sandra, Rogier, Moniek, Natasja, Lifang, Luuk, Luc, Karen, Jacques, Suzie, Inge, Eefje, Robine, Carlijn, Maggie, Suzan, Rick bedankt voor alle praatjes op de gang, lunchwandelingen, etentjes en spelletjesavonden, filmfestivalbezoekjes, hardlooptrainingen, het (beach)volleyballen, schaatsen, en natuurlijk koffiedrinken bij D.E.

Vrienden, familie en schoonfamilie, ik heb genoten van alle etentjes, borrels, weekendjes weg en vakanties. Mijn hockeyteam, shoothappens, bedankt dat jullie me elke woensdagavond en zondag lekker heen en weer lieten rennen op het hockeyveld. Mede-KZVS-ers, het cat-zeilen is een geweldige manier om er even tussen uit te zijn en weekenden op het water door te brengen. Ik ben blij dat jullie er zijn om mee (en tegen) te zeilen en kijk al weer met smart uit naar het nieuwe zeilseizoen. Lieve Leidenaren, ik hoop nog vaak derde kerstdag met jullie te vieren. Maartje, Joost en Hidde, ik vind het altijd weer een feest om bij jullie langs te komen. Maart, bedankt dat je als paranimf naast me wilt staan. Lieve ouders, bedankt voor jullie belangstelling, vertrouwen, onvoorwaardelijke steun en liefde. Het is mooi om te zien dat jullie zo trots op me zijn. Otto, ik ben heel erg blij dat je al zo lang bij me bent. Bedankt voor je steun, eerlijkheid en kritische blik, (soms flauwe) humor, en alle lekkere maaltijden die je voor me hebt bereid de afgelopen jaren... Ik hoop dat we de komende jaren samen nog veel mooie momenten gaan beleven!

CURRICULUM VITAE

Nicolien Thea van Ravesteijn was born on the 12th of June 1981, in Leiden, the Netherlands. In 1999, she completed her secondary education at the Stedelijk Gymnasium Leiden. One year later, in 2000, she started studying 'Human Movement Sciences' at the Vrije Universiteit (VU) in Amsterdam and obtained her Master of Science degree in 2006 with a specialization in 'Movement Coordination'. She wrote a Master's thesis about the influence of cognitive factors on rhythmic bimanual coordination and during her research internship she assessed the effect of a visual illusion on the performance of a ball hitting task. From 2007 to 2008 she worked as a junior researcher at the department of Rehabilitation Medicine, VU University Medical Center (VUmc) in Amsterdam on the evaluation of a questionnaire developed to measure mobility limitations in children with cerebral palsy, the MobQues, assessing its validity and reliability.

Since 2008, she is employed as a researcher at the department of Public Health at the Erasmus University Medical Center in Rotterdam, where she evaluates the effects of breast cancer screening using the microsimulation model MISCAN-Fadia. An important part of the research is performed within the Cancer Intervention and Surveillance Modeling Network (CISNET), which is an international consortium of NCI-sponsored investigators that use statistical modeling to improve our understanding of cancer control interventions in prevention, screening, and treatment and their effects on population trends in incidence and mortality. The research findings of the impact of breast cancer screening on population health are presented in this thesis.

PUBLICATIONS

Mandelblatt JS, **van Ravesteyn NT**, Schechter CB, Chang Y, Huang A, Near AM, de Koning HJ, Jemal A. Which strategies would reduce breast cancer mortality most? Collaborative modeling of optimal screening, treatment, and obesity prevention. *Cancer [in press]*.

Mandelblatt JS, Tosteson AN**, van Ravesteyn NT**. Costs, evidence, and value in the medicare program: comment on "the cost of breast cancer screening in the medicare population". *JAMA Intern Med*. 2013 Feb 11;173(3):227-8.

Chang Y, Schechter CB, **van Ravesteyn NT**, Near AM, Heijnsdijk EA, Adams-Campbell L, Levy D, de Koning HJ, Mandelblatt JS. Collaborative modeling of the impact of obesity on race-specific breast cancer incidence and mortality. *Breast Cancer Res Treat*. 2012 Dec;136(3):823-35.

Mackenbach JP, Lingsma HF, **van Ravesteyn NT**, Kamphuis CB. The population and high-risk approaches to prevention: quantitative estimates of their contribution to population health in the Netherlands, 1970-2010. *Eur J Public Health*. 2012 Aug 8. [Epub ahead of print]

Nelson HD, Zakher B, Cantor A, Fu R, Griffin J, O'Meara ES, Buist DS, Kerlikowske K, **van Ravesteyn NT**, Trentham-Dietz A, Mandelblatt JS, Miglioretti DL. Risk factors for breast cancer for women aged 40 to 49 years: a systematic review and meta-analysis. *Ann Intern Med*. 2012 May 1;156(9):635-48.

van Ravesteyn NT, Miglioretti DL, Stout NK, Lee SJ, Schechter CB, Buist DS, Huang H, Heijnsdijk EA, Trentham-Dietz A, Alagoz O, Near AM, Kerlikowske K, Nelson HD, Mandelblatt JS, de Koning HJ. Tipping the balance of benefits and harms to favor screening mammography starting at age 40 years: a comparative modeling study of risk. *Ann Intern Med*. 2012 May 1;156(9):609-17.

Mandelblatt JS, Cronin KA, Berry DA, Chang Y, de Koning HJ, Lee SJ, Plevritis SK, Schechter CB, Stout NK, **van Ravesteyn NT**, Zelen M, Feuer EJ. Modeling the impact of population screening on breast cancer mortality in the United States. *Breast*. 2011 Oct;20 Suppl 3:S75-81.

van Ravesteyn NT, Heijnsdijk EA, Draisma G, de Koning HJ. Prediction of higher mortality reduction for the UK Breast Screening Frequency Trial: a model-based approach on screening intervals. *Br J Cancer*. 2011 Sep 27;105(7):1082-8.

de Gelder R, Heijnsdijk EA, **van Ravesteyn NT**, Fracheboud J, Draisma G, de Koning HJ. Interpreting overdiagnosis estimates in population-based mammography screening. *Epidemiol Rev*. 2011 Jul;33(1):111-21.

van Ravesteyn NT, Heijnsdijk EA, de Koning HJ. More on screening mammography. *N Engl J Med*. 2011 Jan 20;364(3):282-3; author reply 285-6.

van Ravesteyn NT, Schechter CB, Near AM, Heijnsdijk EA, Stoto MA, Draisma G, de Koning HJ, Mandelblatt JS. Race-specific impact of natural history, mammography screening, and adjuvant treatment on breast cancer mortality rates in the United States. *Cancer Epidemiol Biomarkers Prev*. 2011 Jan;20(1):112-22.

Mandelblatt JS, Cronin KA, Bailey S, Berry DA, de Koning HJ, Draisma G, Huang H, Lee SJ, Munsell M, Plevritis SK, Ravdin P, Schechter CB, Sigal B, Stoto MA, Stout NK, **van Ravesteyn NT**, Venier J, Zelen M, Feuer EJ; Breast Cancer Working Group of the Cancer Intervention and Surveillance Modeling Network. Effects of mammography screening under different screening schedules: model estimates of potential benefits and harms. *Ann Intern Med*. 2009 Nov 17;151(10):738-47.

van Ravesteyn NT, Scholtes VA, Becher JG, Roorda LD, Verschuren O, Dallmeijer AJ. Measuring mobility limitations in children with cerebral palsy: content and construct validity of a mobility questionnaire (MobQues). *Dev Med Child Neurol*. 2010 Oct;52(10):e229-35

van Ravesteyn NT, Dallmeijer AJ, Scholtes VA, Roorda LD, Becher JG. Measuring mobility limitations in children with cerebral palsy: interrater and intrarater reliability of a mobility questionnaire (MobQues). *Dev Med Child Neurol*. 2010 Feb;52(2):194-9.

PHD PORTFOLIO

Summary of PhD training and teaching

Name PhD student: N.T. van Ravesteyn
PhD period: 2008-2013
Erasmus MC Department: Public Health
Promotor(s): Prof.dr. H.J. de Koning

1. PhD training

	Year	Workload (ECTS)
General courses		
- Biomedical English Writing and Communication	2010-2011	4
Specific courses		
Erasmus Summer/Winter Programme, Erasmus MC, Rotterdam		
- Case-control Studies	2008	0.7
- Methods of Public Health Research	2008	0.7
- Survival analysis for clinicians	2009	1.4
- Regression Analysis	2009	1.9
- Demography of Aging	2009	0.7
- Health Economics	2009	0.7
- Genome Wide Association Analysis	2012	1.4
Nihes, Erasmus MC, Rotterdam		
- Planning and evaluation of screening	2008	1.4
Department of Biostatistics, Erasmus MC, Rotterdam		
- Multistate-models and models for competing risks	2009	0.6
Karolinska Institute, Stockholm, Sweden		
- Essentials of descriptive cancer epidemiology	2010	1.2
Centrum voor nascholing Amsterdam (Karin Herrebout)		
- Interactief presenteren	2011	0.3
Nederlands Kanker Instituut - Antoni van Leeuwenhoek		
- Absolute Risk Prediction workshop	2012	0.2
Seminars and workshops		
- Seminars at the department of Public Health, Erasmus MC, Rotterdam	2008-2013	3.6
- "Research-based evaluation of the National Mammography Screening Programme", Oslo, Norway	2008	0.3
- Cancer screening symposium: trials and modeling to guide public health policies, Erasmus MC, Rotterdam	2009	0.3
- BCSC/CISNET/EPC Collaboration Meeting, Portland , USA	2009	0.5
- Symposium Patients, People and Populations: 40 years of Epidemiology at Erasmus, Erasmus MC, Rotterdam	2009	0.2

- APRC meeting, Seattle, USA	2010	0.3
- BCSC-GO meeting, Sonoma, USA	2011	0.3
- Workshop on methods for evaluation of screening, London, UK	2012	0.4
- BCSC-PO1 meeting "Risk Based Breast Cancer Screening in Community Settings", Seattle, USA	2012	0.3
- werkconferentie best practices, code goed gebruik, Utrecht	2012	0.2
- Symposium over 10-jaar borstkankerscreening in Vlaanderen, Hasselt, België	2012	0.1

Presentations

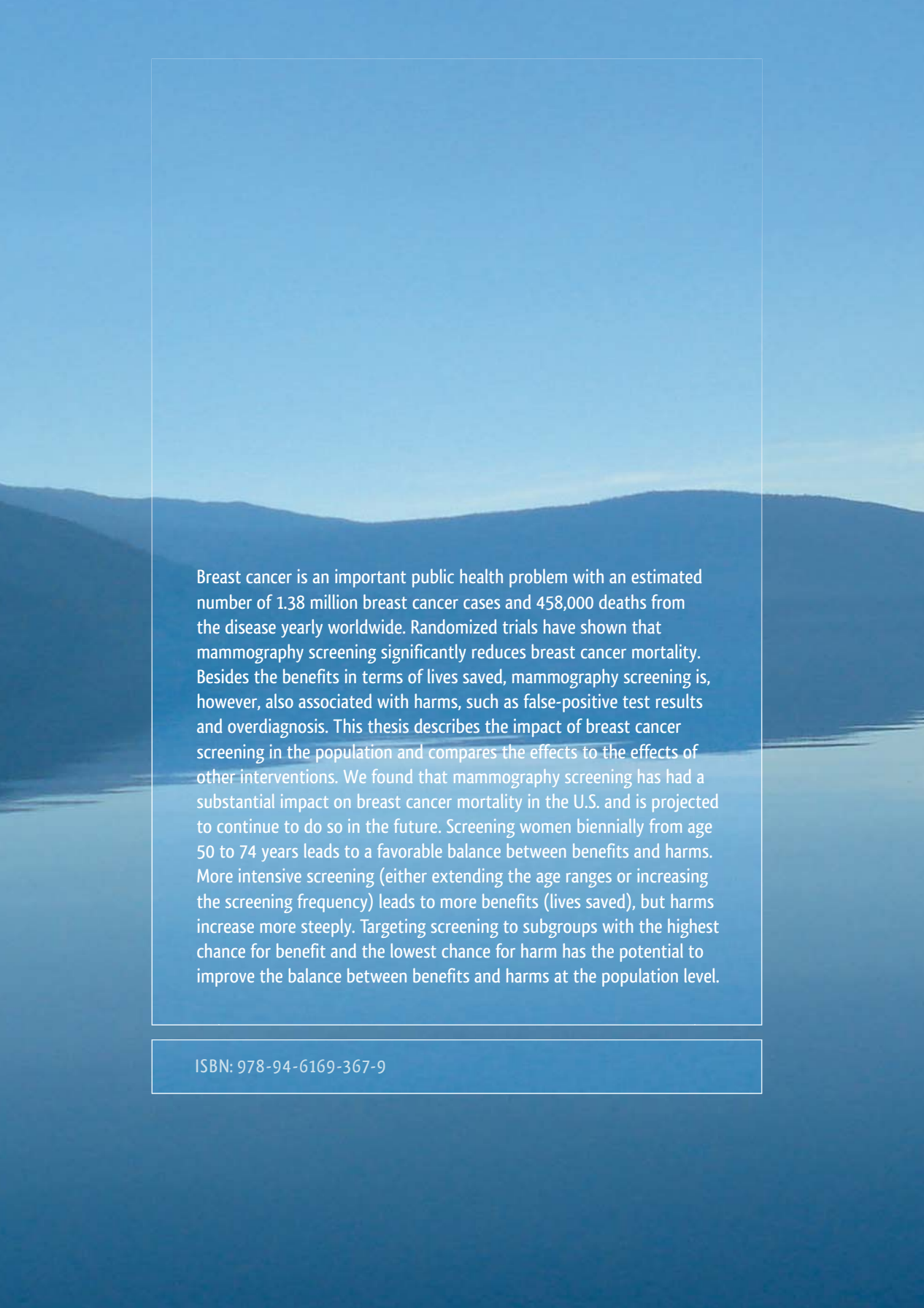
- Presentations at Cancer Intervention and Surveillance Modeling Network (CISNET) meetings, National Cancer Institute, USA	2008-2013	6.0
- Oral presentation. Research meeting at the department of Public Health, Erasmus MC, Rotterdam	2009	0.6
- Oral presentation. European Breast Cancer Conference, Barcelona, Spain	2010	1.0
- Presentations at Breast Cancer Surveillance Consortium meetings, USA	2010-2011	1.2
- Oral [invited] presentation Breast Densitometry Workshop, San Francisco, USA	2011	1.0
- Oral presentation. Society for Medical Decision Making, Chicago, USA	2011	1.0
- Oral presentation. COPD modeling meeting, Amsterdam	2012	0.8
- Oral presentation [in Dutch]. Symposium over 10-jaar borstkankerscreening in Vlaanderen, Hasselt, België	2012	0.6

(Inter)national conferences

- European Breast Cancer Conference, Berlin, Germany	2008	0.8
- European Breast Cancer Conference, Barcelona, Spain	2010	1.0
- Society for Medical Decision Making, Chicago, USA	2011	0.6
- International Breast Densitometry Workshop, San Francisco, USA	2011	0.5

2. Teaching

	Year	**Workload (ECTS)**
Lecturing		
- Sandwichcursus Mammaradiologie & Thoraxradiologie; "DCIS: overdiagnosis versus benefits"	2011	0.7
Supervising practicals and excursions, Tutoring		
- Supervising 3rd year curriculum medical students for a community project, which is part of education theme 3.C 'Arts en volksgezondheid' at the Erasmus MC, Rotterdam	2012	0.6
Supervising Master's thesis		
- Lissanne van Lier: "Costs and health consequences of film and digital mammography for women served through the U.S. National Breast and Cervical Cancer Early Detection Program."	2010-2011	1.5



Breast cancer is an important public health problem with an estimated number of 1.38 million breast cancer cases and 458,000 deaths from the disease yearly worldwide. Randomized trials have shown that mammography screening significantly reduces breast cancer mortality. Besides the benefits in terms of lives saved, mammography screening is, however, also associated with harms, such as false-positive test results and overdiagnosis. This thesis describes the impact of breast cancer screening in the population and compares the effects to the effects of other interventions. We found that mammography screening has had a substantial impact on breast cancer mortality in the U.S. and is projected to continue to do so in the future. Screening women biennially from age 50 to 74 years leads to a favorable balance between benefits and harms. More intensive screening (either extending the age ranges or increasing the screening frequency) leads to more benefits (lives saved), but harms increase more steeply. Targeting screening to subgroups with the highest chance for benefit and the lowest chance for harm has the potential to improve the balance between benefits and harms at the population level.

ISBN: 978-94-6169-367-9