NIH Consensus Development Conference on Breast Cancer Screening for Women Ages 40–49

PROGRAM AND ABSTRACTS

(Online Edition)

NIH Consensus Development Conference
January 21–23, 1997

Natcher Conference Center
National Institutes of Health
Bethesda, Maryland

Sponsored by the National Cancer Institute and the NIH Office of Medical Applications of Research; cosponsored by the National Institute on Aging and the Office of Research on Women's Health of the NIH and the Centers for Disease Control and Prevention.

Continuing Education Sponsorship by
National Institutes of Health
Office of Education

Contents

Introduction to the NIH Consensus Development Conference on Breast Cancer Screening for Women Ages 40–49 1

Agenda 3

Panel 7

Speakers 8

Planning Committee 11

Abstracts 14

I. Introduction and Overview

Breast Cancer Screening Among Women in Their Forties: An Overview of the Issues
Suzanne W. Fletcher, M.D., M.Sc. 15

A Breast Cancer Survivor's Perspective
Zora Kramer Brown 19

What Do Women Want to Know?
Maryann Napoli 22

Screening Fundamentals
Robert A. Smith, Ph.D. 24

II. Screening Benefits

A. Basic Description of Randomized Controlled Clinical Trials Design

Basic Designs of Randomized Clinical Trials of Screening
Freda E. Alexander, M.D. 27

Study Design II
Eugenio Paci, M.D. . 29

B. Individual Randomized Controlled Clinical Trials: Description and Results

Periodic Screening for Breast Cancer: The Health Insurance Plan of Greater New York Randomized Controlled Trial
Sam Shapiro 33

The Edinburgh Randomized Trial of Breast Cancer Screening
Freda E. Alexander, M.D. 41

The Canadian National Breast Screening Study: Update on Breast Cancer Mortality
Anthony B. Miller, M.B., F.R.C.P. 42

Swedish Studies

Recent Results from the Swedish Two-County Trial: The Effects of Age, Histological Type, and Mode of Detection
László Tabár, M.D.; Gunnar Fagerberg; and Stephen W. Duffy, M.Sc. 45

The Malmo Mammographics Screening Trial: Update on Results and a Harm-Benefit Analysis
Ingvar Andersson, M.D., Ph.D.. 48

The Stockholm Mammographic Screening Trial: Risks and Benefits
Jan Frisell, M.D., and Elisabet Lidbrink, M.D. 50

The Gothenburg Breast Screening Trial: Results from 11 Years Followup
Nils Bjurstam, M.D., Ph.D.; Lena Björnel; and Stephen W. Duffy, M.Sc. 52

C. Overview Analyses

Update of the Overview of the Swedish Randomized Trials on Breast Cancer Screening with Mammography
Lennarth Nyström, Stig Wall, LarsErik Rutqvist, Ingvar Andersson, Nils Bjurstam, Gunnar Fagerberg, Jan Frisell, László Tabár, Lars-Gunnar Larsson 54

Variation in the Effect of Breast Screening by Year of Followup
Brian Cox, M.D., Ph.D., F.A.F.P.H.M. 59

The Quality and Interpretation of Mammographic Screening Trials for Women Ages 40–49
Paul Glasziou, M.B.B.S., Ph.D. 63

Efficacy of Screening Mammography: Relative and Absolute Benefit
Karla M. Kerlikowske, M.D. 65

Benefit of Mammography Screening in Women Ages 40–49: Current Evidence from Randomized Controlled Trials
Charles R.. Smart, M.D., F.A.C.S.; R. Edward Hendrick, Ph.D.; James H. Rutledge III, M.S.; and Robert A. Smith, Ph.D... 70

Markov Models for Breast Tumor Progression: Estimates from Empirical Screening Data and Implications for Screening
Stephen W. Duffy, M.Sc., N.E. Day; László Tabár, M.D.; and H.H. Chen . 76

Quantitative Interpretation of Age-Specific Mortality Reductions from Trials by Microsimulation
Harry J. de Koning, Ph.D., and Rob Boer, M.Sc. 78

III.Screening Outcomes: Clinical Experiences

Problems With the Randomized Controlled Trials of Screening and Inappropriate Analysis of Breast Cancer Data
Daniel B. Kopans, M.D., F.A.C.R. 82

Results from the National Breast and Cervical Cancer Early Detection Program, 1991–1995
Nancy C. Lee, M.D. 87

Screening Outcomes: Clinical Experience With Service Screening Using Modern Mammography
Edward A. Sickles, M.D., F.A.C.R.. 89

Outcomes of Modern Screening Mammography
Karla M. Kerlikowske, M.D. 95

Mammography Outcomes in a Practice Setting by Age: Prognostic Factors, Sensitivity, and Positive Biopsy Rate
Michael N. Linver, M.D., F.A.C.R.. 99

IV. Risks in Absolute and Relative Terms

Radiation Risk
Stephen A. Feig, M.D., and R. Edward Hendrick, Ph.D.. 104

Mammography Versus Clinical Examination of the Breasts
Cornelia J. Baines, M.D., F.A.C.E. 109

The Psychosocial Consequences of Mammography
Barbara K. Rimer, Dr.P.H. 113

Variation of Benefits and Harms of Breast Cancer Screening With Age
Russell P. Harris, M.D., M.P.H. 117

V. Other Benefits

Screening for Breast Cancer in Younger Women Ages 40–49
Helena R. Chang, M.D., Ph.D., and Kirby I. Bland, M.D.. 123

Detection and Treatment Trends: A Clinical Experience
Blake Cady, M.D. 125

Increases in Ductal Carcinoma *In Situ* in Relation to Mammography: A Dilemma
Virginia L. Ernster, Ph.D. 129

Introduction to the NIH Consensus Development Conference on Breast Cancer Screening for Women Ages 40–49

A number of randomized clinical trials have shown clearly that early detection of breast cancer by mammography, with and without clinical breast examination at regular intervals ranging from 1 year to 33 months, reduces breast cancer mortality in women ages 50–69 by about a third. However, the picture is not as clear for women 40–49 years of age, and, worldwide, experts continue to examine the data regarding the use of mammography in this age group. Data from the American, Swedish, Canadian, and Edinburgh (U.K.) clinical trials will be presented at the conference in an attempt to help clarify these issues.

This conference will bring together the investigators who have conducted the randomized clinical trials, epidemiologists, statisticians, radiologists, oncologists, and other experts, as well as representatives of the public, to present and discuss the latest data and data analyses.

Following $1^1/_2$ days of presentations and audience discussion, an independent, non-Federal consensus panel will weigh the scientific evidence and write a draft statement that it will present to the audience on the third day. The consensus statement will address the following key questions:

- Is there a reduction in mortality from breast cancer due to screening women ages 40–49 with mammography, with or without physical examination? How large is the benefit? How does this change with age?

- What are the risks associated with screening women ages 40–49 with mammography, and with physical examination? How large are the risks? How do they change with age?

- Are there other benefits? If so, what are they? How do they change with age?

- What is known about how the benefits and risks of breast cancer screening differ based on known risk factors for breast cancer?

- What are the directions for future research?

On the final day of the meeting, the conference and panel chairperson, Leon Gordis, M.D., Professor, Department of Epidemiology, School of Hygiene and Public Health, Associate Dean for Admissions and Academic Affairs, School of Medicine, Johns Hopkins University, will read the draft statement to the conference audience and invite comments and questions. A press conference will follow to allow the panel and chairperson to respond to questions from media representatives.

GENERAL INFORMATION

Conference sessions will be held in the Natcher Conference Center (Building 45), NIH, 9000 Rockville Pike, Bethesda, Maryland. Sessions will run from 8:30 a.m. to 5:30 p.m. on Tuesday, 8 a.m. to 12:45 p.m. on Wednesday, and 9 to 11 a.m. on Thursday. The telephone number for the message center is 301-496-9966.

CAFETERIA

The cafeteria is located on the lobby level and is open daily from 7:00 a.m. to 3:00 p.m.

CONTINUING EDUCATION CREDIT

The purpose of this Consensus Development Conference is to evaluate the benefits and risks of mammography screening in women ages 40–49.

The conference will

- present in open, public sessions state-of-the-art information regarding the benefits and risks of mammography screening in women 40–49 years of age,
- prepare a statement in response to the five specific questions, and
- inform the biomedical research and clinical practice communities and the general public of the conclusions and recommendations of the panel.

The National Institutes of Health is accredited by the Accreditation Council for Continuing Medical Education to sponsor continuing medical education for physicians.

The National Institutes of Health designates this continuing medical education activity for a maximum of 15 credit hours in Category I of the Physician's Recognition Award of the American Medical Association.

SPONSORS

The primary sponsors of this conference are the National Cancer Institute and the NIH Office of Medical Applications of Research. The conference is cosponsored by the National Institute on Aging and the Office of Research on Women's Health of the NIH and the Centers for Disease Control and Prevention. This is the 103rd Consensus Development Conference held by the NIH since the establishment of the Consensus Development Program in 1977.

Agenda

Tuesday, January 21, 1997

8:30 a.m.	Welcome	Richard D. Klausner Director National Cancer Institute Vivian W. Pinn Director Office of Research on Women's Health
	Charge to the Panel	John H. Ferguson Director Office of Medical Applications of Research
	Conference Issues	Leon Gordis Conference and Panel Chairperson
	I. Introduction and Overview	
9:00 a.m.	Breast Cancer Screening Among Women in Their Forties: An Overview of the Issues	Suzanne W. Fletcher
9:15 a.m.	A Breast Cancer Survivor's Perspective	Zora Kramer Brown
9:30 a.m.	What Do Women Want to Know?	Maryann Napoli
9:45 a.m.	Screening Fundamentals	Robert A. Smith
10:00 a.m.	Discussion	
	II. Screening Benefits	
	A. Basic Description of Randomized Controlled Clinical Trials Design	
10:25 a.m.	Basic Designs of Randomized Clinical Trials of Screening	Freda E. Alexander
10:45 a.m.	Study Design II	Eugenio Paci
11:05 a.m.	Discussion	
	B. Individual Randomized Controlled Clinical Trials: Description and Results	
11:30 a.m.	Periodic Screening for Breast Cancer: The Health Insurance Plan of Greater New York Randomized Controlled Trial	Sam Shapiro

11:45 a.m.	The Edinburgh Randomized Trial of Breast Cancer Screening	Freda E. Alexander
12:00 p.m.	The Canadian National Breast Screening Study: Update on Breast Cancer Mortality	Anthony B. Miller
12:15 p.m.	Lunch	
	Swedish Studies	
1:15 p.m.	Recent Results from the Swedish Two-County Trial: The Effects of Age, Histological Type, and Mode of Detection	László Tabár
1:30 p.m.	Results from the Malmö Breast Screening Trial	Ingvar Andersson
1:45 p.m.	The Stockholm Mammographic Screening Trial: Risks and Benefits	Jan Frisell
2:00 p.m.	The Gothenburg Breast Screening Trial: Results from 11 Years Followup	Nils Bjurstam
2:15 p.m.	Discussion	
	C. Overview Analyses	
3:15 p.m.	Update of the Overview of the Swedish Randomized Trials on Breast Cancer Screening with Mammography	Lennarth Nystrom
3:30 p.m.	Variation in the Effect of Breast Screening by Year of Followup	Brian Cox
3:45 p.m.	The Quality and Interpretation of Mammographic Screening Trials for Women Ages 40–49	Paul Glasziou
4:00 p.m.	Efficacy of Screening Mammography: Relative and Absolute Benefit	Karla M. Kerlikowske
4:15 p.m.	Benefit of Mammography Screening in Women Ages 40–49: Current Evidence from Randomized Controlled Trials	Charles R. Smart
4:30 p.m.	Markov Models for Breast Tumor Progression: Estimates from Empirical Screening Data and Implications for Screening	Stephen W. Duffy
4:45 p.m.	Quantitative Interpretation of Age-Specific Mortality Reductions from Trials by Microsimulation	Harry J. de Koning
5:00 p.m.	Discussion	
5:30 p.m.	Adjournment until Wednesday	

Wednesday, January 22, 1997

	III. Screening Outcomes: Clinical Experiences	
8:00 a.m.	Problems With the Randomized Controlled Trials of Screening and Inappropriate Analysis of Breast Cancer Data	Daniel B. Kopans
8:15 a.m.	Results from the National Breast and Cervical Cancer Early Detection Program, 1991–1995	Nancy C. Lee
8:30 a.m.	Screening Outcomes: Clinical Experience With Service Screening Using Modern Mammography	Edward A. Sickles
8:45 a.m.	Outcomes of Modern Screening Mammography	Karla M. Kerlikowske
9:00 a.m.	Mammography Outcomes in a Practice Setting by Age: Prognostic Factors, Sensitivity, and Positive Biopsy Rate	Michael N. Linver
9:15 a.m.	Discussion	
	IV. Risks in Absolute and Relative Terms	
9:45 a.m.	Radiation Risk	Stephen A. Feig
10:00 a.m.	Mammography Versus Clinical Examination of the Breasts	Cornelia J. Baines
10:15 a.m.	The Psychosocial Consequences of Mammography	Barbara K. Rimer
10:30 a.m.	Variation of Benefits and Harms of Breast Cancer Screening With Age	Russell P. Harris
10:50 a.m.	Discussion	
	V. Other Benefits	
11:30 a.m.	Screening for Breast Cancer in Younger Women Ages 40–49	Helena R. Chang
11:45 a.m.	Detection and Treatment Trends: A Clinical Experience	Blake Cady
12:00 p.m.	Increases in Ductal Carcinoma *In Situ* in Relation to Mammography: A Dilemma	Virginia L. Ernster
12:15 p.m.	Discussion	
12:45 p.m.	Adjournment until Thursday	

Thursday, January 23, 1997

9:00 a.m.	Presentation of the Consensus Statement	Leon Gordis Conference and Panel Chairperson
9:30 a.m.	Discussion	
11:00 a.m.	Panel Meets in Executive Session	
1:00 p.m.	Press Conference	
2:00 p.m.	Adjournment	

Panel

Panel and Conference
Chairperson: Leon Gordis, M.D.
Professor
Department of Epidemiology
School of Hygiene and Public Health
Associate Dean for Admissions and Academic Affairs
School of Medicine
Johns Hopkins University
Baltimore, Maryland

Donald A. Berry, Ph.D.
Professor
Institute of Statistics and Decision Science
Duke University
Durham, North Carolina

Susan Y. Chu, Ph.D., M.P.H.
Associate Director
Center for Health Studies
Group Health Cooperative
Seattle, Washington

Laurie L. Fajardo, M.D.
Professor of Radiology and
Vice Chair for Research
Department of Radiology
University of Virginia
Charlottesville, Virginia

David G. Hoel, Ph.D.
Professor and Chairman
Department of Biometry and Epidemiology
Medical University of South Carolina
Charleston, South Carolina

Leslie Laufman, M.D.
President
Hematology Oncology Consultants
Worthington, Ohio

Constance A. Rufenbarger
Project Development
The Catherine Peachey Fund, Inc.
Warsaw, Indiana

Julia R. Scott
President and CEO
National Black Women's Health Project, Inc.
Washington, DC

Daniel C. Sullivan, M.D.
Associate Professor of Radiology
University of Pennsylvania Medical Center
Philadelphia, Pennsylvania

John H. Wasson, M.D.
Herman O. West Professor of Geriatrics
Center for the Aging
Dartmouth Medical School
Hanover, New Hampshire

Carolyn L. Westhoff, M.D., M.S.
Associate Professor
Obstetrics, Gynecology, and Public Health
Columbia University College of Physicians
and Surgeons
New York, New York

Ruthann Zern, M.D.
Obstetrician/Gynecologist
Private Practice
St. Joseph's Hospital
Greater Baltimore Medical Center
Towson, Maryland

Speakers

Freda E. Alexander, M.D.
Epidemiologist
Department of Public Health Sciences
The University of Edinburgh
Medical School
Edinburgh, Scotland

Ingvar Andersson
Department of Diagnostic Radiology
University Hospital of Malmö
Malmö, Sweden

Cornelia J. Baines, M.D., M.S.C., F.A.C.E.
Associate Professor
Department of Preventive Medicine
and Biostatistics
Faculty of Medicine
University of Toronto
Toronto, Ontario, Canada

Nils Bjurstam, M.D., Ph.D.
Radiology Clinic
Section of Mammography
NÄL Hospital
Trollhättan, Sweden

Zora Kramer Brown
Founder and Chairperson
Breast Cancer Resource Committee
Washington, DC

Blake Cady, M.D.
Professor of Surgery, Harvard School of
Medicine and
Interim Chief of Surgery, Surgical Oncology
Division
Department of Surgery
Beth Israel Deaconess Medical Center
Boston, Massachusetts

Helena R. Chang, M.D., Ph.D., F.A.C.S.
Surgeon and Director of the Hybridoma
Laboratory
Associate Professor of Surgery and
Pathobiology Program
Roger Williams Medical Center
Brown University
Providence, Rhode Island

Brian Cox, M.D., Ph.D., F.A.F.P.H.M.
Senior Research Fellow and Public Health
Physician
Department of Preventive and Social
Medicine
University of Otago Medical School
Dunedin, New Zealand

Harry J. de Koning, Ph.D.
Assistant Professor of Public
Health/Medical Technology Assessment
Department of Public Health
Erasmus University
Rotterdam, The Netherlands

Stephen W. Duffy, M.Sc.
Senior Scientist
MRC Biostatistics Unit
Institute of Public Health
University Forvie Site
Cambridge University
Cambridge, United Kingdom

Virginia L. Ernster, Ph.D.
Professor and Vice Chair
Department of Epidemiology
and Biostatistics
Associate Director of the Cancer Center
School of Medicine
University of California, San Francisco
San Francisco, California

Stephen A. Feig, M.D.
Professor of Radiology
Jefferson Medical College
Director, Breast Imaging
Department of Radiology
Thomas Jefferson University Hospital
Philadelphia, Pennsylvania

Suzanne W. Fletcher, M.D., M.Sc.
Professor
Primary Care Division
Department of Ambulatory Care and Prevention
Harvard Medical School and Harvard Pilgrim Health Care
Department of Epidemiology
Harvard School of Public Health
Boston, Massachusetts

Jan Frisell, M.D.
Assistant Professor
Department of Surgery and Oncology
Stockholm South Hospital
Stockholm, Sweden

Paul Glasziou, M.B.B.S., Ph.D.
Senior Lecturer in Clinical Epidemiology
Department of Social and Preventive Medicine
University of Queensland Medical School
Herston, Queensland, Australia

Russell P. Harris, M.D., M.P.H.
Assistant Professor of Medicine
Division of General Medicine and Clinical Epidemiology and
Co-Director, Program on Health Promotion and Disease Prevention
University of North Carolina at Chapel Hill School of Medicine
Chapel Hill, North Carolina

Karla M. Kerlikowske, M.D.
Associate Director
Women Veterans Comprehensive Health Center
Veterans Affairs Medical Center
Assistant Professor
Department of Medicine, Epidemiology, and Biostatistics
University of California, San Francisco
San Francisco, California

Daniel B. Kopans, M.D., F.A.C.R.
Associate Professor of Radiology
Harvard Medical School
Director of Breast Imaging
Department of Radiology
Massachusetts General Hospital
Boston, Massachusetts

Nancy C. Lee, M.D.
Associate Director for Science
Division of Cancer Prevention and Control
Centers for Disease Control and Prevention
Atlanta, Georgia

Michael N. Linver, M.D., F.A.C.R.
Director of Mammography
X-Ray Associates of New Mexico, PC
Clinical Associate Professor
Department of Radiology
University of New Mexico School of Medicine
Albuquerque, New Mexico

Anthony B. Miller, M.B., F.R.C.P.
Director, National Breast Screening Study
Professor and Chairman
Department of Preventive Medicine and Biostatistics
University of Toronto
Toronto, Ontario, Canada

Maryann Napoli
Associate Director
Center for Medical Consumers
New York, New York

Lennarth Nystrom
Biostatistician/Epidemiologist
Department of Epidemiology and Public Health
Umea University
Umea, Sweden

Eugenio Paci, M.D.
Epidemiologia Unit
Center for the Study and Prevention of Cancer
Florence, Italy

Barbara K. Rimer, Dr.P.H.
Director, Cancer Prevention, Detection, and Control Research
Cancer Control Department
Duke Comprehensive Cancer Center
Duke University Medical Center
Durham, North Carolina

Sam Shapiro
Professor Emeritus
Health Policy and Management
School of Hygiene and Public Health
Johns Hopkins University
Baltimore, Maryland

Edward A. Sickles, M.D., F.A.C.R.
Professor of Radiology
Chief, Breast Imaging Section
Department of Radiology
University of California Medical Center
University of California, San Francisco
San Francisco, California

Charles R. Smart, M.D., F.A.C.S.
Surgeon
Salt Lake City, Utah

Robert A. Smith, Ph.D.
Senior Director
Cancer Control Department
American Cancer Society
Atlanta, Georgia

László Tabár, M.D.
Associate Professor and Director
Department of Mammography
Falun Central Hospital
Falun, Sweden

Planning Committee

Chairperson: John K. Gohagan, Ph.D.
Branch Chief
Early Detection Branch
National Cancer Institute
National Institutes of Health
Bethesda, Maryland

Jeffrey S. Abrams, M.D.
Senior Investigator
Cancer Therapy Evaluation Program
Division of Cancer Treatment, Diagnosis, and Centers
National Cancer Institute
National Institutes of Health
Bethesda, Maryland

Anne Bavier, M.N., F.A.A.N.
Deputy Director
Office of Research on Women's Health
Office of the Director
National Institutes of Health
Bethesda, Maryland

Frank Bellino, Ph.D.
Health Scientist Administrator
Biology of Aging Program
National Institute on Aging
National Institutes of Health
Bethesda, Maryland

Jerry M. Elliott
Program Management and Analysis Officer
Office of Medical Applications of Research
National Institutes of Health
Bethesda, Maryland

Virginia L. Ernster, Ph.D.
Professor and Vice Chair
Department of Epidemiology and Biostatistics
School of Medicine
Associate Director of the Cancer Center
University of California, San Francisco
San Francisco, California

John H. Ferguson, M.D.
Director
Office of Medical Applications of Research
National Institutes of Health
Bethesda, Maryland

Leslie G. Ford, M.D.
Associate Director
Early Detection and Community Oncology Program
Division of Cancer Prevention and Control
National Cancer Institute
National Institutes of Health
Bethesda, Maryland

Leon Gordis, M.D.
Conference and Panel Chairperson
Professor
Department of Epidemiology
School of Hygiene and Public Health
Associate Dean for Admissions and Academic Affairs
School of Medicine
Johns Hopkins University
Baltimore, Maryland

William H. Hall
Director of Communications
Office of Medical Applications of Research
National Institutes of Health
Bethesda, Maryland

Douglas B. Kamerow, M.D., M.P.H.
Director
Office of the Forum for Quality and Effectiveness in Health Care
Agency for Health Care Policy and Research
Rockville, Maryland

Daniel B. Kopans, M.D., F.A.C.R.
Associate Professor of Radiology
Harvard Medical School
Director of Breast Imaging
Department of Radiology
Massachusetts General Hospital
Boston, Massachusetts

Barnett S. Kramer, M.D., M.P.H.
Deputy Director
Division of Cancer Prevention and Control
National Cancer Institute
National Institutes of Health
Bethesda, Maryland

Amy S. Langer, M.B.A.
Executive Director
National Alliance of Breast Cancer Organizations
New York, New York

Elaine Lee
Program Analyst
Planning, Evaluation, and Analysis Branch
National Cancer Institute
National Institutes of Health
Bethesda, Maryland

Nancy C. Lee, M.D.
Associate Director for Science
Division of Cancer Prevention and Control
Centers for Disease Control and Prevention
Atlanta, Georgia

Carl M. Mansfield, M.D., D.Sc., F.A.C.R., F.A.C.N.M.
Associate Director
Radiation Research Program
Division of Cancer Treatment, Diagnosis, and Centers
National Cancer Institute
National Institutes of Health
Bethesda, Maryland

Anthony B. Miller, M.B., F.R.C.P.
Director, National Breast Screening Study
Professor and Chairman
Department of Preventive Medicine and Biostatistics
University of Toronto
Toronto, Ontario, Canada

Sue Moss, Ph.D.
Cancer Screening and Evaluation Unit
Institute of Cancer Research
Royal Cancer Hospital
Sutton, Surrey, England

Nancy Nelson
Science Writer
National Cancer Institute
National Institutes of Health
Bethesda, Maryland

Gillian Newstead, M.D.
Director of Breast Imaging
Breast Imaging Center
New York University
New York, New York

Cherie Nichols, M.B.A.
Chief
Planning, Evaluation, and Analysis Branch
National Cancer Institute
National Institutes of Health
Bethesda, Maryland

Vivian W. Pinn, M.D.
Director
Office of Research on Women's Health
Office of the Director
National Institutes of Health
Bethesda, Maryland
Alan Rabson, M.D.
Deputy Director
National Cancer Institute
National Institutes of Health
Bethesda, Maryland

Barbara K. Rimer, Dr.P.H.
Director, Cancer Prevention,
Detection, and Control Research
Cancer Control Department
Duke Comprehensive Cancer Center
Duke University Medical Center
Durham, North Carolina

Sam Shapiro
Professor Emeritus
Health Policy and Management
School of Hygiene and Public Health
Johns Hopkins University
Baltimore, Maryland

Edward A. Sickles, M.D., F.A.C.R.
Professor of Radiology
Chief, Breast Imaging Section
Department of Radiology
University of California Medical Center
University of California, San Francisco
San Francisco, California

Robert A. Smith, Ph.D.
Senior Director
Cancer Control Department
American Cancer Society
Atlanta, Georgia
Edward Sondik, Ph.D.
Director
National Center for Health Statistics
Centers for Disease Control and Prevention
Hyattsville, Maryland

László Tabár, M.D.
Associate Professor and Director
Department of Mammography
Falun Central Hospital
Falun, Sweden

Robert E. Tarone, Ph.D.
Mathematical Statistician
Division of Cancer Epidemiology and
Genetics
National Cancer Institute
National Institutes of Health
Bethesda, Maryland

Rosemary Yancik, Ph.D.
Chief, Cancer Section
Geriatrics Program
National Institute on Aging
National Institutes of Health
Bethesda, Maryland

Abstracts

The following abstracts of presentations to the NIH Consensus Development Conference on Breast Cancer Screening for Women Ages 40–49 were furnished by presenters in advance of the conference. The abstract for the presentation on the Results of the Malmö Breast Screening Trial, by Ingvar Andersson, does not appear. This book is designed for the use of panelists and participants in the conference and as a pertinent reference document for anyone interested in the conference deliberations. We are grateful to the authors who have summarized their materials and made them available in a timely fashion.

John K. Gohagan, Ph.D.
Branch Chief
Early Detection Branch
National Cancer Institute
National Institutes of Health
Bethesda, Maryland

Jerry M. Elliott
Program Management and Analysis Officer
Office of Medical Applications
of Research
National Institutes of Health
Bethesda, Maryland

Breast Cancer Screening Among Women in Their Forties: An Overview of the Issues

Suzanne W. Fletcher, M.D., M.Sc.

Introduction

Although 85 percent of breast cancers occur in women after they reach the age of 50, breast cancer is the number one cause of cancer death for women in their forties. Each year, for every 100,000 women in their forties, 163 are diagnosed with breast cancer and 30 die of the disease.[1] Paradoxically, despite the importance of breast cancer for women in their forties, the disease mercifully is uncommon in this age group. For every 1,000 women entering their fifth decade, approximately 16 will develop breast cancer at some time before their 50th birthday. Eight to 10 of these 16 women will survive the cancer, with or without screening, partly because of recent therapeutic advances. Thus, screening trials and screening activity are directed at the 6–8 women in every 1,000 who might be saved by earlier detection during their forties. If screening decreases mortality by as much as 25 percent, it would save approximately 2 out of the 1,000 women.

Any fatal illness striking persons in the prime of life is a terrible occurrence, but breast cancer is doubly so because it threatens not only a woman's life but also an emotionally and sexually important part of her body. Fear of breast cancer is so great that women in their forties overestimate their risk of dying of breast cancer 20-fold and their risk of developing breast cancer 6-fold.[2] With such a terrifying disease, it is important to find better ways to cure and prevent this disease.

What can screening, especially screening with mammography, contribute to the control of breast cancer in women in their forties? Three issues should be addressed when considering any kind of screening intervention: benefits, possible adverse effects, and costs. This conference focuses on mortality benefits and possible adverse effects of breast cancer screening. There is one talk examining breast conservation, a benefit other than mortality. There is no discussion of costs or cost-effectiveness.

Mortality Benefits

Most attention has been given to mortality benefits of breast cancer screening. Eight randomized controlled trials of mammography, with or without clinical breast examination, have been conducted in four countries. At the National Cancer Institute (NCI) International Workshop on Breast Cancer Screening in 1993, all trial results found mortality benefits (two with statistically significant results), among women ages 50–69.[3] A meta-analysis of all the trials combined found a 34 percent reduction in breast cancer mortality after 7 years of followup.[4] However, the findings among younger women were less clear. The meta-analysis showed no effect at 7 years of followup, the Health Insurance Plan (HIP) study showed a 25 percent benefit after 10–18 years of followup, and the combined Swedish trials showed a nonsignificant 10–13 percent benefit at 12 years.[5]

In April 1996, an update of the studies was reported. In the eight studies with women ages 40–49 (in two trials, women were ages 45–49), five showed mortality benefits after 15–16 years of followup and three showed no benefit after 10–16 years. Meta-analyses demonstrated a statistically significant mortality benefit of 23, 24, or 15 percent, depending on which trials were included. These important new findings stimulated the formation of this consensus development conference.

As results of the trials continue to accrue, it appears that the time required to demonstrate beneficial screening effect varies by age. Whereas mortality differences between screened and unscreened groups began to emerge after only a few years of followup for women ages 50–69, most of these same studies show effects more slowly for women in their forties. In the combined Swedish studies, mortality rates were similar in the invited and control groups during the first 8–10 years of followup, after which a beneficial effect of screening began to emerge, and grew to be statistically significant at 16 years of followup. The same trend occurred in the HIP study. The cause of this time difference in effect by age is not yet clear. Perhaps screening picks up such early cancers in younger women that it takes longer for mortality benefits to be shown.[6] On the other hand, perhaps the effect is due to women entering these studies during their forties aging into their fifties during the course of the study, when screening benefit becomes apparent.[7,8] Analyses are needed to determine the degree to which each explanation could account for the time delay seen in the younger age group.

All published analyses of results are reported according to the age of women at entry into the trial, not the age at the time of diagnosis of breast cancer. This approach is necessary to preserve the comparability of the screened and control groups. Nevertheless, when there is the possibility that effect of breast cancer screening varies by age, information about the age of diagnosis is needed. Most trials have not yet provided this information. The issue is especially important in the two trials in which women only as young as 45 years old were included. In the HIP study, Shapiro and colleagues demonstrated that screened women ages 45–49 at entry into the study gained benefit when their cancers were diagnosed after age 50 but not when their cancers were diagnosed before age 50.[9] On the other hand, Tabar and Duffy reported the relative mortality in the Swedish Two-County Trial was .95 for women after they turned age 50 and .85 for women before age 50. They also reported that 36 percent of cancers found in the women in the forties group were diagnosed after the women turned 50.[10] Information from the other trials is needed.

Why would a screening test for breast cancer have differential effects by age? Part of the explanation may be the lower accuracy of screening tests in younger women.[3] Also, breast cancer growth rates may differ by age of the woman. Tabar and colleagues found that the mean sojourn time (time in the preclinical detectable state) was 1.25 years for women in their forties and 3.03 years for women in their fifties.[11] Whether and how estrogen levels and menopause, rather than age per se, influence effectiveness of breast cancer screening remains unclear and needs to be determined. It is also important to determine the effect of screening in groups at high risk for breast cancer.

Adverse Effects

Possible adverse effects of breast cancer screening include the question of radiation risk, discomfort during mammography, and adverse psychosocial sequelae of false-positive mammograms, all of which will be addressed in the conference. Most previous work has analyzed the degree of accuracy of screening tests and the related problems of false-positive and false-negative results. Breast cancer screening in younger women is neither as sensitive nor as specific as that in older women.[3] In the United States, 11 percent of all screening mammograms require followup investigations.[12] More than 90 percent of such mammograms are false-positives; the resultant numbers of additional procedures, including biopsies and psychological distress can be substantial.[13,14] Because breast cancer screening is periodically repeated, the percentage of women ultimately experiencing a false-positive mammogram may be substantial. The cumulative false-positive rate of mammography over an extended period of time needs study.

As more breast cancers are diagnosed through screening, the percentage of ductal carcinomas *in situ* (DCIS) is rising, as will be discussed at the conference. In the United States, DCIS accounted for approximately 12 percent of all breast cancers in 1991, with a slightly higher percentage in women under 50 years of age.[1] In one screening program, 43 percent of cancers detected among women in their forties were DCIS, and concerns have been raised about the possibility that early lesions such as DCIS could lead to overdiagnosis of breast cancer.[15] There is an urgent need for studies of the natural history and appropriate management of DCIS.

In sum, progress has been made in understanding the mortality benefits of a breast cancer screening program for women in their forties. If routine screening occurred for all women in their forties, results of randomized trials to date suggest that between 0 and 3 women per thousand will be saved. Less clear are other benefits that might occur and the hazards, especially those caused by false-positive results and possible overdiagnosis because of DCIS. Finally, costs cannot be ignored. Women need information about all these issues. They rightly demand to be involved in an important decision about their lives and their bodies. Ultimately, it is the job of medical science to search for new and better ways to maintain and promote health, and along the way to share to the best of our abilities the very complicated facts as we understand them. Armed with facts, women can then apply their own set of values to cope with the important problem of breast cancer.

References

1. Ries LAG, Miller BA, Hankey BF, Kosary CL, Harras A, Edwards BK (eds). SEER cancer statistics review, 1973-1991: Tables and graphs, National Cancer Institute. NIH Pub. No. 942789. Bethesda, MD, 1994.

2. Black WC, Nease RF, Tosteson ANA. Perceptions of breast cancer risk and screening effectiveness in women younger than 50 years of age. J Natl Cancer Inst 1995;87:720–731.

3. Fletcher SW, Black W, Harris R, Rimer BK, Shapiro S. Report of the International Workshop on Screening for Breast Cancer. J Natl Cancer Inst 1993;85:1644–56.

4. Elwood JM, Cox B, Richardson AK. The effectiveness of breast cancer screening by mammography in younger women. Online J Cur Clin Trials, 1993.

5. Nystrom L, Rutqvist LE, Wall S, et al. Breast cancer screening with mammography: Overview of Swedish randomized trials. Lancet 1993;341:973–8.

6. Kopans DB. Mammography screening and the controversy concerning women aged 40 to 49. Radiol Clin North Am 1995;33:1273–90.

7. Kerlikowske K, Grady D, Rubin SM, et al. Efficacy of screening mammography. A meta-analysis. JAMA 1995;273:149–54.

8. Fletcher SW. Why question screening mammography for women in their forties? Radiol Clin North Am 1995;33:1259–71.

9. Shapiro S, Vanet W, Strax P, Vanet L. Periodic screening for breast cancer: The Health Insurance Plan Project and its sequelae, 1963–1986. Baltimore: The Johns Hopkins University Press, 1988. p. 81.

10. Tabar L, Duffy SW, Chen HH. Re: Quantitative interpretation of age-specific mortality reductions from the Swedish breast cancer-screening trials (letter). J Nat Cancer Inst 1996;88:52–3.

11. Tabar L, Fazerberg 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.

12. Brown ML, Houn F, Sickles EA, Kessler L. Screening mammography in community practice: Positive predictive value of abnormal findings and yield of follow-up procedures. AJR 1995;165:1373–7.

13. Lidbrink E, Elfving J, Frisell J, Jonsson E. Neglected aspects of false-positive findings of mammography in breast cancer screening: analysis of false-positive cases from the Stockholm trial. BMJ 1996;312:273–6.

14. Lerman C, Tock B, Rimer B, et al. Psychological and behavioral implications of abnormal mammograms. Ann Intem Med 1991;114:657–61.

15. Ernster VL, Barclay J, Kerlikowske K, Grady D, Henderson C. Incidence of and treatment for ductal carcinoma *in situ* of the breast. JAMA 1996;275:913–8.

A Breast Cancer Survivor's Perspective

Zora Kramer Brown

As a survivor, activist, and founder of the Breast Cancer Resource Committee, my work is dedicated to providing women, in particular African-American women, with vital information regarding the importance of early detection and treatment for breast cancer. I have become dismayed and disturbed by the controversy in the scientific community regarding the debate on mammograms for women under age 50.

In my case, my family history played an important role in my education on the topic. My sisters, mother, grandmother, great-grandmother, and great aunts have all fought breast cancer and as a result we are more sensitive and knowledgeable about the dangers of not taking precautions and the importance of preventive measures such as mammography screening before age 50.

I had my first mammogram at age 21, and was diagnosed with breast cancer 10 years later. As an African-American, I am doubly offended that the tools for diagnosing breast cancer in its early stages are being taken away from the women of my race. The fastest growing group of women affected is African-American women under age 40. According to a study conducted and published in the *Proceedings of the American Society of Clinical Oncology* in May 1994, by Dr. Robert Siegel, interim director of the George Washington University Cancer Center in 1995, tumors in black women treated at the university hospital in recent years were occurring in younger black women and taking a more aggressive form.

Dr. Elizabeth A. Patterson, a University of Pennsylvania Hospital radiologist, recommends that black women have a baseline mammogram at age 35 and annual mammogram beginning at age 40. Based on published findings and other unpublished data, the National Medical Association and the American Cancer Society continue to endorse mammography screening for women in their forties. Because of improved technology, better trained technicians, and awareness of the need for screening, we are beginning to see clear benefits for screening and treating women under age 50.

Despite all of the above, there are still some members of the scientific community who suggest there is no benefit for early mammography screening. The American Cancer Society notes that the incidence of breast cancer increases with age, and approximately 77 percent of women with new diagnoses of breast cancer each year are over age 50. When age groups over age 50 are lumped together, it seems a case can be made for screening women 50 and older.

Looking more closely at the breast cancer data, the American Cancer Society estimated that women between the ages of 40 to 49 will develop approximately 33,400 new cases of breast cancer in 1996 or 18.1 percent of all new breast cancer cases for the year, while women in the 50 to 59 age group will be diagnosed with 30,900 new cases of breast cancer or 16.8 percent. These facts are rarely brought forth in the debate.

The incidence of breast cancer is actually higher for the 40–49 age group than for the 50–59 age group. Yet, both are not equally protected with early mammography screening. In fact, the incidence of new breast cancer cases does not sharply increase until women reach the 60 plus age group.

Today, there should be no excuses for not authorizing mammograms for women between ages 40 and 49. It is deplorable to deny this age group the chance to survive breast cancer with simple mammography screening that can detect the cancer before it becomes fatal.

The death rate is highest among women under age 50. In addition, although white women are at greater risk of getting cancer, black women are more likely to die from it. In fact, the breast cancer mortality rate for black women is 50 percent higher than that for white women. Breast cancer is the leading cause of cancer death among black women. Isn't this a figure that could logically be reduced with easily obtainable baseline information acquired through mammography screening?

I look at mammogram screening in the reverse of the scientific community—not as a scientist but as a survivor. If screening can prevent any number of deaths it is valuable. Mammography is important because it is the only scientific method of finding cancer at its earliest, most treatable stage. Mammography has greatly improved through the years. Since the late 1980's the image quality of mammograms has increased significantly, and since 1993 breast cancer screening facilities have been required to meet specific standards of quality in order to offer mammography. Mammograms can accurately diagnose 90 percent of breast cancers. Mammography is the single most effective method of screening—detecting cancer several years before physical symptoms are apparent.

Many in the scientific community have taken the low road on the debate that younger women do not need mammograms. In fact, mammograms are not designed to treat breast cancer, just to find its existence. Higher mortality rates occur in younger women generally because detection came too late and often because a mammogram was never prescribed and therefore the younger woman was never educated as to possible early detection methods.

The Breast Cancer Resource Committee runs several programs including a support group called **Rise-Sister-Rise.** **Rise-Sister-Rise** began in 1993 as a support group for black women diagnosed with breast cancer. Today, with more than 250 members, the group provides services, fellowship, and discussions of its members, the average age of whom is 42.

Approximately 85 percent of the women in the group have found their breast tumors through mammography screening. In most of the cases the women were sent for a mammogram by their physicians and therefore the cancer was generally found at the treatable stages. These fortunate women were often sent for mammograms because they appeared to have risk factors or symptoms that suggested screening would be important. Diagnosis of breast cancer was verified by the mammogram. The vast majority of these women were under age 50. As a result, many of these women are now aware of the importance of mammogram screening and have been able to educate other women about the importance of early detection and treatment.

Unfortunately there are still too many unhappy endings and cases where we have been too late—due to poor judgment on behalf of the scientific community about the benefits of early screening and many women's lack of knowledge about risk factors and a perception that breast cancer cannot affect women under age 50. For example, African-American women practice self-examinations more frequently than do white women. According to Dr. Melvin Gaskins, clinical oncologist at Howard University Cancer Center, this is because of the belief that self-examination provides the same accuracy and detection as mammography. This is far from the truth. In fact only a small percentage of breast cancer is detected through self examination, while a mammogram has the ability to find the disease in almost 90 percent of cases.

Although there have been advances and there is more knowledge about early detection, diagnosis is of no use if treatment is not available. Only 45 percent of women who need mammograms receive them, and Medicare and Medicaid plans authorize a mammogram only every 2 years—not frequent enough for high-risk women. The scientific community has failed in its ability to treat women effectively; instead of finding cures they make the woman under age 50 the sacrificial lamb because we have been unable to effectively treat breast cancer for this age group. Breast cancer has life and death consequences for women, not for the scientific community where the incidence of breast cancer is a statistic. Surely women under age 50 can also benefit by having their lives spared through early detection and treatment.

The fact that the scientific community does not have enough clinical trial research data on African-American women and does not understand the biology of how the disease affects African American women is of grave concern to myself and others. There is a compelling case to be made for finding the reason why so many African-American women under age 40 are aggressively affected by the disease, why more younger women die from breast cancer each year, and how we can begin to diagnose breast cancer earlier. An important step is earlier mammography screening.

This year, 42,610 young women of all races under age 40 will develop breast cancer. The scientific community deems they are too young to be screened for breast cancer. Many of them will die from breast cancer because a diagnosis of cancer will come too late. Many of these lives could possibly be spared.

In the United States, 1.3 million women are breast cancer survivors. Let's increase those numbers and give all women at high risk—and, at a minimum, those women between ages 40 and 49—the chance to live longer lives through the same early detection and treatment afforded to older women.

What Do Women Want to Know?

Maryann Napoli

It seems presumptuous to address my assigned topic, "What Do Women Want to Know?" so I'll say at the outset that I can speak of what I hear from the women who come to the Center for Medical Consumers. Breast cancer brings many to our medical library, which is open to the public in order to promote informed decision-making. I can also speak of what I learn from the growing number of breast cancer advocacy organizations around the country.[1] And lastly, I speak for myself. As a medical writer, I follow the literature on breast cancer and early detection. As a consumer advocate, I have also followed the selling of mammography screening to women, ever since the early 1970s, when the Breast Cancer Detection Demonstration Project (BCDDP) introduced the concept of mammography screening at 27 medical centers around the country. Over 280,000 women over age 35 years took part in the BCDDP, which was sponsored by the American Cancer Society (ACS) and the National Cancer Institute (NCI).

My organization, the Center for Medical Consumers, is founded on the belief that people should be encouraged to base their medical treatment decisions on the published evidence, preferably from randomized, controlled trials. We also believe that screening decisions should be held to a higher standard of evidence because they affect a healthy population.

The NCI made the right decision in 1993 when it decided to stop advising mammography screening for women in their forties, but this didn't seem to change many opinions. Women had already been sold the idea that early detection of breast cancer is good. To most, it is simply inconceivable that finding a tumor early could be anything but beneficial. If mammography doesn't save lives, many women reason, then at the very least early detection will allow for less drastic treatment.

The two most common reactions I heard from women at that time were: "I'll still have mammograms just to play it safe." and "What can we do to protect ourselves, if they take away mammography?" Those women who were always skeptical about the value of mammography for younger women—a minority, I suspect—saw the NCI decision as confirmation.

These reactions must be viewed against the backdrop of the "public education" surrounding the above-mentioned BCDDP and the more recent breast cancer awareness activities. The portrayal of younger women in breast cancer awareness ads, the use of *one in nine* and *one in eight* statistics, the women's magazines featuring the personal stories of young breast cancer survivors—all have contributed to an exaggerated idea of the odds of developing breast cancer while under age 50 years. Put that heightened awareness together with the exaggerated "public health" message—*early detection equals virtual cure*—and you have a lot of women who will remain committed to mammography screening no matter what the consensus panel decides. The far more influential information source for the public is the ACS, not the NCI, not the consensus development conferences. This is unfortunate because the ACS has a long history of overselling the value of screening and not warning of its downside. For example, women are largely in the dark about the unknowns surrounding carcinoma *in situ.*

In summary, most women probably don't know what they should be asking about mammography screening for those under age 50.

Most professional organizations with mammography screening guidelines advise women to start at age 40 years. (Many women and doctors, by the way, remain unaware that the ACS base-line mammogram recommendation has been withdrawn.) Women justifiably find it confusing when expert panels or guidelines committees look at the same evidence (or lack of it) and come to entirely different conclusions. It would suggest that expert panels do not make decisions entirely on the basis of scientific evidence.

At this point, I would like to change the title of my talk to: "What Do Women *Need* to Know?" A consensus pronouncement isn't enough unless you also educate the public about science, about differences in the quality of evidence, about how science is an ongoing process, and that what we believe today may be contradicted by tomorrow's research, etc. Share the uncertainties with us, and we won't be so shocked when a panel of experts decides to make a revision.

That's the consumer advocate side of me talking. But there's also the medical writer side of me asking: Does anyone really care about scientific evidence? Do we pay attention to randomized controlled trial (RCT) findings only when they support our preconceived ideas? I'm including doctors in my questions. Look how long it took surgeons and radiologists to let go of the ultraradical mastectomy, the radical mastectomy, and postmastectomy radiation therapy—just to give a few examples.

When the National Breast Screening Study of Canada was published, its design and mammographic techniques were assailed by American radiologists. The debate over study design that spilled into the popular media was difficult for the public to follow. Few people have the time or the skills to evaluate medical research. The suggestion that Canadian mammography techniques are behind ours, however, seemed plausible to a lot of women. But I personally found the mammography-has-improved argument troubling. Does this mean that medical technologies should never be subjected to RCTs because the findings will always be obsolete by the time they are published?

I have been asked to address the last of the questions to be considered by the Consensus Panel: "What are the directions for future research?" I believe it is time to turn researchers away from mammography and give priority to prevention. This is also the view of the National Breast Cancer Coalition, which represents a large number of breast cancer advocacy organizations around the country. Many breast cancer survivor/activists have expressed variations on this theme: "Yes, we will continue to undergo mammography screening, but researchers must find better ways to detect early breast cancers because mammography does not help a huge portion of the female population. We need to know more about the causes of breast cancer."

I look forward to the consensus panel's decision and to learning how it was developed.

Reference

1. Reese-Coulbourne J, National Breast Cancer Coalition, Washington, D.C.; Yaker A, executive director, Share, New York City; Wiener B, director, Women's Cancer Resource Center, Mineapolis; and the newsletters of Breast Cancer Action, San Francisco, and the Women's Community Cancer Project, Cambridge, Massachusetts.

Screening Fundamentals

Robert A. Smith, Ph.D.

As a disease control strategy and policy, the goal of breast cancer screening is to reduce morbidity and mortality by distinguishing those individuals in an asymptomatic population who are likely and not likely to have breast cancer.[1] The emphasis on *likelihood* is important because inherent in the concept of screening is that a person identified as likely to have the disease in question becomes then a candidate for further diagnostic testing and, if necessary, treatment. The emphasis on likelihood is also important because screening programs have inherent limitations according to the criteria described below; thus, whereas the majority of interpretations are correct, inevitably a small percentage of individuals will be incorrectly identified as having or not having the disease. The advantages of screening an asymptomatic population is that the test can identify preclinical disease with sufficient *lead time*, that is, the time before the expected onset of symptoms, to alter the natural and more adverse course of disease.

To be an effective disease control strategy, a screening program should meet fundamental criteria in three areas: (1) characteristics of the disease, (2) effectiveness of early versus later treatment, and (3) characteristics of the screening test, specifically, its accuracy and reliability, but also costs and acceptability to the target population.[2] It would be ideal if there were conventional benchmarks for these criteria, but this is not the case.

With respect to the disease, in order to screen large numbers of well people, the disease should represent a significant public health burden. This burden may be a function of any one or combination of the following—deaths, morbidity, and/or premature mortality. For most, breast cancer meets these criteria well enough. Breast cancer is the most common malignancy diagnosed among women and the second leading cause of mortality from cancer. In 1996, the American Cancer Society estimates that 184,300 women will be diagnosed with breast cancer and that 44,300 women will die from this disease.[3] Breast cancer is also a leading cause of premature mortality among women.[4] In fact, the decision to include women age 40 and older in the Health Insurance Plan (HIP) of Greater New York randomized trial of breast cancer screening was based on the observation that women diagnosed with breast cancer between ages 40–49 contributed 34 percent of the total years of potential life lost (YPLL) due to breast cancer.[5] At present, due to trends in aging (in particular, the maturation of the postwar birth cohort), in 1996 more cases of breast cancer will be diagnosed among women ages 40–49 than in women ages 50–59 (33,400 versus 30,900), even though age-specific rates are lower.[6,7]

Beyond disease burden, the disease must also meet certain criteria related to its preclinical phase.[1] First, the preclinical condition should be predictive of a reasonable probability of progression to clinical symptoms if left untreated. Second, the disease should have a detectable, preclinical phase, estimated as the mean sojourn time.[1,8] The sojourn time is the estimated maximum duration of the detectable preclinical phase, and is the basis for establishing screening intervals within which beneficial lead times are attainable.[9] Thus, it is axiomatic that screening intervals be less than the estimated mean sojourn time. Third, the sojourn time must be of sufficient length to ensure a reasonable level of disease prevalence, both for the disease to be detectable *and* to offer the opportunity for detection at a point when medical intervention can make a difference in its natural history. If the sojourn time is short, then there will be poor coincidence between the occasion of screening tests in an asymptomatic population and the ability to detect the disease. It has been

estimated that the mean sojourn time for women ages 40–49 is 1.7 years, and for women ages 50–69, between 3.3 and 3.8 years.[10] Finally, there should be sufficient evidence that treatment for early-stage disease offers significant benefits compared with treatment at a later stage. The benefits of breast cancer treatment at earlier versus later stages is well established, although evidence is stronger for women age 50 and older compared with women ages 40–49.[10–13]

Provided that the disease in question meets the characteristics described above, the test must meet acceptable criteria for accuracy and reliability. In other words, it must do a reasonably good job at correctly distinguishing those who probably have the disease from those who probably do not. The reliability of screening tests is often overlooked among screening test characteristics, but reproducibility of results obviously is an important factor—a test should give the same result if applied repeatedly to an individual with or without the disease.[1] Beyond this, a screening test should meet reasonable performance measures, including the cancer detection rate, sensitivity, specificity, and positive predictive value (PPV). These measures are defined by end results in the context of a breast cancer screening program. By convention, the basic measurements for calculating these outcome measures are as follows: A true-positive (TP) can be defined as breast cancer diagnosed within 1 year after recommendation for a biopsy after an abnormal mammogram. A true-negative (TN) can be defined as no evidence of breast cancer within 1 year of a normal mammogram. A false-negative (FN) can be defined as a cancer diagnosed within 1 year of a normal mammogram. Finally, a false-positive (FP) can be defined several ways, each according to the criteria that there is no evidence of breast cancer within 1 year after the definition of a positive finding, and each relevant to the focus of evaluation in a screening program. FP_1 is based on cases recalled for additional imaging evaluation after an abnormal screening mammogram; FP_2 is based on cases referred for biopsy or surgical consultation after an abnormal mammogram; and FP_3 is based is based on cases that have undergone biopsy after an abnormal mammogram. Each false- positive measurement in turn represents additional progression into the diagnostic process.[14]

Sensitivity is a measure of the probability of detecting a cancer when a cancer exists, or the proportion of patients found to have cancer within 1 year of screening who were identified as having an abnormality at the time of screening. Sensitivity is estimated by TP/ (TP + FN). Specificity is a measure of the probability of correctly identifying an individual as not having cancer when no cancer exists, or the proportion of patients found to not have cancer within 1 year of a normal screening examination. Specificity is estimated as TN/(TN + FP). The PPV varies according to the definition of a false-positive result, and is the proportion of cases correctly identified as having cancer among all cases identified as positive according to the three definitions listed above.[14]

The goal of a screening program is to achieve uniformly high sensitivity and specificity, and the relative importance of accuracy for either of these measures is a function of the severity of an error, for both the individual and the cost of the screening program. From a measurement standpoint, the sensitivity and specificity of mammography are influenced by several factors, including quality control of the screening tests, interpretation thresholds, and screening interval. Thus any assessment of existing estimates must consider the characteristics of the screening program from which it is derived.[15] For this reason, constant monitoring of the performance of a screening program is essential to determine those dimensions of sensitivity and specificity inherent in the interplay between the disease and the technology at hand, and that which may be influenced by improvements in technique and operation.

References

1. Morrison AS. Screening in chronic disease. New York: Oxford University Press, 1992.

2. Cole P, Morrison AS. Basic issues in population screening for cancer. J Natl Cancer Inst 1980; 64:1263–72.

3. American Cancer Society. Facts and figures. Atlanta: American Cancer Society, 1996.

4. Centers for Disease Control. Premature mortality due to breast cancer–United States, 1984. Morb Mortal Wkly Rep 1987;36(4):736–9.

5. Shapiro S, Venet W, Strax P, Venet L. Periodic screening for breast cancer: The Health Insurance Plan Project and its sequelae, 1963–1986. Baltimore: Johns Hopkins Press, 1988.

6. American Cancer Society. Breast cancer facts and figures, 1996. Atlanta: American Cancer Society, 1996.

7. Smith RA. Epidemiology of breast cancer. In: Kopans DB, Mendelson EB, eds. Syllabus: A categorical course in breast imaging. Chicago: Radiological Society of North America, 1995.

8. Walter SD, Day NE. Estimation of the duration of the preclinical state using data. Am J Epidemiol 1983;118:865–86.

9. Duffy SW, Chen HH, Tabar L, Day NE. Estimation of mean sojourn time in breast cancer screening using a markov chain model of both entry to and exit from the preclinical detectable phase. Stat Med 1995;14:1531–43.

10. Tabar L, Fagerberg G, Chen RH, Duffy SW, Smart CR, Gad A, Smith RA. Efficacy of breast screening by age. New results from the Swedish two county trial. Cancer 1995;75:2412–9.

11. Hurley SF, Kaldor JM. The benefits and risks of mammographic screening for breast cancer. Epidemiol Rev 1992;14:101–30.

12. Nystrom L, Rutqvist LE, Wall S, Lindgren A, Lindqvist M, Ryden S, Anderson I, Bjurstam N, Fagerberg G, Frisell J, Tabar L, Larsson LG. Breast cancer screening with mammography: overview of Swedish randomized trials. Lancet 1993;341:973–8.

13. Smart CR, Hendrick RE, Rutledge JH, Smith RA. Benefit of mammography screening in women aged 40–49: current evidence from randomized controlled trials. Cancer 1995;75:1619–26.

14. Bassett LW, Hendrick RE, Bassford TL, et al. Quality determinants of mammography. Clinical Practice Guideline No. 13. AHCPR Publication No. 95–0632. Rockville, MD: Agency for Health Care Policy and Research, Public Health Service, U.S. Department of Health and Human Services, 1994.

15. Chen HH, Duffy SW. A Markov chain method to estimate the tumor progression rate from preclinical to clinical phase, sensitivity and positive predictive value for mammography in breast cancer screening. The Statistician 1996;45:1–11.

Basic Designs of Randomized Clinical Trials of Screening

Freda E. Alexander, M.D.

To avoid lead time bias, length-biased sampling and selection bias evaluation of population screening (e.g., for breast cancer) must be based on randomized clinical trials (RCTs). The design basics of such trials are well established. A study population is identified and randomized into two arms, which either receive (the intervention arm) or do not receive (the control arm) an offer of screening under the protocol to be evaluated. All other health care and all therapy for breast cancer should be independent of the study arm. The entire study population is followed for a (usually lengthy) time period, after which disease-specific mortality in the two trial arms from the date of randomization to the end of followup is compared. Reduced mortality in the arm randomized to receive an offer of screening is evidence of the beneficial effect of screening.

A number of features of this basic design deserve attention.

Identification of the Study Population

Women who have already been diagnosed with breast cancer cannot benefit from screening and are invariably excluded from the study population, although identifying these ineligible women is not always straightforward. Two general choices of study population have been used: first, a geographical population, or one that is representative of this; and second, a volunteer group. The use of a volunteer group has been rare in trials of breast cancer screening, but is now frequently used for trials of other cancer screening. The main advantage is that acceptance rates in the intervention arm will be higher. Disadvantages (see below) include an increased possibility of contamination in the control arm, difficulties in ensuring that randomization is blind, and requirements for some minimal screening of the entire study population.

A further problem is that the results are not necessarily generalizable. This, in fact, may arise even for the first choice of study population. These results will apply to the general population at the time and place of study, but need not extend to other times and places. In particular, changes with time, or by country, of the underlying breast cancer incidence, stage at presentation, or survival rates mean that results of trials cannot necessarily be generalized.

Randomization

Individual randomization is the ideal, but logistical and ethical issues arise when large populations of healthy individuals are involved. Many trials of breast cancer screening have used cluster randomization. The effect can be to reduce the efficiency of the randomization. A basic requirement of RCTs of therapy is that randomization be blind; that is, that allocation to trial arm should be conducted without knowledge of the clinical status of the individual. For screening trials with the first choice of study population, this presents no problem; but when the study population consists of volunteers who may have had clinical examination prior to consent and prior to randomization, it is essential that blindness is seen to be achieved.

Contamination, Compliance, and Prescreening

This benefit of screening can only apply to women who are screened when compared with those who are not. The requirements of the RCT mean that analyses are conducted on an “intention-to-

treat" basis. Maximum effect would be seen if all women in the intervention arm and none in the control arm were screened. This would, in fact, provide an accurate estimate of the benefit of screening. The observed differences between the two arms of the trial will give diluted effect estimates (accompanied by loss of statistical power) if either of the following occur: women in the intervention arm do not accept the offer of screening (low compliance), or women in the control arm find alternative sources of screening (contamination). Quantifying compliance is relatively straightforward, although estimating its impact is difficult. Quantifying contamination is almost impossible.

If all members of the study population are prescreened, the effect will be similar to that of contamination.

Statistical Power and Subgroup Analysis

The statistical power of RCTs of screening, as of therapy, is based on the number of events expected in the two arms of the trial. Since, unlike therapeutic trials, the study population is initially disease free, this leads to a requirement for both very large numbers (25,000–100,000 or more) in the study population and long-term (7 years or more) followup. These numbers are required to provide adequate statistical power to detect an effect; higher numbers are required to provide precise estimates of the effect and to permit adequate power for subgroup analyses (e.g., women aged less than 50 years at entry).

End-points

The end-point of interest is, invariably, disease-specific death (or an estimate of this derived from the use of 'surrogate' or interim end-points). Ascertaining all relevant deaths and validating their status is critical in the design of screening RCTs. It is possible for biases to arise at both points and essential that this be avoided. Women in the intervention arm (especially screen-detected cases) may be more likely to be treated in specific centers, and this has the potential to introduce the confounding of the trial arm with management. This may be unavoidable, but monitoring is mandatory.

Objectives

Finally, there is a tension between two objectives of screening trials. The first objective (as for Phase 2 clinical trials) is to determine whether screening can reduce mortality. This requires optimal or maximal screening (in terms of frequency, number of views, personnel involved, biopsy decisions, etc.). The second objective is to provide information that can be interpreted in terms of disease natural history and cost-effectiveness. These two do not always lead to the same design choices.

Study Design II

Eugenio Paci, M.D.

Introduction

Seven randomized controlled trials (RCTs) have been carried out in the United States, Sweden, Canada, and Scotland to assess the efficacy of breast cancer screening. Some preliminary remarks need to be made before the discussion of specific design issues.

1. The first RCT, the Health Insurance Plan (HIP) of Greater New York study, started in December 1963 to determine "whether periodic breast cancer screening with mammography and clinical examination of the breast holds substantial promise for lowering mortality in the female population from breast cancer."[1] The screening was offered annually to the invited group and ended after four screening rounds. The HIP study was designed to assess the effect of screening independently by the age at entry, but the question of a differential efficacy of screening by age was evident very soon, and the data interpretation was difficult because of small numbers. During the early seventies—the period of the planning phase and start of the Swedish trials—the appreciation of a differential impact of mammography in younger women became increasingly clear. In the Malmo trial (start: 1976) the age at entry was postponed to 45; in the Two-County Study (TCS) the age range was 40–74, but the interscreening interval was shorter for women ages 40–49 at entry (24 months). The Canadian National Breast Screening Study (NBSS-I), which began in 1980, was the only study specifically designed to solve the question posed by the results of the HIP study: the efficacy of screening (mammography and clinical examination) in younger women. The Gothenburg trial, the most recent Swedish trial (start: 1982), was also addressed to study women at younger ages (40–59), and adopted a shorter interval (18 months) than that used in the previous Swedish trials.

The first conclusion is that, although only the NBSS-1 was expressly designed to assess the efficacy at younger ages (40–49), the question of breast cancer screening in younger women arose very soon in the history of the breast cancer screening evaluation and conditioned choices in the study design of the following trials. The choice of 49 as the cutoff point was at the root of the unresolved debate on the age at breast cancer diagnosis. Because the benefit for younger women was evident after several years, the finding was thought to possibly relate to screening after 50 of women enrolled in their late forties. The analysis of the data considering age at diagnosis in the HIP study is available, and the TCS data were recently published. The interpretation of the findings is not straightforward, largely because the open question of the possible interaction between screening and the physiological and reproductive history of the woman. Data on reproductive life and menopausal status over time are not available in these studies. The group 45+, which is about half of the women enrolled in their forties in the RCTs, had the majority of their screening tests around the age of menopause, and there are suggestions today of the possible negative interaction between the woman's individual status and screening history. On this basis the contribution of the RCTs to the understanding of the efficacy of screening in women both in premenopausal and in perimenopausal status has very limited statistical power.

2. In parallel with the evaluation of the efficacy in terms of mortality reduction, the screening process has been studied in terms of the relationship between the biological natural history of the tumor and the screening protocol. In 1968, Hutchison and Shapiro presented data on the "Lead Time Gained by Diagnostic Screening for Breast Cancer."[2] Important statistical modeling has been done since on the relationship between the modification of the natural history of the tumor and the

mortality reduction. Statistical modeling and histopathologic case definition allowed for an in-depth analysis of surrogates as predictors of the mortality reduction in the TCS. The results presented in the last Falun Conference[3] suggest new possibilities of research based on this kind of analysis.

3. The screening test was both physical examination and mammography in the American (HIP and NBSS-I) studies but not in the European studies. The protocol of the American studies considered two-view mammography, whereas the first group of Swedish trials (excluding only the latest one, the Gothenburg trial) was in general based on a single-view mammography. This confirmed the prevalent impression of a more parsimonious attitude in Europe, which has been the occasion of debate also with reference to the propensity to biopsies. The quality of the mammograms has always had great relevance in the design of the studies. Quality has changed very much over the last 20 years, and because of this technological modification the comparison between old and recent trials is difficult. At the same time, the different performance of mammography in younger, perimenopausal, and postmenopausal women has been documented and studied. The combined impact of mammographic technical problems and breast characteristics in younger women has been considered as a possible contributor to the lower efficacy of screening in younger women. The issue of the radiological quality of the mammograms has been especially debated in relation to the NBSS-I (with many discussions and disagreements within the expert panel). In recent clinical series in Western countries, the proportion of nonpalpable breast cancer cases is growing, and a high proportion of the cases occurring in the control group population is of very small tumors. For this reason, as time passes, the contribution of clinical examination to cancer detection might be less important in these countries.

The estimates of the possible impact on the mortality reduction of the specific design choices (number of mammographic views, interscreening interval, addition of physical examination) have been analyzed by Kerlikowske.[4] However, it is extremely problematic to estimate the confounding effect due to the changing quality of mammography over time. Looking at performance indicators such as interval cancer rates, detection rates, and advanced or DCIS cancer rates instead of mortality, this analysis might be possible but in any case extremely difficult because of the huge variability of the data characteristics between studies.

The Screening Protocol

Three basic elements are especially relevant in the trial design of the breast cancer screening protocol:

1. Screening test
   1.1 physical examination
   1.2 mammographic modalities (number of views, quality control)

2. Interscreening interval

3. Number of rounds

Other issues of relevant interest are:

4. Case definition

5. Assessment of positive screening results

6. Risk factors

Table 1 shows the characteristics of the protocol adopted in each RCT regarding elements 1, 2, and 3 above:

TABLE 1. Characteristics of Randomized Controlled Trial Protocols

Study	Mammography Views	Clinical Exam	Interval (months)
1963 HIP	2	yes	12
1976 Malmo	2 or 1	no	21
1977 TCS	1	no	24
1979 Edinburgh	2,1	no	24
1980 NBSS-1	2	yes	12
1981 Stockholm	1	no	26
1982 Gothenburg	2	no	18

Note: Swedish data from Rutqvist.[5]

4. All the trials, although addressed to mortality reduction, collected information on the occurrence of breast cancer cases in the invited-to-screening and the not-invited-to-screening groups. Data are available on the main process indicators considered relevant for the evaluation of screening process. Based on screen-detected, interval, and clinical-detected cancer cases, measures of performance (sensitivity, specificity, and predictive values) can be calculated also using statistical modeling. The pathologic definition of cases (pTNM, Grade) varied in different trials, and data were not collected according to specific protocols. Data on the pathologic characteristics of the cases in the HIP study were available retrospectively. Until now only the TCS has offered material rich enough for an in-depth evaluation of the relationship between pathologic characteristics of tumors and mortality reduction.

5. The proportion of women recalled for assessment because of a positive result at the screening test is a fundamental parameter for evaluation of the human and economic cost of screening. As reported by Rutqvist at Falun[5] the percentage of younger women recalled for assessment was in the order of 4–6 percent in the Swedish trials, and among them from 0.2 to 0.9 percent were referred for biopsy.

6. The question of risk factors for breast cancer has changed in the last few years: New developments in genetics could be of interest for the selection of high-risk groups, especially in younger women. In the RCTs carried out until now, only the HIP and NBSS-1 published results on the risk profile of enrolled women. Other trials (the population-based ones) did not provide information on the risk patterns of the enrolled population.

Comment

The RCTs that have been carried out until now were designed to solve the question of the screening efficacy independently by age, and only the NBSS-I studied specifically younger women. These results provide qualitative evidence of an impact of screening in younger women on long-term mortality reduction. However, the trials, because of their design characteristics, were not able to resolve the dilemma “age at screening/age at diagnosis.” No trial was specifically designed to study the interaction between the woman’s personal history and screening protocol.

In breast cancer research, the issue of lifestyle and reproductive life of as modifiers of the woman's risk pattern has received large attention in recent years, and the influence of these factors on the woman's breast cancer risk is the object of several investigations. So far, breast cancer screening has been studied by looking at the natural history of the disease and not at the personal history of the woman. The possible interaction of the efficacy of the screening procedure with the risk pattern and the changing physiology of the woman could be an issue for a better understanding of the trial findings and for further research. Because of these considerations, new research proposals should consider the follow-up of women starting screening in premenopausal age (40–42) and suggest high-quality, two-view, annual mammograms. This protocol would allow for a better understanding of the modification of the screening performance in perimenopausal age.

References

1. Shapiro S, Venet W, Strax P, Venet L. Periodic screening for breast cancer. The John Hopkins University Press, 1988.

2. Hutchison G, Shapiro S. Lead time gained by diagnostic screening for breast cancer. J Natl Cancer Inst 1968; 41:665–81.

3. Duffy SW, Chen HH. Parameters of screening practice and outcome variation among programmes. Proceedings of the Falun Conference, 1996.

4. Kerlikowske K, Grady D, Rubin SM, Sandrok C, Ernster VL. Efficacy of screening mammography. JAMA 1995; 273:149–54.

5. Rutqvist LE. Summary of the characteristics of the Swedish trials. Proceedings of the Falun Conference, 1996.

Periodic Screening for Breast Cancer: The Health Insurance Plan of Greater New York Randomized Controlled Trial

Sam Shapiro

The Health Insurance Plan (HIP) of Greater New York study began in December 1963 and was the first randomized controlled trial to test whether annual screening for breast cancer led to a reduction in breast cancer mortality.

Highlights of the Study's Design

1. Women ages 40–64 years were included in the trial. Through randomization about 31,000 were allocated to the study group to whom screening was offered and another 31,000 women, the control group, continued to receive usual care.

2. The screening schedule included an initial examination and three reexaminations at annual intervals for those who had the base line examination; 67 percent appeared for the base line; response rates among them in future examinations were high; women ages 40–49 had the highest response rate, about 69 percent.

3. Each examination consisted of film mammography (cephalocaudal and lateral views of each breast) and a clinical breast examination (CBE) by a physician, usually a surgeon; mammography and CBE were conducted independently.

4. Overlapping sources of information were used to identify breast cancer cases and deaths from all causes, including breast cancer. These included HIP records, Blue Cross files, death record files from New York City and State and other states to which migration may have occurred, the National Death Index, New York State Cancer Registry, periodic surveys of the study and control groups.

Selected Methodological Issues

1. Special attention was directed at whether there were biases in the random allocation that might limit the comparisons between the study and control groups. No differences were found in a survey of personal characteristics; also, the rates of breast cancer at the end of 10 years of followup and mortality from all causes of death except breast cancer were similar in the study and control groups (Table 1). Differences were present between the study women who were screened and those who refused screening; the latter group had a much higher general mortality rate and lower breast cancer incidence rate, indicating the need to combine both groups in making comparisons with the control group.

TABLE 1. Mortality from All Causes Excluding Breast Cancer and Breast Cancer Detection Rates: 10-Year Followup After Entry			
Rate	Intervals from entry		
	10 years	1–5 years	6–10 years
Deaths/10,000 person-years			
Total Study	68.6	56.3	81.4
Screened	56.8	42.9	71.1
Refused screening	93.0	83.7	102.7
Control	68.9	58.2	80.1
Breast cancers/1,000 person-years			
Total Study	2.11	2.05	2.18
Screened	2.24	2.26	2.21
Refused screening	1.86	1.61	2.13
Control	2.09	1.95	2.22

2. The number of breast cancer cases detected was almost equal in the study and control groups at the end of 5 years from entry (i.e., about 1_ years after the last women were screened in their followup examinations); at 5 years there were 304 breast cancers histologically confirmed in the study group and 295 in the controls group; at 6 years the numbers were 367 and 364 breast cancers in the two groups respectively; and at 7 years there were 426 and 439 breast cancers in the two groups. Most of the results of the trial are based on the cases detected within 5 years; very similar results are found when the data include the breast cancers diagnosed in years 6 and 7.

3. Lead time estimates were derived. The study group was compared with the control group with respect to dates of diagnosis of breast cancer. The difference between the mean times to diagnosis for the two groups reflects both the proportion of cases detected at screening and the time interval by which diagnosis was advanced in each case detected at screening. Lead time by age at entry follows (a summary is provided in Table 2):

   Total - 10.4 months (SE - 4.5)
   40–44 - 15.8 months (SE - 22.0)
   45–49 - Indeterminate
   50–54 - 25.2 months (SE - 8.6)
   55–59 - 18.7 months (SE - 8.1)
   60+ - Indeterminate

TABLE 2. Lead Time and Percent Reduction in Breast Cancer Mortality Among Women by Age at Entry (18 years from entry)			
Age at Entry	Lead Time (Months)	Deaths	PYLL
Total	10.4 (4.5)[a]	22.7	25.4
40–49	5.2 (5.9)[a]	24.6	20.3
50–59	21.9 (9.6)[a]	23.0	30.8
60–64	Indet.[b]	16.7	22.3

[a] Standard error due to sampling
[b] No clear evidence of a greater early case finding rate in study subjects than in controls

4. Rules were established for assigning breast cancer as the cause of death. This was done because of the problem in dealing with death certificate information in classifying underlying cause of death for research purposes. Two physicians classified designated cases (reasonably certain and questionable) as to whether breast cancer was the underlying cause; differences of opinion were resolved through consultation.

Results of Screening Trial

1. Table 3 gives the distribution of histologically confirmed breast cancers detected during the first 5 years from entry for the study group by source of diagnosis; also shown is the number of breast cancers in the control group; 74 percent of the cases among the study group of women were diagnosed among those who had been screened at least once; more cases were found through the rescreenings than at the initial examination; about 15 percent were detected in the 12-month interval since their last screening.

2. A higher proportion of breast cancers were detected through the clinical examination than through mammography; this was especially true for the women under 50 years of age (Table 4).

3. Among women ages 40–64 at entry, screening resulted in about a 30 percent reduction in mortality from breast cancer during the first 10 years of followup from entry; by the end of 18 years the reduction was close to 25 percent (Table 5). Figure 1 plots the data for breast cancer deaths among women who had breast cancer in the first 5 years and in the first 7 years after entry. It is clear that the same relationships apply to both sets of curves.

4. A favorable effect of screening appeared appreciably later among women ages 40–49 at entry than among women above this age. At 10 years from entry, mortality differentials between the study and control groups were relatively lower at ages 40–49 than at ages 50–59 but were at a similar level at 18 years of followup (Table 6). The later appearance of a possible reduction in mortality among women ages 40–49 than those ages 50–59 is seen in Figure 2.

TABLE 3. Breast Cancer Cases Histologically Confirmed Study and Control

	Number	Percent	Rate
Study (Total)	304	100.0	2.05
Screened	225	74.0	2.26
Detected on Screening	132	43.4	
Interval	93	30.6	0.92
(<12 months)	(45)	(14.8)	—
(>12 months)	(48)	(15.9)	—
Refused	79	26.0	1.61
Control	295	—	1.95

Note: Case detection first 5 years after entry

TABLE 4. Histologically Confirmed Breast Cancer Cases by Modality of Detection at Screening				
Modality of detection	Total	Age at entry		
		40–49	50–59	60–64
Total	100.0	100.0	100.0	100.0
MM only	33.3	25.0	38.8	32.0
Clinical only	44.7	57.5	40.3	36.0
MM and Clinical	22.0	17.5	20.9	32.0

TABLE 5. Breast Cancer Deaths Among Women Diagnosed in Specified Intervals from Entry, Study and Control Groups			
Year of diagnosis after entry	Interval from entry to breast cancer death		
	5 years	10 years	18 years
1–5			
Study Group	39	95	126
Control Group	63	133	163
Percent Difference	38.1	28.6	22.7
1–7			
Study Group		123	180
Control Group		174	236
Percent Difference		29.3	23.7

Note: In years 1–5, there were 307 breast cancers in the study group and 301 in the control group; in years 1–7, 431 and 448 were in the two groups respectively.

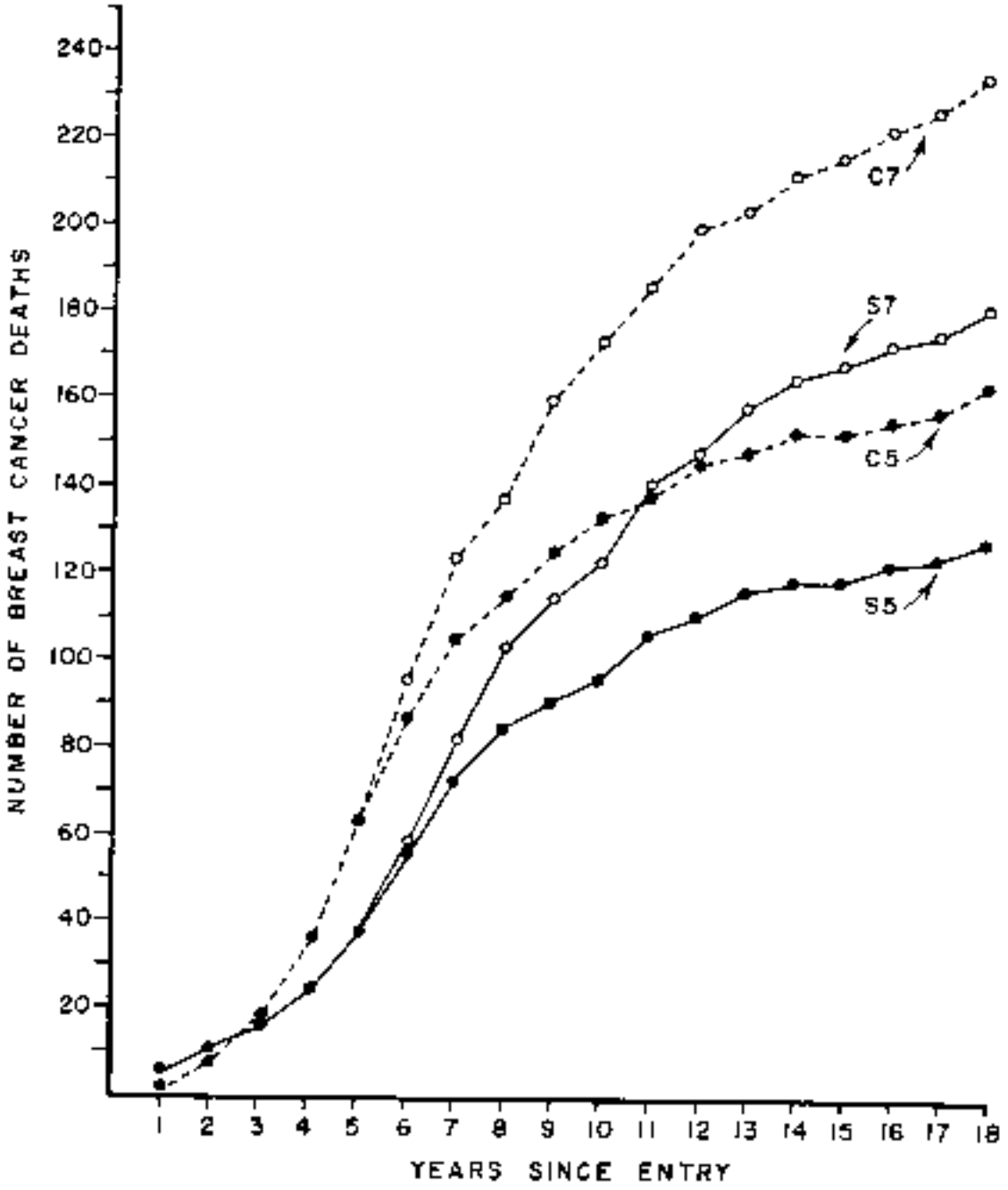

FIGURE 1. Cumulative Number of Deaths Due to Breast Cancer by Interval Since Entry: All Ages, Study and Control Groups (Breast Cancers Diagnosed Within 5 and 7 Years After Entry)

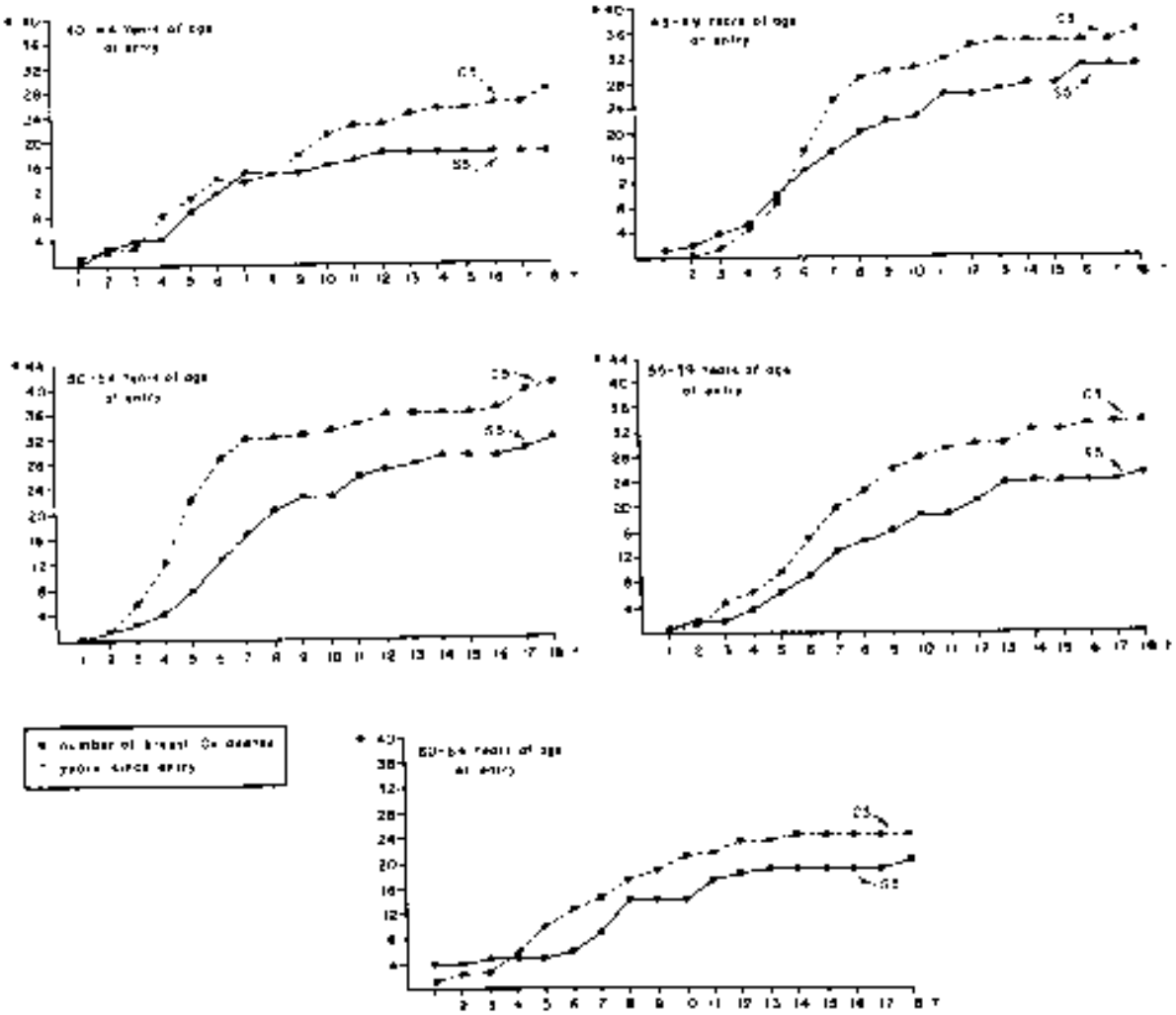

FIGURE 2. Cumulative Number of Deaths Due to Breast Cancer by Interval Since Entry and Age at Entry: Study and Control Groups (Breast Cancers Diagnosed Within 5 Years After Entry)

TABLE 6. Percent Reduction in Breast Cancer Deaths Among Study Group Women by Age at Entry and Selected Intervals After Entry			
Age at entry	Interval from entry to death, year		
	5 years	10 years	18 years
Total	38.1	28.6	22.7
40–49	5.0	23.5	24.6
40–44	a	23.8	35.7
45–49	a	23.3	16.2
50–59	54.5	31.1	23.0
50–54	65.2	30.3	22.0
55–59	a	32.1	24.2
60–64	a	33.3	16.7

[a] Not calculated, small numbers.

5. Much of the gain after 18 years of followup among women 40–49 is due to breast cancer cases detected when these women were 50–54. Limiting the experience to women who were still 40–49 at time of detection reduces the decrease in breast cancer mortality in this age group from 25 to 14 percent (Table 7). Among women 45–49 at entry and at diagnosis more deaths from breast cancer occurred in the study group, 18 versus 13 in the control group.

There are restrictions on drawing hard conclusions from these data but the reduction in the decrease in mortality casts doubt on the ability to conclude from the HIP study that initiation of screening under the age of 50 is efficacious.

TABLE 7. Breast Cancer Deaths Among Women Ages 40–49 at Entry, by Age at Diagnosis and Specified Interval from Entry, Study and Control Groups		
Age at diagnosis in years	Deaths within 18 years from entry	
	Study	Control
40–49[a]	18	28
40–44	7	10
45–49	11	18
45–54[b]	31	37
45–49	18	14
50–54	13	23

Note: Deaths were those due to breast cancer among cases diagnosed within 5 years after entry. Includes deaths among cases histologically confirmed plus deaths among women with breast cancer as the underlying cause but with no histologically confirmed diagnosis before death.

[a] Age at entry, 40–44 years

[b] Age at entry, 45–49 years

References

1. Shapiro S, Venet W, Strax P, et al. Ten-to-Fourteen Year Effect of Screening on Breast Cancer Mortality. 1982. J Natl Cancer Inst 69:349-55.

2. Shapiro S, Venet W, Strax P, et al. Selection, Followup, and Analysis in the Health Insurance Plan Study: A Randomized Trial with Breast Cancer Screening. 1985. In: Selection, Followup, and Analysis in Prospective Studies: A Workshop. Garfinkel L, Ochs O, and Mushinski M, eds. NIH Publication 85-2713; National Cancer Institute Monograph 67. Washington, DC: Department of Health and Human Services, Public Health Service.

3. Shapiro S, Venet W, Strax P, et al. Current Results of the Breast Cancer Screening Randomized Trial: The Health Insurance Plan (HIP) of Greater New York Study. 1988. In: Screening for Breast Cancer. Day N, Miller A, eds. Toronto: Hans Huber.

4. Shapiro S, Venet W, Strax P, et al. Periodic Screening for Breast Cancer: The Health Insurance Plan Project and Its Sequelae, 1963-1986. Baltimore: The Johns Hopkins University Press. 1993.

The Edinburgh Randomized Trials of Breast Cancer Screening

Freda E. Alexander, M.D.

The Edinburgh Randomized Trial of Breast Cancer Screening[1] recruited 44,288 women ages 45–64 between 1978 and 1987 (the initial cohort) and a further 12,142 ages 45–49, between 1982 and 1985 (updates). Randomization was based on 'clusters' defined by primary health care practices. The screening protocol included four biennial mammography examinations and annual clinical examination over the same period (reduced numbers of both for updates). Followup for breast cancer incidence and all-cause mortality is systematically obtained via flagging with the UK National Health Service central registries. Results based on 10 years of followup for the initial cohort and 6–8 years for updates have been published.[2] In the total study, breast cancer mortality was reduced by 18 percent in women offered screening (95% confidence interval, 11 percent increase to 39 percent decrease). For women ages 45–49 at entry, the reduction was 22 percent with a wide 95 percent confidence interval (31 percent increase to 54 percent decrease).

Followup times of 14 years (initial cohort) and 10–12 years (updates) are now available, and results based on these will be reported. In addition, a nonrandomized comparison of breast cancer mortality according to age at first screening (45–53 years) will be presented.

References

1. Roberts MM, Alexander FE, Anderson TJ, et al. The Edinburgh randomized trial of screening for breast cancer: description of method. Br J Cancer 1984;47:1–6.

2. Alexander FE, Anderson RJ, Brown HK, Forrest APM, Hepburn W, Kirkpatrick AE, McDonald C, Muir BB, Prescott RJ, Shepherd SM, Smith A, Warner J. The Edinburgh randomized trial of breast cancer screening: results after 10 years of follow-up. Br J Cancer 1994;70:542–8.

The Canadian National Breast Screening Study: Update on Breast Cancer Mortality

Anthony B. Miller, M.B., F.R.C.P.

The Canadian National Breast Screening Study (CNBSS) is an individually randomized trial designed to evaluate the efficacy of the combination of annual mammography, physical examination of the breasts, and the teaching of breast self-examination in the reduction of mortality from breast cancer in women ages 40–49 years on entry to the study.[1]

Women with no previous history of breast cancer and no mammogram in the previous 12 months were eligible, providing they signed an informed consent form. A total of 50,430 women ages 40–49 years were enrolled from January 1980 through March 1985 from 15 centers across Canada. Randomization was performed by the local coordinators by reference to prearranged lists. Women were randomly assigned either to mammography and physical examination of the breasts (the MP allocation) or to a control group receiving usual care in the context of the Canadian health care system (the UC allocation) after an initial physical examination. In the MP allocation, five annual screenings were offered to the majority of the participants; those enrolled in the last year of recruitment in the individual centers were only offered four annual screens. The participants in the UC group received annual questionnaires over the same time period. Compliance with attendance for rescreening or return of questionnaires was excellent, exceeding 90 percent in both groups. Breast cancer mortality has been ascertained by record linkage to the Canadian National Mortality Data Base (CNMDB), initially to December 31, 1988, and more recently to December 31, 1993. Thus, participants have been followed for a mean of 10.5 years, with a range of 8.75 to 13 years.

In our published 7-year mortality report,[2] we demonstrated that participants were well balanced with respect to age, marital status, number of live births, reproductive status, education, family history of breast cancer, and place of birth. Compared with the age- and sex-matched Canadian population, participants were more likely to be married, have fewer children, have had more education and be in professional occupations, smoke less, and be born in North America. Subsequently, the validity of the randomization process has been challenged,[3] on the basis of an excess of women with breast cancer with four or more nodes involved detected as a result of the initial screening examination in the MP allocation compared with the UC. However, an independent assessment of the validity of our randomization process carried out by Bailar and MacMahon (in preparation) for the National Cancer Institute of Canada, with particular concentration on the centers where the excess appeared to be concentrated, has found no evidence of any deliberate falsification of randomization such to ensure that more women with “advanced” breast cancers were placed in the MP arm. Further, an independent validation of data from the Manitoba screening center has found no evidence of falsification there either.[4]

We address this issue in Table 1. For women to have been deliberately placed in the MP arm, it would have been necessary for the examiner to have identified an abnormality. All women with such suspect abnormalities were referred to the NBSS review clinic to be assessed by the study surgeon. Table 1 demonstrates that such referrals were similar across the allocations within the study centers.

TABLE 1. Women Ages 40–49 Years on Entry with Abnormalities Detected on Clinical Examination at Screen 1 and Referred to Review			
	Screening Center	MP	UC
1.	Mount Sinai Hospital, Toronto	420	488
2.	Saint Sacrement Hospital, Quebec	643	675
3.	Notre Dame Hospital, Montreal	458	469
4.	Henderson General Hospital, Hamilton	218	243
5.	Health Science Centre, Winnipeg	265	231
6.	Cancer Centre, Vancouver	399	371
7.	Ottawa Civic Hospital	85	89
8.	Ottawa General Hospital	73	68
9.	Hotel Dieu Hospital, Montreal	128	179
10.	Halifax General Hospital	237	214
11.	Westminster Hospital, London	167	162
12.	Cross Cancer Institute, Edmonton	129	136
13.	St. Michael's Hospital, Toronto	39	50
14.	Red Deer General Hospital	58	42
15.	Tom Baker Cancer Centre, Calgary	179	184
	Total, all centers	3,498	3,601

Note: MP = Mammography and physical examination of the breasts; UC = Usual care received by the control group (in the context of the Canadian Health Care System)

Variables that became apparent as a result of the screening process (including the process of diagnosis) were called pseudo-variables by Prorok, and are biased. Nodal status is one such variable. Whether a woman with a physical "abnormality" was referred for subsequent diagnosis (and biopsy) was influenced by the availability of mammograms in the MP group and their nonavailability in the UC group. Several women with four or more nodes were probably unrecognized in the UC group, and many were not even recognized subsequently, as their cancers spread, and they were more likely to be treated in centers where careful extensive nodal dissection was not the norm. Some, because their cancers were advanced, may not have had nodal dissection at all, whereas, in the MP group, many of the so-called advanced cancers were small, with limited involvement of the individual nodes, even though four or more were found to be involved.

In the 7-year report,[2] there were 38 deaths from breast cancer in the MP and 28 in the UC allocation. The ratio of the proportion of breast cancer deaths in the MP allocation compared with that in the UC was 1.36 (95% confidence interval 0.84, 2.21). Using the data from the CNMDB to December 31, 1993, the numbers of breast cancer deaths are now 52 in each arm. Although the nonsignificant excess of breast cancer mortality found previously in the MP arm has now disappeared, there is still no evidence of any reduction in breast cancer mortality consequent to the use of mammography in the CNBSS. We can also update the breast cancer mortality findings derived from our routine annual followup of all the breast cancers ascertained in the study, and included in the summary report from the March 1996 meeting in Falun, Sweden, resulting in 78 in the MP arm and 73 in the UC group. These additional deaths known to us are not the final tally for the 1–15-year report currently under way, and we regard them as less valid than the unbiased data derived from the linkage to the CNMDB, as there was a greater ascertainment of cancers with mammography screening. Only when we are able to evaluate the findings from the record linkage to the Canadian National Cancer Registry, currently under way, will we be sure that any bias from this source has been eliminated.

Conclusion

The study is internally valid and there is no evidence of bias in allocation. Screening of women ages 40–49 with yearly mammography and physical examination has had no impact on mortality from breast cancer in a period of observation of 8.75 up to 13 years from entry.

References

1. Miller AB, Howe GR, Wall C. The National Study of Breast Cancer Screening: protocol for a Canadian randomized controlled trial of screening for breast cancer in women. Clin Invest Med 1981;4:227–58.

2. Miller AB, Baines CJ, To T, Wall C, et al. Canadian National Breast Screening Study: 1. Breast cancer detection and death rates among women 40–49 years. Can Med Assoc J 1992;147:1459–76.

3. Tarone RE. The excess of patients with advanced breast cancer in young women screened with mammography in the Canadian National Breast Screening Study. Cancer 1995;75:997–1003.

4. Cohen MA, Kaufert PA, MacWilliam L, Tate RB. Using an alternative data source to examine randomization in the Canadian National Breast Screening Study. J Clin Epidemiol 1996;49:1039–44.

Recent Results from the Swedish Two-County Trial: The Effects of Age, Histological Type, and Mode of Detection

László Tabár, M.D., Gunnar Fagerberg, and Stephen W. Duffy, M.Sc.

In this presentation, data from the Swedish Two-County Trial[1,2] are used to investigate the roles of age and histological type,[3] including malignancy grade, in determining the effect of screening on mortality from breast cancer. Particular attention is paid to explaining differences in results by age and county. The effect on mortality is assessed from a randomized trial perspective, based on intention to treat, that is, the formal mortality comparison is between an uninvited control group and a group invited to screening, regardless of acceptance or otherwise of the screening invitation. It is therefore of interest to consider the contributions to mortality from cancers detected at screening, interval cancers, and cancers in women who were invited but did not attend.

The results to December 1994 show an overall reduction (both counties, all ages) in mortality of 30 percent (RR= 0.70, 95% CI 0.60–0.83). In women ages 40–49 at randomization, the reduction was 12 percent (RR= 0.88, 95% CI 0.58–1.34), and in women ages 50–74 the reduction was 33 percent (RR=0.67, 95% CI 0.56–0.80). Within age groups, differences between the two counties were also observed. In the 40–49 age group, the mortality reduction in W-county was 27 percent, and in E-county there was an increase in mortality of 2 percent. In women ages 50–74 at randomization, the reductions in mortality were similar in the two counties.

The role of detection mode was found to be a contributory factor to the difference in mortality reduction between the two counties in the 40–49 age group. Table 1 shows the numbers of breast cancers and breast cancer deaths by detection mode and county for this age group.

TABLE 1. Breast Cancers and Deaths by County, Ages 40–49

Detection mode	W-county		E-county	
	Deaths/cancers	Fatality (%)	Deaths/cancers	Fatality (%)
Control before screen	15/37	41	23/78	26
Control first screen	1/12	8	3/35	9
Control Total	16/49	33	26/113	23
Prevalence screen	3/19	16	2/20	10
Incidence screen	8/49	16	7/61	11
Interval	11/44	25	11/47	23
Refuser	0/1	0	4/9	44
Before screening	0/1	0	1/5	20
Study group total	22/114	20	25/142	18
Total	38/163	23	51/255	20

The major difference in the study group cancers between the two counties is the four deaths in refusers and the one death in those cancers diagnosed between randomization and commencement of screening in E-county. Without these deaths, the relative risk of 1.02 observed in E-county would have been 0.81. Screening had no opportunity of preventing these five deaths. There were no corresponding deaths in these detection modes in W-county. It has been shown that prediction of the mortality reductions in the two counties from the numbers of tumors by detection mode gives a close approximation to the observed reductions.[4]

If we consider the cumulative mortality over time, the reduction in mortality in both counties begins to appear at around 5 years after randomization in women ages 50–74. In women ages 40–49 in W-county, the benefit appears at about 8 years after randomization and is not observed at all in E-county, partly due to the deaths in refusers and those cancers diagnosed between randomization and screening, and to the low fatality rate in the control group in E-county. The cumulative mortality from ductal grade 3 carcinoma shows a mortality reduction occurring from around five years in both counties in women ages 50–74, and occurring at 8 years after randomization in women ages 40–49 in W-county only. For ductal grade 2, lobular and medullary carcinoma, the reduction in mortality begins to appear at around 8 years in the 50–74 age group and at around nine years in women ages 40–49. Thus, reducing mortality from ductal grade 3 carcinoma means preventing deaths that would have occurred in the next 5 years and onwards, whereas reducing mortality from ductal grade 2, lobular and medullary carcinoma, prevents deaths that would have occurred 8 years or more later. These results indicate that the delayed and lesser effect of screening on breast cancer mortality in women ages 40–49 is at least partly due to a lesser effect on the grade 3 cancers, possibly due to the 2-year interscreening interval being too long.

Our previous research suggests that some tumors dedifferentiate; that is, some tumors that have malignancy grade 1 may if left untreated progress to grade 2 or 3, and that this tendency is more marked in the age group 40–49.[3] Table 2 shows the percentage of grade 3 tumors by size and age group.

TABLE 2. Malignancy Grade of Tumors by Tumor Size and Age

Tumor size (mm)	40–49		50–59		60–69	
	Grade 1–2 (%)	Grade 3 (%)	Grade 1–2 (%)	Grade 3 (%)	Grade 1–2 (%)	Grade 3 (%)
1–9	38 (81)	9 (19)	74 (85)	13 (15)	126 (88)	17 (12)
10–14	56 (70)	24 (30)	102 (74)	36 (26)	138 (76)	44 (24)
15–19	30 (46)	35 (54)	86 (69)	38 (31)	103 (62)	63 (38)
20–29	33 (45)	40 (55)	69 (48)	75 (52)	94 (51)	90 (49)
30+	16 (33)	33 (67)	22 (38)	36 (62)	40 (37)	68 (63)

In the 40–49 age group, more than 50 percent of tumors of size 15 mm or more are of grade 3, whereas in tumors in women age 50 or more, this only occurs for tumors of size 20 mm or more. This suggests a potential for improvement in terms of improving the malignancy grade by early detection in women ages 40–49. This potential, however, cannot be achieved by screening with a 2-year interval.

The delayed effect on mortality in the 40–49 age group is due to later breast cancer deaths, and not to diagnosis after age 50, as has been suggested.[5,6] The deaths and person-years are shown in Table 3, by age at randomization and age at diagnosis. The majority of the tumors were diagnosed before age 50 and most of the mortality benefit is due to this group of cancers. This is in agreement with results from the Health Insurance Plan (HIP) of Greater New York study.[7]

TABLE 3. Mortality Results by Age at Randomization and Age at Diagnosis, Women Ages 40–49			
Age at randomization	Trial Arm	Deaths (person-years)	
		Age at diagnosis 40–49	Age at diagnosis 50+
40–44	Study	16 (73431.3)	0 (59429.1)
40–44	Control	17 (59304.0)	1 (47628.6)
40–44	Relative risk	0.76	—
45–49	Study	14 (24457.5)	15 (106756.5)
45–49	Control	9 (18886.4)	12 (81920.6)
45–49	Relative risk	1.20	0.96

References

1. Tabár 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.

2. Tabár L, Fagerberg G, Chen HH, et al. Efficacy of breast cancer screening by age: New results from the Swedish Two-County Trial. Cancer 1995; 75:2507–17.

3. Tabár L, Fagerberg G, Chen HH, Duffy SW, Gad A. Tumour development, histology and grade of breast cancers: Prognosis and progression. Int J Cancer 1996; 66:413–9.

4. Tabár L, Fagerberg G, Chen HH, Duffy SW. Screening for breast cancer in women aged under 50: mode of detection, incidence, fatality and histology. J Med Screening 1995;2:94–8.

5. de Koning HJ, Boer R, Warmerdam PG, Beemsterboer PMM, van der Maas PJ. Quantitative interpretation of age-specific mortality reductions from the Swedish breast cancer screening trials. J Natl Cancer Inst 1995;87:1217–23.

6. Tabár L, Duffy SW, Chen HH. Re: quantitative interpretation of age-specific mortality reductions from the Swedish breast cancer-screening trials. J Natl Cancer Inst 1996;88:52–3.

7. Shapiro S, Venet W, Strax P, Venet L. Periodic screening for breast cancer: The health insurance plan project and its sequelae, 1963–1986. Baltimore: Johns Hopkins University Press, 1988.

The Malmö Mammographic Screening Trial: Update on Results and a Harm–Benefit Analysis

Ingvar Andersson, M.D., Ph.D.

The Malmö Mammographic Screening Trial (MMST) was initiated in the fall of 1996. The initial cohort (MMST I) comprised approximately 21,000 women born in 1908 through 1932 (aged approximately 45–69 at entry). After the initial cohort, another cohort of about 17,000 women was entered into the study (MMST II). These women were aged 45–48 at entry. They were entered into the study between 1978 and 1990.

The MMST is population based. The randomization was done on an individual basis. Women were invited birth year by birth year. The control group women were not contacted.

The Malmö population is a purely urban population. The attendance rate in MMST has been relatively low compared with other Swedish studies (74 percent in the first round, approximately 70 percent in the subsequent rounds). The attendance has been higher in younger than in older women (in the first round, about 80 percent among women aged 45–54 and 70 percent among women aged 54–69).

The nonattenders represent a complex group. Our experience is, however, that, on average, this group has presented with more advanced breast cancers and a worse prognosis than the control group.

It is also the experience in the MMST that the interval cancer group had a greater risk of dying from breast cancer than the control group cancers (relative risk about 2.3). The rate of interval cancers has been relatively higher in the younger age group (below 55) than in the older, and the proportion of interval cancer deaths has been higher in the younger than in the older age group.

Mortality

The initial study cohort (MMST I) is part of the Swedish overview. The breast cancer mortality in the MMST I will be presented in the overview of the Swedish studies. New preliminary mortality data from MMST II will be presented. The mortality is determined by data linkage to the Swedish Cause of Death Registry.

The experience so far clearly suggests a reduction of breast cancer mortality among women younger than age 50 years at entry. Screening for breast cancer is, however, a complex intervention, and benefits have to be weighed against adverse effects.

Harm–Benefit Analysis

A harm–benefit analysis has been performed. This was an attempt to quantify in absolute terms some positive and negative effects of screening (Table 1). The number of prevented deaths per 100,000 woman-years has been calculated from the Gothenburg data. It was assumed that the number of women

TABLE 1. Effects of Screening Women Ages 40–49	
Positive effects/100,000 WY	Negative effects/100,000 WY
Prevented deaths 10	False positives 1,250
Prevented metastatic disease10	Surgery, benign disease 56
Conservative surgery36	Clinically insignificant cancer 10
	Radiation-induced breast cancer death1

who did not have metastatic disease equaled the number of prevented deaths. Also, the Gothenburg data were used to calculate the rate of false positives, i.e., the number of women who had to be recalled for work-up. In the majority of cases, the work-up was limited to additional mammographic views, after which the suspicion of malignancy could be dismissed. Only a minority, 56 per 100,000 woman-years, had to undergo a surgical biopsy.

The number of women undergoing conservative surgery as a result of screening as well as the number of women with clinically insignificant cancer were calculated from MMST II data. In this context, clinically insignificant cancer denotes cancer that in the absence of screening would either not have been detected or if detected later would not metastasize. The rate of DCIS and highly differentiated tubular carcinoma in the invited and control groups was calculated. Fifty percent of the difference was arbitrarily chosen as an estimate of the rate of clinically insignificant cancer detected as a result of screening.

Although somewhat hypothetical at very low doses, the risk of inducing breast cancer by radiation can not be completely ignored. The rate of radiation-induced breast cancer death was calculated assuming biannual, two-view mammography, an average glandular dose of 2 mGy per view, a linear dose–response relationship, and an age-dependent risk.

The monetary cost and the lead time years have not been quantified in this analysis.

Conclusions

1. Breast cancer mortality can be reduced in the under 50 age group by mammographic screening.

2. General screening recommendations should take into consideration advantages as well as disadvantages of the intervention.

References

Andersson I, Aspegren K, Janzon L, Landberg T, Lindholm K, Linell F, et al. Mammographic screening and mortality from breast cancer: The Malmö mammographic screening trial. BMJ 1988;297:943–8.

Nyström L, Rutqvist LE, Wall S, Lindgren A, Lindqvist M, Rydén S, Andersson I, Bjurstam N, Fagerberg G, Fresell J, Tabár L, Larsson LG. Breast cancer screening with mammography: overview of Swedish randomized trials. Lancet 1993;341, 8851:973–8.

The Stockholm Mammographic Screening Trial: Risks and Benefits

Jan Frisell, M.D., and Elisabet Lidbrink, M.D.

Background

Results from several randomized mammography screening trials have shown that it is possible to reduce mortality by early discovery of breast cancer in a mammographic screening program, at least for women over 50 years of age.

Purpose

To present data on mortality in breast cancer in study and control groups of the Stockholm trial after 11.4 years of follow-up, and also to present some of the side effects with mammographic screening in the 40–49 age group.

Methods

The Stockholm mammographic screening trial started in March 1981; 40, 318 women ages 40–64 entered a randomized trial of breast cancer screening by single-view mammography alone versus no intervention in a control group of 20,000 women. Attendance was 81 percent in the first two screening rounds, and the attendance rate was equal in all age groups. Two screening rounds were performed, and the first and second screening intervals were 28 and 24 months, respectively. During 1986, the control group was invited once to screening and the study was ended.

Results

Mortality

In 40–49 age group, 118 and 59 cases of breast cancer were diagnosed in the study and control groups, respectively. After a mean follow-up of 11.4 years, the number of breast cancer deaths in the respective groups was 24 and 12. The relative risk of death in breast cancer was 1.08 (95% confidence interval 0.54–2.17). No mortality reduction was seen in this age group, in contrast to a significant mortality reduction among women over age 50 years. The breakpoint for benefit in this study seemed to be at 50 years, but this tendency is uncertain because of the low statistical power in the analyses of small subgroups. Updated results from the Swedish overview which include the Stockholm trial have shown a 24-percent mortality reduction in the 40–49 age group, near significance.

False-positives and costs

The recall rate for clinical examination, fine-needle biopsy, and additional X-rays after a complete mammography was 0.8 percent in the Stockholm study. The recall rate was equal in the 40–49 and 50–59 age groups. The number of false-positives was higher in the 40–49 age group, and number of cancers found in this age group was lower than that for women over 50 years. The examinations of the false-positives generate both psychological discomfort and large costs. Forty-one percent and 45 percent of the costs of the false-positives in the first and second round, respectively, resulted from examining women who were under age 50 when they entered the trial.

Interval cancers

The incidence of interval cancers in the Stockholm study was 1.8 and 2.0 breast cancers/1,000 women/24 months in the first and second interval, respectively. There was a significant larger number of younger women, ages 40–49, in the group of interval cancers and the mortality was higher than that for women over age 50 years. In the total study there was a significantly better survival among the interval cancers compared with the control cancers. The subgroup analysis showed that this better survival was seen only among women over 50 years, in the 40–49 age group survival was equal to the control group. The mortality of younger women was dominant among the interval cancers.

Summary

In the Stockholm mammographic screening trial, a significant reduction in breast cancer mortality was found in women over age 50 years, but not in the 40–49 age group. It seems to be a breakpoint of benefit in this study at 50 years, but this tendency in uncertain because of low statistical power in the subgroup analysis. Large overview studies are needed to answer the question of whether mammography screening reduces mortality in the 40–49 age group. The side effects of mammography screening among younger women such as false-positive cases, costs, nonattendance, and mortality of interval cancers need further studies.

References

1. Shapiro S, Strax P, Venet L. Periodic breast cancer screening in reducing mortality from breast cancer. JAMA 1971;215:1777–85.

2. Tabár L, Fagerberg G, Gad A, et al. Reduction in mortality from breast cancer after mass screening with mammography. Lancet 1985;1:829–32.

3. Nyström L, Rutqvist LE, Wall S, et al. Breast cancer screening with mammography: Overview of Swedish randomized trials. Lancet 1993;341:973–8.

4. Frisell J, Glas U, Hellström L, Somell A. Randomized mammographic screening for breast cancer in Stockholm. Design, first rounds results, and comparisons. Breast Cancer Res Treat 1986;8:45–54.

5. Frisell J, Eklund G, Hellström L, et al. Randomized study of mammographic screening - preliminary report on mortality in the Stockholm trial. Breast Cancer Res Treat 1991;18:49–56.

6. Lindbrink E, Elfving J, Frisell J, Jonsson E. Neglected aspects of false-positive findings in mammographic screening. BMJ 1996;312:273–6.

7. Lidbrink E, Levi L, Pettersson J, et al. Single-view screening mammography: psychological, endocrine and immunological effects of recalling for a complete three-view examination. Eur J Cancer 1995;31A(6):932–3.

8. Frisell J, von Rosen A, Wiege et al. Interval cancer and survival in a randomized breast cancer screening trial in Stockholm. Breast Cancer Res Treat 1992;24:11–6.

The Gothenburg Breast Screening Trial: Results from 11 Years Followup

Nils Bjurstam, M.D., Ph.D.; Lena Björneld; and Stephen W. Duffy, M.Sc.

In 1983 and 1984, the female population of Gothenburg ages 39–59 was randomized to invitation to screening every 18 months (study group) or to no invitation (control group). In order to deliver screening every 18 months with the resources available, the study group had to be no larger than 21,000, and therefore the population of 51,611 women with no previous breast cancer was randomized in the approximate ratio of 1.4:1. The study group comprised 21,650 women and the control group 29,961 women.

Screening was by two-view, single-read mammography.[1] Attendance rate at first screening round in the study group was 83 percent. After five rounds of screening in the women ages 39–49 at randomization, and after three rounds in women ages 50–59, the control group was screened. In this presentation we report on mortality in women ages 39–49 at randomization from breast cancers diagnosed in both groups up to immediately after the first screen of the control group, with followup to 12 years after the start of the trial.

Breast cancer deaths were those deaths classed as having breast cancer present in the Swedish Cause of Death Register. Breast cancer mortality was compared between the study and control group using Poisson regression.[2]

Number of subjects in the 39–49 age group and breast cancer cases are shown by trial arm in Table 1. The corresponding breast cancer deaths and person-years are given in Table 2, with the relative risk (RR) of breast cancer death, for the study compared with that for the control group. The mortality reduction was 44 percent (RR=0.56, 95% CI 0.32–0.98).

TABLE 1. Numbers of Subjects and Breast Cancer Cases, Women Ages 39–49 at Randomization, by Trial Arm

Trial arm	Breast cancer cases	Number of subjects
Study	148	11,724
Control	196	14,217

TABLE 2. Breast Cancer Mortality, in Women Ages 39–49 at Randomization, by Trial Arm

Trial arm	Breast cancer deaths	Person-years	RR	95% CI
Study	18	138402	0.56	0.32–0.98
Control	39	168025	1.00	–

The mortality reduction is surprisingly large for women under age 50, a population in which screening for breast cancer is known to be difficult.[1] It should be noted that the mortality reduction is consistent with the reduction in advanced tumors. In this age group there were 39 node-positive tumors in the study group and 74 in the control group, giving a relative risk of node-positive breast cancer of 0.64 (95% CI 0.43–0.95).

Table 3 shows the cancers diagnosed by detection mode, with the breast cancer deaths and fatality rates. As one would expect, very low fatality rates were observed for screen-detected cancers. Interval cancers and cancers in nonattenders were associated with fatality rates slightly higher than

those in the control group.

TABLE 3. Breast Cancers and Breast Cancer Deaths by Detection Mode, Women Ages 39–49 at Randomization

Detection mode	Number of DCIS cases	Number of invasive cancers	Breast cancer deaths	Fatality (%)
First screen	10	9	1	5
Subsequent screen	11	56	1	2
Interval cancer	2	33	10	29
Nonattenders	2	25	6	22
Total study group	25	123	18	12
Control	14	182	39	20

Note: DCIS = ductal carcinoma ***in situ***

In terms of age at diagnosis, of the 57 breast cancer deaths in the 39–49 age group at randomization, 37 (65 percent) were ages 39–49 at diagnosis. For the study group, the figures are 13 of 18 ages 39–49 at diagnosis (72 percent), and for the control group 24 of 39 (62 percent).

References

1. Nystrom L, Rutqvist L-E, Wall S, et al. Breast cancer screening with mammography: overview of Swedish randomized trials. Lancet 1993;341:973–8.

2. Breslow NE, Day NE. Statistical methods in cancer research II: The design and analysis of cohort studies. Lyon: International Agency for Research on Cancer, 1987.

Update of the Overview of the Swedish Randomized Trials on Breast Cancer Screening with Mammography

Lennarth Nyström, Stig Wall, LarsErik Rutqvist, Ingvar Andersson, Nils Bjurstam, Gunnar Fagerberg, Jan Frisell, László Tabár, Lars-Gunnar Larsson

Introduction

The aim of the present study was to update the four randomized controlled Swedish trials on breast cancer screening with mammography performed in Malmö, Kopparberg/Östergötland (Two-county trial), Stockholm, and Gothenburg adding 4 more years of followup to gain more precision in the estimates and to be able to study the long-term effects of breast cancer screening with mammography. The analysis will focus on the age group 40–49 years at randomization.

Material and Methods

Study Design

The basic characteristics of the four randomized trials on mammography screening in Sweden have been extensively presented before, and a summary was presented in our first report from the overview.[1]

Invited Women

Initially each screening center sent a magnetic tape containing information concerning their cohort to the administrative center of this study at the Department of Epidemiology and Public Health in Umeå. The cohorts were merged and linked to the six Regional Cancer Registers to identify cases with breast cancer diagnosed between 1958 and 1993, the Swedish Cause of Death Register to identify women who died between 1951 and 1993 and the cause of death according to Statistics Sweden.

Exclusion Criteria

All analyses were based on exact age at randomization, in spite of the fact that most trials for practical reasons have used year of birth cohorts. Thus, 5,143 women ages 38–39 years at randomization were excluded (Kopparberg= 1,148, Östergötland=1,296, Stockholm=680 and Gothenburg=2,019) as we focused on the age group 40–49 years at randomization.

Since invitation to breast screening cannot be expected to essentially influence the clinical course of the disease among those who already have breast cancer, women with breast cancer diagnosed before the date of randomization, according to the Swedish Cancer Register, were excluded from the cohorts (invited group (IG) = 272; control group (CG) = 256).

Determination of Cause of Death

In the overview[1] cause of death determination was made by an independent end point committee (EPC) consisting of four physicians who blindly reviewed medical records, autopsy protocols, cause of death certificates, and histopathology reports for all deceased breast cancer cases, that is, cases

reported to the Cancer Register with breast cancer (ICD=174) after randomization and deceased before date of followup. Later the relative risk (RR) estimates according to the EPC were compared to the RR according to the Cause of Death Register at Statistics Sweden.[2] The RR for breast cancer as the underlying cause of death in the age groups 40–49, 50–59, 60–69, 70–74, and 40–74, respectively, was according to the EPC/Statistics Sweden 0.90/0.95, 0.72/0.76, 0.69/0.69, 0.98/1.05, and 0.77/0.80. Thus RRs determined by these methods were very similar, however, with a slight tendency toward higher values when Statistics Sweden was used. As cause of death determination according to Statistics Sweden at least must be regarded as a conservative estimate of the effect of screening, we decided to use it in the present study.

Models for Analysis

In four of the five screening centers women included in the CG were later also invited to screening. Therefore two different models were used for evaluation—the "followup" model and the "evaluation" model.[1] The former model included all breast cancer deaths that occurred among women with a primary diagnosis after the date of randomization, and before a common fixed study end-point at December 31, 1993. The latter model ignored breast cancer deaths among women whose primary tumor according to the Cancer Register was diagnosed after completion of the first screening round for the control group. In the followup until 1989 the "followup" model and the "evaluation" model showed, as expected, very similar results. In the followup until 1993 only the evaluation model was used, as the duration from the date for the completion of the first screening round and the date of followup had increased considerably.

Statistical Methods

Statistical and epidemiological data analyses have been performed using the QUEST software program.[3] Relative risks (RR) have been calculated using the density method, where the person-time experience of the cohort by time interval of followup is used as a basis for estimating the mortality rates in breast cancer. Weighted RRs and confidence intervals (CI) have been calculated using Mantel-Haenszel procedures.

Results

The number of women by age at randomization and screening center is presented in Table 1. The median followup time was 12.8 years varying from 15.5 (Malmö) to 9.9 (Gothenburg) (Table 2).

TABLE 1. Number of Women Ages 40–49 Years at Randomization in the Invited (IG) and Control Groups (CG) by Screening Center

Screening Center	Age at Randomization					
	40–44		45–49		40–49	
	IG	CG	IG	CG	IG	CG
Malmö			3,945	4,017	3,945	4,017
Kopparberg	4,595	2,478	5,055	2,531	9,650	5,009
Östergötland	5,157	5,337	5,062	5,074	10,240	10,411
Stockholm	7,517	4,495	6,668	3,470	14,185	7,985
Gothenburg	5,664	7,106	5,157	5,995	10,821	13,101
Overview	22,954	129,416	25,887	21,087	48,841	40,503

TABLE 2. Median Followup Time (Years) from Date of Randomization to Date of Followup (12/31/93) by Screening Center and Age at Randomization

Screening Center	Followup Time (years)	
	Median	Lower-Upper Limit
Malmö	15.5	15.3–16.8
Kopparberg	15.2	13.9–16.5
Östergötland	14.2	12.8–15.6
Stockholm	11.9	10.6–12.8
Gothenburg	9.9	9.7–10.3
Overview	12.8	9.7–16.8

The women in the IG and CG were followed for 616,264 and 506,358 person-years, respectively. During that period 104 and 111 breast cancer deaths respectively occurred corresponding to a mortality reduction of 23 percent (RR=0.77; 95% CI: 0.59–1.01). The effect was similar in the age cohorts 40–44 and 45–49 years at randomization, 26 percent and 21 percent, respectively.

The reduction of the breast cancer mortality could also be illustrated by looking at the curves for the cumulative mortality/100,000 person-years by time since randomization. For the age group 40–49 years at randomization the curves start to diverge after about 6 years and continue to diverge at 15 years of followup.

TABLE 3. Number of 1,000 Person-Years and Number of Cases with Breast Cancer as the Underlying Cause of Death According to Statistics Sweden in Invited (IG) and Control Groups (CG) by Screening Center and Age at Randomization. Relative Risks (RR) and 95 % Confidence Limits (CI)

Screening Center	No. of 1,000 Person-Years		No. of Deaths		RR	95% CI
	IG	CG	IG	CG		
Malmö	61	62	15	23	0.67	0.35–1.27
Kopparberg	144	75	23	18	0.67	0.37–1.22
Östergötland	143	147	27	27	1.02	0.59–1.77
Stockholm	162	94	23	10	1.34	0.64–2.80
Gothenburg	106	129	16	33	0.59	0.33–1.06
Overview						
40–44	283	235	39	44	0.74	0.48–1.14
45–49	334	272	65	67	0.79	0.56–1.11
40–49	616	506	104	111	0.77	0.59–1.01

Discussion

A strength of the trials is that it has been shown, based on the followup until 1989,[4] that the cause of death pattern in the IG was, except for breast cancer, very similar to that in the CG showing that the groups were comparable. Besides that, the total mortality including the breast cancer mortality, was in the CG almost identical to that of Swedish women in general. The same was true, with the exception of breast cancer, for the IG. This confirms that the trial cohorts are representative of Swedish women indicating that the quantitative results from these trials may safely be generalized to the Swedish population.

An alternative approach to estimate the effect of mammography screening was recently applied on the followup data until 1989.[5] By applying an indirect method for estimation of the breast cancer related mortality in the breast cancer subcohorts using official national cause of death statistics according to Statistics Sweden as a reference, estimates very similar to the traditional comparison of the breast cancer mortality in the IG and CG were received. This analysis further strengthens the previous report[1] of a beneficial effect of mammography screening.

The extended followup of the cohort until 1993 raises the question of whether it will be possible to show a statistically significant effect of screening in the age group 40–49 years when the younger trials in Gothenburg and Stockholm have been followed for 15 years. However, following a cohort ages 40–49 years at entry for 15 years one can question how much of the effect of screening can be referred to screening before age 50. However, Tabár et al.[6] found in their analysis of the 40–49 year age group in the Two-County trial that "For cancers diagnosed before age 50 years, the relative mortality adjusted for age at randomization and county, was 0.85. For cancers diagnosed after age 50 years, the relative mortality was 0.95."

To conclude, this followup of the four randomized controlled trials on mammography screening

for breast cancer in Sweden until 1993 indicates a possible effect in women ages 40–49 years at randomization. The question of whether this effect is due to screening after the age of 50 requires further study.

References

1. Nyström L, Rutqvist LE, Wall S, Lindgren A, Lindqvist M, Rydén S, Andersson I, Bjurstam N, Fagerberg G, Frisell J, Tabár L, Larsson L-G. Breast cancer screening with mammography: overview of Swedish randomized studies. Lancet 1993;341:973–8.

2. Nyström L, Larsson L-G, Rutqvist LE, Lindgren A, Lindqvist M, Rydém S, Andersson I, Bjurstam N, Fagerberg G, Frisell J, Tabár L. Determination of cause of death among breast cancer cases in the Swedish randomized mammography screening trials. A comparison between official statistics and validation by an endpoint committee. Acta Onc 1995;34:145–32.

3. Gustafsson L. QUEST. A program system for statistical and epidemiological data analysis. Umeå University, Umeå, Sweden. January 1990.

4. Nyström L, Larsson L-G, Wall S, Rutqvist LE, Andersson I, Bjurstam N, Fagerberg G, Frisell J, Tabár L. The overview of the Swedish randomized mammography trials. The total mortality pattern and the representativity of the study cohorts. J Med Screen 1996;3;85–7.

5. Larsson L-G, Nyström L, Wall S, Rutqvist LE, Andersson I, Bjurstam N, Fagerberg G, Frisell J, Tabár L. The Swedish randomized mammography screening trials. An analysis of their effect on the breast cancer related excess mortality. J Med Screen 1996;3:129–32.

6. Tabár L, Duffy SW, Chen H-H. Re: Quantitative interpretation of age-specific mortality reductions from the Swedish breast cancer screening trials. J Natl Cancer Inst 1996;88:52–3.

Variation in the Effect of Breast Screening by Year of Followup

Brian Cox, M.D., Ph.D., F.A.F.P.H.M.

Background

Several meta-analyses and an overview of the effect of breast screening have been undertaken. Not all have specifically accounted for the variation in effect by length of followup despite apparent differences between younger and older women. If all studies are combined regardless of the length of followup, those studies with the longest followup will have an undue effect on the overall measure of effect in any meta-analysis.

Purpose

To assess the effectiveness of mammography screening by year from the start of screening for breast cancer. Published results of all randomized controlled trials (RCTs) were used to conduct a meta-analysis assessing breast cancer mortality for each year after the start of screening for women offered screening compared with women not offered screening. Because only RCTs can be sufficiently unaffected by the important biases of lead time, length bias, and selection bias in assessing the effects of screening, they were the only studies included. This analysis is an extension of a previous analysis[1] and includes data from the Gothenburg trial[2,3] and some earlier unpublished data from the Edinburgh and Canadian trials.

Methods

The results of mammography screening of several types and frequency with or without physical examination or breast self-examination were combined. A meta-analysis combining data for each individual year of followup was undertaken.

The data allowed the calculation of combined breast cancer mortality rates for each year of followup for five [4–10] of the seven studies and combined cumulative breast cancer mortality rates up to each year of followup for six [4–11] of the seven studies in both the intervention and the control groups among younger and older women. At or about 7 years of followup, data were available from all seven studies, while at about 10 years of followup data were available from five studies. Summary relative risks of breast cancer mortality for women offered screening compared with women not offered screening were calculated for younger and older women.[12] A weighted arithmetic average cumulative breast cancer mortality rate was calculated from the cumulative breast cancer mortality rates up to each year of followup from six of the seven trials weighted by the number of person-years at 7 years in each study separately for the intervention and control groups. This enabled estimation of the relative risk of breast cancer mortality among women offered screening compared with women not offered screening up to each year of followup; however, confidence intervals could not be calculated. Heterogeneity of the crude relative risks associated with screening up to 7 years of followup between younger and older women was evaluated.[13] Because the design of the Canadian trial was slightly different from that of the other RCTs, the analysis was conducted with and without inclusion of the results of this trial.

Logistic regression of the yearly breast cancer mortality rates of the intervention and control groups for the five studies with available yearly followup data was also used to assess differences in the effect of screening between the groups.[14]

Results

There was no significant reduction in breast cancer mortality in younger women (women under about 50 years of age), RR=1.04 (95% CI 0.81–1.33), with 143 deaths from breast cancer among younger women offered screening and 127 deaths among those not offered screening. Exclusion of the Canadian trial reduced the summary relative risk to 0.95 (95% CI 0.72–1.25), exclusion of the Malmo trial alone reduced the summary relative risk to 1.01 (95% CI 0.78–1.31), and exclusion of both the Canadian and Malmo trials reduced the summary relative risk to 0.89 (95% CI 0.66–1.20). The results of the Canadian trials were not particularly dissimilar to the results of other trials at an equivalent length of followup. At longer followup (10 years), there was a nonsignificant reduction in mortality in younger women offered mammography, RR=0.81 (95%CI 0.63–1.03), based on fewer trials; but in several trials, a nonsignificant excess of mortality from breast cancer was seen at short followup times.

In contrast, for older women at 7 years of followup, breast cancer mortality was reduced in those offered screening in all seven trials, with a summary relative risk of 0.74 (95% CI 0.62–0.87), with 269 deaths from breast cancer among older women offered screening and 327 deaths among those not offered screening, which differed little from the crude relative risk of 0.71. Exclusion of the Canadian trial reduced the summary relative risk to 0.71 (95% CI 0.60–0.85). The effect of screening in younger women was significantly different from that in older women (chi-square=6.44 on 1 degree of freedom, p=0.01, based on the crude relative risks).

Age group (younger versus older women), study, and year of followup were the main variables assessed in the logistic regression model for the five studies with available data for each year of followup. Significant interaction between outcome and age was present (deviance=3.89, 1 degree of freedom, $p<0.05$). Interaction between outcome and study with adjustment for the length of followup was not significant for either younger or older women, suggesting that results between studies were not significantly heterogeneous within each age category.

Conclusions

The meta-analysis did not adjust for differences in screening frequency and variation due to the cluster sampling used in the Swedish Two–County, Gothenburg, and Edinburgh studies. Adjustment for the cluster sampling method of subject selection has not been found to greatly alter the results in the Swedish Two–County study.[5] Adjustment for socioeconomic status for the results of the Edinburgh study[15] or the use of a random effects model would not be expected to greatly alter the results of this meta-analysis. These results suggest a biological difference in the effect of screening between younger and older women. Specifically designed studies and an analysis of the results of the effectiveness of the earlier screening rounds of present studies are required to clarify this.[16,17]

Acknowledgments

The assistance of Dr. Freda Alexander in providing data from the Edinburgh trial and Professor Anthony Miller in providing data from the Canadian trial are gratefully acknowledged. Helpful comments on an unpublished manuscript upon which some of this abstract is based from Professors Mark Elwood and Anthony Miller and Dr. Ann Richardson are also gratefully acknowledged.

References

1. Elwood JM, Cox B, Richardson AK. The effectiveness of breast cancer screening by mammography in younger women. Online J Curr Clin Trials 1993; document 32, para 1–195.

2. Nystrom L, Rutqvist LE, Wall S, Lindgren A, Lindqvist M, Ryden S, et al. Breast cancer screening with mammography: overview of Swedish randomized trials. Lancet 1993;341:973–8.

3. Nystrom L, Larsson L. Breast cancer screening with mammography [letter]. Lancet 1993;341:1531–2.

4. Shapiro S, Venet W, Strax P, Venet L. Periodic screening for breast cancer: the Health Insurance Plan Project and its sequelae, 1963–1986. Baltimore: Johns Hopkins University Press, 1988.

5. 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.

6. Miller AB, Baines CJ, To T, Wall C. Canadian National Breast Screening Study: 1. Breast cancer detection and death rates among women aged 40 to 49 years. Can Med Assoc J 1992;147:1459–76.

7. Miller AB, Baines CJ, To T, Wall C. Canadian National Breast Screening Study: 2. Breast cancer detection and death rates among women aged 50 to 59 years. Can Med Assoc J 1992;147:1477–88.

8. Roberts MM, Alexander FE, Anderson TJ, et al. Edinburgh trial of screening for breast cancer: mortality at seven years. Lancet 1990;335:241–6.

9. Alexander F, Anderson TJ, Brown HK, et al. The Edinburgh randomized trial of breast cancer screening: results after 10 years of followup. Br J Cancer 1994;70:542–8.

10. Andersson I, Aspergren K, Janzon L, et al. Mammographic screening and mortality from breast cancer: the Malmo mammographic screening trial. Br Med J 1988;297:943–8.

11. Frisell J, Eklund G, Hellstrom L, Lidbrink E, Rutqvist L, Somell A. Randomized study of mammography screening—preliminary report on mortality in the Stockholm trial. Breast Cancer Res Treat 1991;18:49–56.

12. Mantel N, Haenszel W. Statistical aspects of the analysis of data from retrospective studies of disease. J Natl Cancer Inst 1959;22:719–48.

13. Breslow NE. Elementary methods of cohort analysis. Int J Epidemiol 1984;13:112–5.

14. Payne CD, ed. The GLIM system: release 3.77. Oxford: Numerical Algorithms Group, 1985.

15. Alexander F, Roberts MM, Lutz W, Hepburn W. Randomization by cluster and the problem of social class bias. J Epidemiol Community Health 1989;43:29–36.

16. Hakama M. Breast cancer screening with mammography [letter]. Lancet 1993;341:1531.

17. Cox B. Benefit of mammography screening in women ages 40–49 years: current evidence from randomized controlled trials [letter]. Cancer 1996;78:572–3.

The Quality and Interpretation of Mammographic Screening Trials for Women Ages 40–49

Paul Glasziou, M.B.B.S., Ph.D.

Aim

To carry out a systematic review of the quality and results of all randomized trials of mammographic screening that included women aged less than 50 years.[1]

Methods

Reports of randomized trials of mammographic screening were identified via MEDLINE and checks of the bibliographies of retrieved articles and reviews. Identified trials were assessed for the following design features: (a) method of randomization, (b) documented comparability of baseline data, (c) standardized criteria for breast cancer death, (d) blinded review of cause of death, (e) completeness of followup, and (f) use of an "intention to treat analysis." To blind the two assessors, only the methods section and part of the results section describing baseline data and followup were provided; any material that could either identify the trial, its outcome, authors, or country was removed or blacked out by a research assistant.

Results

Eight randomized trials were identified, seven of which included women aged less than 50 years: four Swedish trials,[2] the Edinburgh trial,[3] the Canadian trial,[4] and the Health Insurance Plan (HIP) of Greater New York trial.[5] The quality of trials was generally high, with the Canadian study being among the most meticulous in design. A total of almost 160,000 women had been randomized. At the time of our analysis, the combined estimate of relative risk was 0.95 (95% CI: 0.77, 1.18), that is, a statistically nonsignificant relative risk reduction of 5 percent. However, recent updates of the data[6] would now suggest a 17-percent reduction (95% CI: 0–31). Adjustment for the cluster randomization of two trials, and for degree of compliance with screening, did not substantially change this result.

Discussion

These analyses suggest a modes benefit for women less than 50 years of age. The results are not explained by the quality of the trials or by the quality of the radiology. What does this mean to a 40-year-old woman considering screening? As Harris et al.[7] suggest: if 10,000 40-year-old women were screened, 500 would show an abnormality; of these 500, about 12 would have breast cancer while 100–200 of those with false-positive results would undergo an invasive procedure. Of the 12 with breast cancer, about 6 would eventually die from that cancer, but by screening we may have prevented perhaps one of these premature deaths (using the more recent figure of 17 percent above). Is that worthwhile? This is a difficult question, and we recommend that women in this age group intending to be screened be fully informed of these results in terms of absolute benefit.

References

1. Glasziou PP, Woodward AJ, Mahon C. Mammographic screening trials for women aged under 50: a quality assessment and meta-analysis. Med J Aust 1995;162:625–9.

2. Nystrom L, Rutqvist L, Wall S, et al. Breast cancer screening with mammography: overview of Swedish randomized trials. The Lancet 1993;341:973–8.

3. Roberts M, Alexander F, Anderson T. Edinburgh trial of screening for breast cancer: mortality at seven years. The Lancet 1990;335:241–6.

4. Miller A, Baines C, et al. Canadian national breast screening study: 1. breast cancer detection and death rates among women aged 40 to 49 years. Can Med Assoc J 1992;147:1459–88.

5. Venet L, Shapiro S, Roeser R, Venet W, Strax P. Ten-to-fourteen year effect of screening on breast cancer mortality. J Natl Cancer Inst 1982;349–55.

6. Conference Proceedings, Falun, Sweden, March 1996.

7. Harris R, Leininger L. Clinical strategies for breast cancer weighing and using the evidence. Ann Intern Med 1995;539:539–47.

Efficacy of Screening Mammography: Relative and Absolute Benefit

Karla M. Kerlikowske, M.D.

In evaluating the controversy concerning regular screening mammography for women ages 40–49, it is important to remember that the goal of screening is to reduce the likelihood of death in a person who has the disease. Randomized controlled trials are the most unbiased means to assess whether a screening test reduces the likelihood of death in a person who has the disease and are considered the "gold standard" when evaluating the efficacy of screening tests. There is good evidence from such trials that mammography reduces breast cancer mortality in women ages 50–69.[1–4] A meta-analysis of data in women age 50 and older from eight randomized controlled screening mammography studies demonstrates an overall significant -27 percent (95% CI -6 to -37%) reduction in breast cancer mortality after 7–9 years from the initiation of screening (Figure 1).[1] Of note, despite various study populations and interventions (different screening intervals, number of mammographic views, screening with or without clinical breast examination), all trials show a reduction in breast cancer mortality.

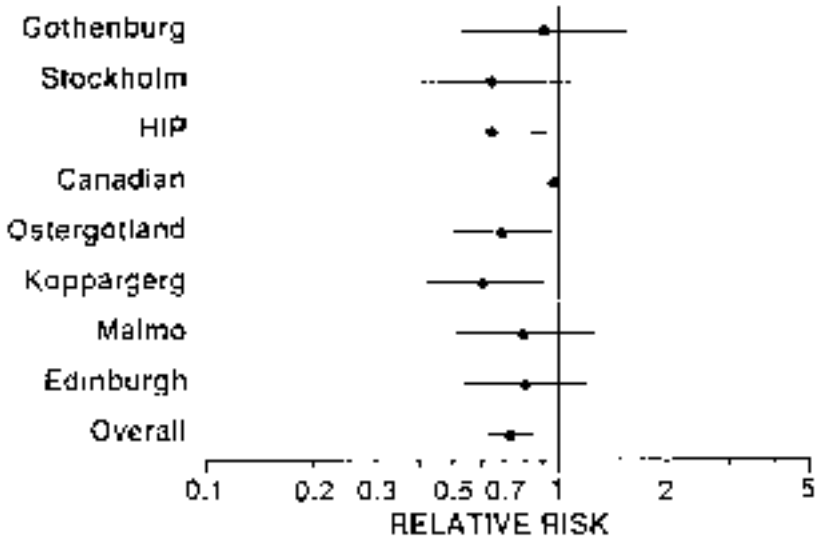

Figure 1. Reduction in Breast Cancer Mortality in Women Ages 50–74 After 7 to 9 Years of Followup[1]

In comparison, for women ages 40–49 from the same eight studies, no pattern or trend is present after 7–9 years from the initiation of screening (Figure 2). Four of the eight trials report a nonsignificant *increase* in breast cancer mortality, while four report a nonsignificant *decrease*, indicating a lack of statistically significant benefit or harm from screening mammography.[1] When data from the eight studies are combined using meta-analytic techniques, the overall summary estimate shows a nonsignificant +2 percent (95% CI -18 to +27%) increase in breast cancer mortality after 7–9 years from the initiation of screening (Figure 2). This means, whether a woman's breast cancer is mammographically or clinically detected, the risk of death from breast cancer is the same for screened and unscreened women for the first 7–9 years.

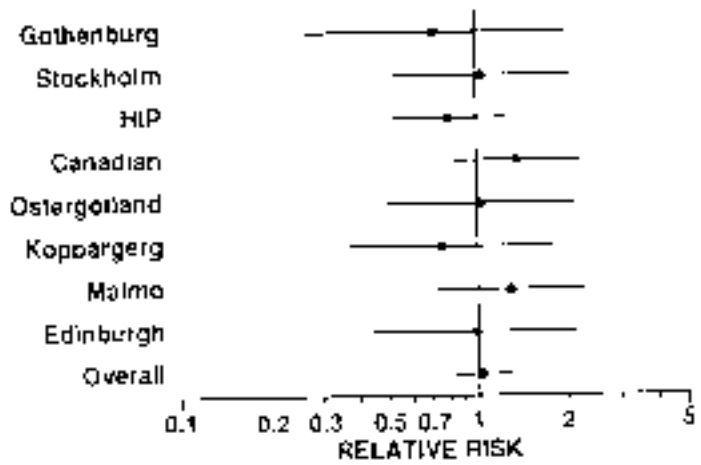

Figure 2. Reduction in Breast Cancer Mortality in Women Ages 40–49 After 7 to 9 Years of Followup[1]

Importantly, despite the diverse study populations and interventions of the various screening mammography trials, the combined results of the eight randomized, controlled trials are homogeneous (Chi-square test for homogeneity; $P = 0.5$), indicating little variability among the studies. In contrast, when data reported from 10–12 years from the initiation of screening are examined, a pattern begins to emerge, with four of five studies having a relative risk estimate to the left of one indicating a reduction in breast cancer mortality. However, all of the confidence intervals overlap one (Figure 3).

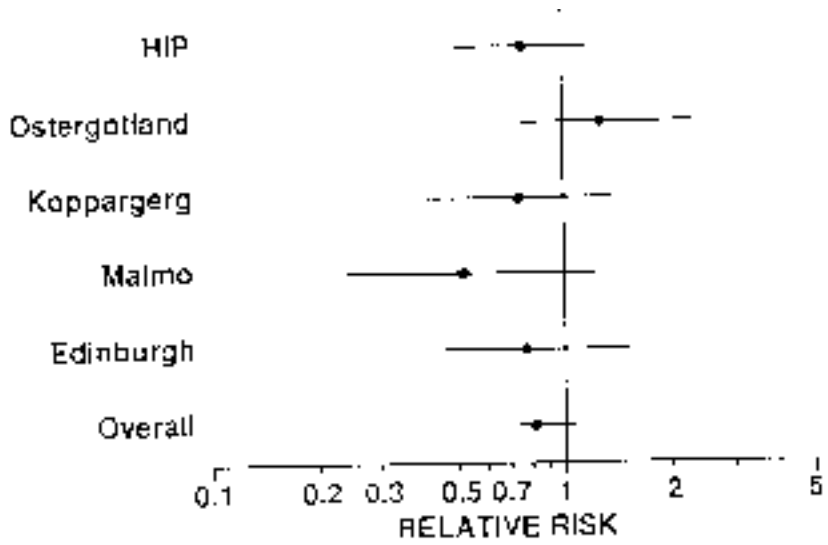

Figure 3. Reduction in Breast Cancer Mortality in Women Ages 40–49 After 10 to 12 Years of Followup[1]

When the five studies are combined, there is a trend toward a reduction in breast cancer mortality with an overall nonsignificant reduction of approximately -17 percent (95% CI -35 to +6%).[1] Pooled data from the five Swedish trials and results from the Health Insurance Plan (HIP) of Greater New York trial also have suggested an emerging benefit from screening mammography in younger women that does not occur until 8–10 years after the initiation of screening.[2–5]

It is unclear why any potential benefit from screening mammography in women ages 40–49 should be delayed a decade. Some have argued that there is insufficient statistical power to detect an early reduction in breast cancer mortality in women of these ages since there are too few women enrolled in the randomized controlled trials. However, if the explanation were simply lack of statistical power, then a reduction in breast cancer mortality should begin to appear after 4–5 years from the initiation of screening, as in women ages 50–69 (Figure 4)[2–4] and become statistically significant with longer followup. However, this does not appear to be the case since the data do not show a gradual separation of the mortality curves between screened and unscreened groups (Figure 5).[2–5] In fact, the data show slightly higher breast cancer mortality among women undergoing screening mammography the first 7–9 years after the initiation of screening.[1–5] One explanation for the delayed benefit could be the fact that most breast cancers detected among women who *start* screening at age 40–49 years are actually *detected* at or after age 50, when mammography is known to be efficacious. Only the HIP trial has published screening mammography results by age at detection. It showed that all of the decrease in breast cancer mortality among screened women ages 45–49 at entry occurred in those who had breast cancer detected at ages 50–54.[6] In addition, most women in the Edinburgh and Malmo trials, which also showed no benefit 7–9 from the initiation of screening but a trend toward a delayed benefit after 10–12,[1] were also probably age 50 or older when their breast cancer was diagnosed, since the youngest age of women at the start of screening was 45 years old.

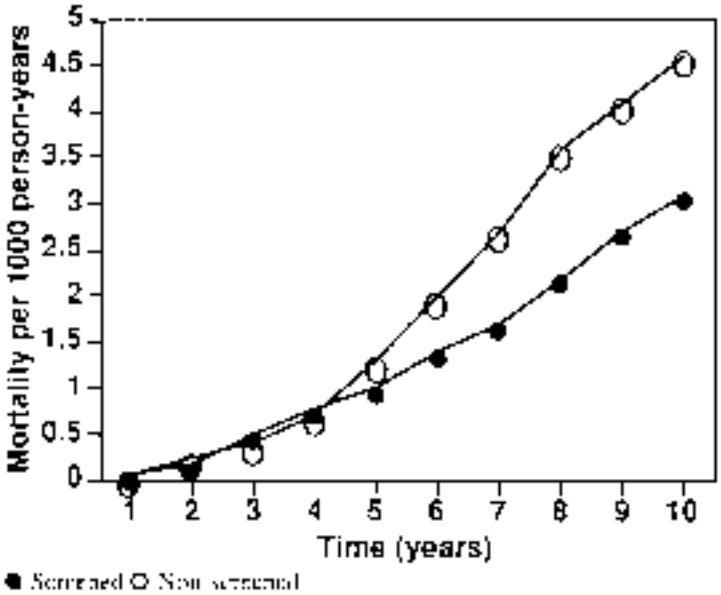

Figure 4. Cumulative Mortality in Screened Versus Nonscreened Women Ages 50–69[4]

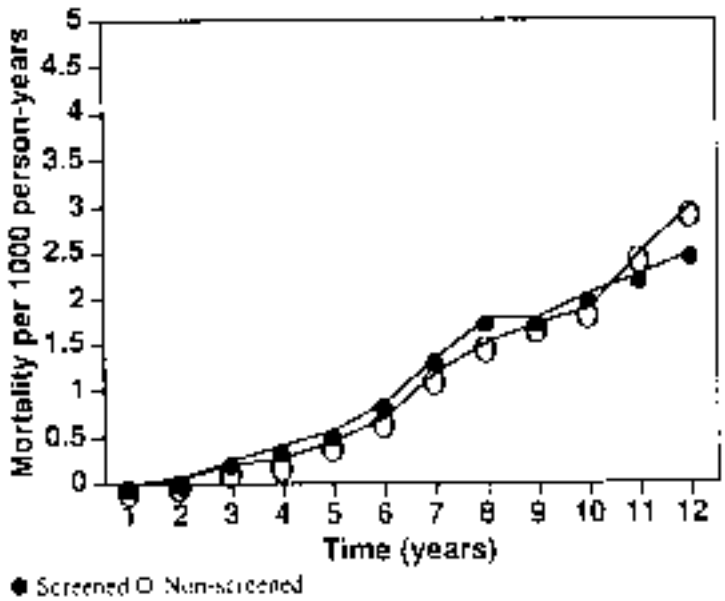

Figure 5. Cumulative Mortality in Screened Versus Nonscreened Women Ages 40–49[2]

Why is mammography efficacious as early as 4–5 years after the initiation of screening in older women? One explanation is that, among women age 50 and older, the sensitivity of mammography to detect invasive cancer is relatively high, resulting in few missed cancers. The relatively high sensitivity of mammography among older women is probably because a higher proportion have primarily fatty breast density, allowing easy detection of breast cancer and, because tumor growth rates are not as rapid as in younger women, allowing sufficient time for detection of small tumors.[7,8] Thus, among women age 50 and older, mammography detects the majority of tumors and detects them when they are more curable than if these tumors were detected clinically. In contrast, the sensitivity of screening mammography to detect invasive breast cancer is lower among women ages 40–49 compared with women age 50 and older (75 percent versus 93 percent).[8] The lower sensitivity in younger women may result from a greater proportion of invasive breast cancers being aggressive and, therefore, growing more rapidly, resulting in more interval cancers between regular screening examinations. Consequently, among women ages 40–49, the proportion of slow growing tumors detected by mammography is probably small, which may account for the marginal and delayed benefit. To better discern who is actually benefiting from the delayed mortality reduction observed among women who started screening in their forties (i.e., women age 50 and older, postmenopausal women, or women in their forties), trial data would need to be analyzed according to age at diagnosis and menopausal status.

Assuming that all of the delayed benefit in breast cancer mortality among women who initiated screening at age 40 results from detecting cancer in their forties, 3,330 forty-year-old women would have to be screened regularly between ages 40–49 to save one life from breast cancer. In comparison, 260 fifty-year-old women would need to be screened regularly from ages 50–69 to save one life from breast cancer. The more than 10-fold difference in the number needed to screen to save one breast cancer life is due to the lower incidence of breast cancer among women ages 40–49, the delay in benefit from screening, and the lower relative risk reduction compared with women age 50 and older. If only a proportion (30 percent) of the delayed benefit occurs because breast cancer was detected when women are in their forties, as reported by de Koning,[9] the number of 40-year-old

women needed to screen to save one life from breast cancer would be even larger.

In summary, based on the results of meta-analyses, there is no reduction in breast cancer mortality 7–9 years after the initiation of screening among women ages 40–49 years who undergo screening mammography. It is important to emphasize that, if screening mammography is effective in reducing breast cancer deaths among women ages 40–49 years, the reduction in deaths does not occur for at least a decade following the initiation of screening and appears to be smaller than the reduction observed in women age 50 and older. Given that the incidence of breast cancer is lower among women ages 40–49 and the potential benefit smaller and delayed, the absolute number of lives saved in this age group is likely to be much less than in women age 50 and older.

References

1. Kerlikowske K, Grady D, Rubin S, Sandrock C, Ernster V. Efficacy of screening mammography: a meta-analysis. JAMA 1995;273:149–54.

2. Nystrom L, Rutquvist LE, Wall S, et al. Breast cancer screening with mammography: overview of Swedish randomized trials. Lancet 1993;341:973–8.

3. Shapiro S. Periodic screening for breast cancer: the Health Insurance Plan project and its sequelae, 1963–1986. Baltimore, MD: Johns Hopkins University Press, Baltimore, 1988.

4. Tabar L, Fagerberg G, Duffy S, Day N, Gas A, Grontoft O. Update of the Swedish two county program of mammographic screening trial. Radiol Clin North Am 1992;30:187–210.

5. Tabar L, Faberberg G, Chen H, et al. Efficacy of breast cancer screening by age: new results from the Swedish two-county trial. Cancer 1995;75:2507–17.

6. Shapiro S, Venet W, Strax P, Venet L, Roeser R. Ten- to fourteen-year effect of screening on breast cancer mortality. J Natl Cancer Inst 1982;69:349–55.

7. Moskowitz M. Breast cancer: age-specific growth rates and screening strategies. Radiology 1986;161:37–41.

8. Kerlikowske K, Grady D, Barclay J, Sickles EA, Ernster V. Effect of age, breast density, and family history on the sensitivity of first screening mammography. JAMA 1996;276:33.

9. de Koning HJ, Boer R, Warmerdam, PG, Beemsterboer PMM, van der Maas PJ. Quantitative interpretation of age-specific mortality reductions from the Swedish breast cancer-screening trials. J Natl Cancer Inst 1995;87:1217–23.

Benefit of Mammography Screening in Women Ages 40–49: Current Evidence from Randomized Controlled Trials[1]

Charles R. Smart, M.D., F.A.C.S.; R. Edward Hendrick, Ph.D.;
James H. Rutledge III, M.S.; and Robert A. Smith, Ph.D.

Summary

Eight randomized controlled trials (RCTs) of screening mammography were conducted involving women ages 40–49 years at entry. Current data gathered for periods ranging from 7 to 18 years of followup are available from these trials. Meta-analyses were performed using a Mantel-Haenszel estimator method to combine current followup data from the eight RCTs of mammography that included women ages 40–49 years. Combining all current data on women ages 40–49 years at entry into the trials yielded a 16-percent benefit from screening mammography, without statistical significance at the 95% confidence level. Combining all data on women ages 40–49 years at entry, excluding results from the Canadian National Breast Screening Study (NBSS), yielded a 24-percent benefit for women invited for screening, with statistical significance at the 95% confidence level.

Background

A delayed demonstration of breast cancer mortality reduction is to be expected in women ages 40–49 for the following reasons:

1. Breast cancer incidence and mortality rates are lower in women ages 40–49 years.

2. The number of women ages 40–49 years included in the eight RCTs is less than one-third of the total.

3. The lead time of mammography is shorter in women ages 40–49 years yielding fewer cases.

4. The sensitivity of mammography appears to be lower in women ages 40–49 years.

5. The higher rates of *in situ* carcinoma in women ages 40–49 years and the slow progression to invasive carcinoma require a longer time to evaluate mortality differences.

6. Women ages 40–49 years have fewer involved positive lymph nodes and consequently have better survival and lower mortality rates than women ages 50–69 years.[2]

These factors make it more difficult to detect differences between study and control groups in women ages 40–49 years compared with women 50 years and older in the first 7 years of followup. Hence, more time is needed to manifest a statistically significant mortality reduction in women ages 40–49 years. Because of the delayed benefit of screening mammography in women ages 40–49 years, some have suggested that the benefit of mammography in RCTs may result from screening at or after the age of 50 years, even though women are younger than age 50 years when they enter the trials. This concern can be addressed only by the analysis of detailed data on the age at breast cancer diagnosis among women in both invited and control groups, data that are not available in published form from any RCT. Analysis of this question using detailed trial data could reintroduce biases, such as lead time bias, which are not present in RCTs that define participants on the basis of age at entry.

In October 1993, the Fletcher report[3] was issued summarizing the findings of the National Cancer Institute (NCI) International Workshop. This report relied on a meta-analysis of RCT data published the day of the NCI Workshop.[4] That meta-analysis included data from each trial at 7 years of followup, even though additional years of followup were available for six of the eight RCTs (see Table 1). The Fletcher report concluded that RCTs showed no benefit from screening mammography in women ages 40–49 years 5–7 years after entry into the trial, and uncertain or marginal benefit at 10–12 years after entry.

TABLE 1. Randomized Controlled Trials of Mammography

Trial (dates)	Age at Entry (yrs)	Regimen	Screening Frequency	Yrs F/U	No. of women: Invited	No. of women: Control	RR (95 % CI)
HIP Study (1963–69)	40–64	2 Vw MM + CBE	Annually 4 rounds	18	30,131	30,565	0.77[a] (0.61–0.97)
Malmo (1976–86)	45–69	1 or 2 Vw MM	18–24 mo, 5 rounds	12	20,695	20,783	0.81 (0.62–1.07)
Kopparberg (1977–85)	40–74	1 Vw MM	24–33 mo, 4 rounds	13	38,562	18,478	0.60 (0.46–0.79)
Ostergotland (1977–85)	40–74	1 Vw MM	24–33 mo, 4 rounds	13	38,405	37,145	0.78 (0.60–1.01)
Edinburgh (1979–88)	45–64	1 or 2 Vw MM	24 mo, 4 rounds	10	23,226	21,904	0.84[b] (0.63–1.12)
Stockholm (1981–85)	40–64	1 Vw MM	28 mo, 2 rounds	8	38,525	20,651	0.80 (0.53–1.22)
Gothenburg (1982–88)	40–59	2 Vw MM	18 mo, 4 rounds	7	20,724	28,809	0.86 (0.54–1.37)
NBSS–1 (1980–87)	40–49	2 Vw MM + CBE	Annually 5 rounds	7	25,214	25,216	1.36 (0.84–2.21)
NBSS–2 (1980–87)	50–59	2 Vw MM + CBE vs. CBE	Annually 5 rounds	7	19,711	19,694	0.97 (0.62–1.52)

Note: RR = relative risk; CI = confidence interval; F/U = followup; Vw = View; CBE = clinical breast examination; MM = mammography.
[a] From References 2 and 11 (for 18 years of followup). Other sources report slightly different results for this trial. For example, Reference 8 reports RR = 0.79 (0.62–0.99) at 18 years of followup.
[b] From Reference 3. Reference 7 reports RR = 0.85 (0.65–1.12).

Since the NCI's International Workshop, several updated results have been published or presented from the eight RCTs that included women ages 40–49 years. A reanalysis of the five Swedish RCTs after independent evaluation of deaths by a blinded end-point committee has provided revised data on all Swedish trials, with extended followup results from the Malmo,[5] Kopparberg, and Ostergotland trials.[6] New data with 11 years of followup from the Edinburgh trial were presented at the International Union Against Cancer Meeting on Breast Cancer Screening in Pre-menopausal Women in Developed Countries in Geneva in October 1993.[7,8] New data with 10 years of followup from the Gothenburg trial were presented at the National Breast Cancer Conference in May 1994.[9]

Methods

Recognition of the limited statistical power in individual RCTs, especially among women ages 40–49 years, has led some investigators to perform meta-analyses of RCT results.[4,8] Meta-analyses benefit by combining results from different studies, thereby increasing the total number of deaths and women-years (number of women (the number of years of followup) in the collective study and control groups. Just as the evaluation of an individual RCT benefits from more complete and longer duration of followup, meta-analyses of combined RCT results also benefit from longer term followup and more complete information from existing, individual studies. Meta-analyses, however, are not panaceas. Meta-analyses provide a mathematical

methodology for combining similar data (e.g., mortality data) from different trials. Meta-analyses are blind, however, to potential differences in design, conduct, quality, and completeness of data among studies.

Is it reasonable to consider the results of a meta-analysis that excludes the NBSS-1 trial,[10, 11] the only trial that was specifically designed to study screening in women ages 40–49 years? Because of several design differences between the NBSS-1 trial and all other RCTs that included women ages 40–49 years,[11] we believe that the answer is yes. First, the NBSS-1 trial invited volunteers to the entire trial, who were then randomized to the study or control groups; all other trials included all women of appropriate ages from a defined population group. Second, the NBSS-1 trial prescreened both the study group and the control group with a clinical breast exam; no other trial prescreened participants. Third, the prescreening occurring in the NBSS-1 trial before randomization of women leaves open the possibility of bias in the group assignment process; such bias was not possible in other RCTs. Whatever the explanation, the presence of nearly four times the rate of advanced disease in the study group biases the outcome. All of these features distinguish the NBSS-1 study from other previously conducted RCTs that included women ages 40–49 years. Further investigation is warranted to explain these differences.

Table 1 summarizes all of the breast cancer RCTs, indicating the period during which screening was
conducted, the ages of participants, the screening techniques, the periodicity of screening, the sample sizes for study and control groups, and the resulting relative risks and 95% confidence intervals for each trial. Table 2 summarizes the published subgroup data from each RCT relevant to screening of women ages 40–49 years, including the number of women entering each arm of the trials in this age subgroup. The present meta-analyses of current RCT data for women ages 40–49 years at entry have been performed using the Mantel-Haenszel estimator method for combining data from different trials. Table 3 lists input data to the RCT meta-analyses.

TABLE 2. Summary of Randomized Controlled Trial Results for Women Ages 40–49 Years

	Screening		Yrs F/U	No. of women			
Study Dates	Regimen	Frequency		Invited	Control	RR	(95% CI)
HIP Study[12] (1963–69)	2 Vw MM + CBE	Annually 4 rounds	18	14,432	14,701	0.77[a]	(0.53–1.11)
Malmo[5] (1976–86)	1 or 2 Vw MM	18–24 mo, 5 rounds	12	3658[a]	3679[a]	0.51[a]	(0.22–1.17)
Kopparberg[6] (1977–85)	1 Vw MM	24 mo, 4 rounds	13	9582	5031	0.73	(0.37–1.41)
Ostergotland[6] (1977–85)	1 Vw MM	24 mo, 4 rounds	13	10,262	10,573	1.02	(0.52–1.99)
Edinburgh[7] (1979–88)	1 or 2 Vw MM	24 mo, 4 rounds	11	5913[b]	5810[b]	0.78[b]	(0.46–1.51)
Stockholm[5] (1981–85)	1 Vw MM	28 mo, 2 rounds	8	14,375	7103	1.04[c]	(0.53–2.05)
Gothenburg[9] (1982–88)	2 Vw MM	18 mo, 4 rounds	7	10,600	12,800	0.73[d]	(0.27–1.97)
NBSS–1[10] (1980–87)	2 Vw MM + CBE	12 mo, 5 rounds	7	25,214	25,216	1.36	(0.84–2.21)

Note: RR = relative risk; CI = confidence interval; F/U = followup; Vw = View; CBE = clinical breast examination; MM = mammography.

[a]From Reference 11. Reference 7 reported RR = 0.78 (0.52–1.18).

[b]Includes only women ages 45–49 years at entry.

[c]Use of the raw data on deaths and on women-years of followup in each group results in a slightly different RR and 95% CI than that published.

[d]This RR and 95% CI corresponds to 7 years of followup. At 10 years of followup, the RR and 95% CI are estimated to be 0.60 (0.34–1.08).

TABLE 3. Data Used in Meta-analyses of Women Ages 40–49 Years

Screening Study	Number of Women-Years		Number of Breast Cancer Deaths	
	Invited	Control	Invited	Control
HIP Study	248,454	253,085	49	65
Malmo[a]	46,000	47,000	8	16
Kopparberg	107,000	56,000	22	16
Ostergotland	104,000	106,000	23	23
Edinburgh[a]	56,750	54,588	17	21
Stockholm	107,000	64,000	20	12
Gothenburg - A	64,000	77,000	6	10
Gothenburg - B	90,753	109,179	17	34
NBSS-1	173,474	173,488	38	28

[a]Includes only women ages 45–49 at entry.

Results

Figure 1 summarizes the individual RCT data and meta-analysis results with and without the NBSS-1 trial. A meta-analysis combining all trials except NBSS-1 at 7 years of followup gave a relative risk of 0.99 (0.74–1.32). At current followup, which ranges from 7 to 18 years, averaging 10.4 years, five studies show a relative risk less than 0.8 for women ages 40–49 years at entry (Table 2), and meta-analysis combining all data for women ages 40–49 years except NBSS-1 gives a relative risk of 0.76 (0.62–0.95). Five of the eight trials suggest benefit, although no single trial shows statistically significant benefit at the 95% confidence level. The failure to achieve a statistically significant difference in mortality between the study and control groups in a single RCT may be due to (1) a true lack of benefit from screening women in this age group, (2) an ineffective screening protocol,
(3) inadequate numbers of women enrolled, or (4) inadequate years of followup.

The true benefit of mammography today is likely to exceed the benefit demonstrated in RCTs for at least two reasons: (1) RCTs test the efficacy of the offer of mammography to a predefined study group compared with a predefined control group. In RCTs that measured compliance among women offered screening, compliance rates to obtain the first screening mammogram ranged from 61 to 89 percent. Assuming benefit exists, the true benefit to women who receive regular screening mammography will be higher than the benefit demonstrated among women who were offered mammography in the RCTs. (2) The technology of mammography has improved considerably since the time of even the most recent RCTs. Women receiving regular, high-quality mammography today are more likely to have their cancers detected at smaller sizes and at earlier stages than women who participated in the eight RCTs.[13]

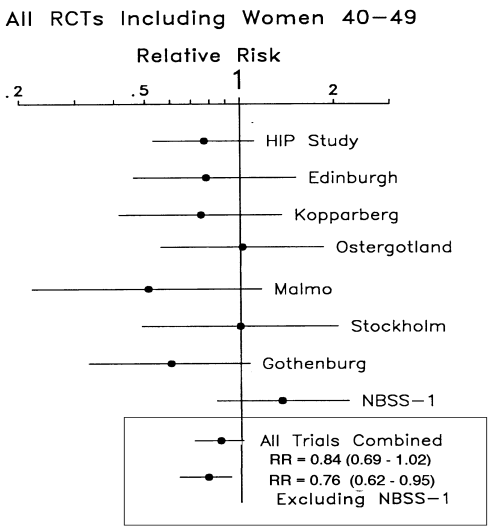

FIGURE 1. Individual RCT Data and Meta-analysis Results With and Without NBSS-1

Conclusions

A statistically significant reduction in mortality is shown at the 95% confidence level for women ages 40–49 years if the most recent data from all eight RCTs are used, with the exclusion of the NBSS-1 data. This result suggests that mammography was effective in women ages 40–49 years at entry to screening, even with noncompliance of some women in the study groups and lower quality mammography than exists today. Even greater benefit should exist today from regular screening of women ages 40–49 years than has been demonstrated by the collective results of the eight RCTs.

References

1. Smart CR, Hendrick RE, Rutledge III JH, Smith RA. Benefit of mammography screening in women ages 40 to 49 years. Current evidence from randomized controlled trials. Cancer 1995;75:1619–26.

2. Byrne C, Smart CR, Cherk C, Hartmann WH. Survival advantage differences by age: evaluation of the extended follow-up of the Breast Cancer Detection Demonstration Project. Cancer 1994;74:301–10.

3. Fletcher SW, Black W, Harris R, Rimer BK, Shapiro S. Report of the International Workshop on Screening for Breast Cancer. J Natl Cancer Inst 1993;85:1644–56.

4. Elwood JM, Cox B, Richardson AK. The effectiveness of breast cancer screening by mammography in younger women. J Curr Clin Trials (online) 1993;1:32–227.

5. Nystrom L, Rutqvist LE, Wall S, Lindgren A, Lindqvist M, Ryden S, et al. Breast cancer screening with mammography: overview of the Swedish randomized trials. Lancet 1993;341:973–8.

6. Tabár L, Fagerberg G, Chen H, 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(10):2507–17.

7. Eckhardt S, Badellino F, Murphy GP. UICC meeting on breast cancer screening in pre-menopausal women in developed countries. Int J Cancer 1994;56:1–5.

8. Wald N, Chamberlain F, Hackshaw A. Report of the European Society for Mastology: Breast Cancer Screening Evaluation Committee (1993). Breast 1993;2:209–16.

9. Bjurstam N, Bjorneld L. Mammography screening in women aged 40–49 years at entry: results of the randomized, controlled trial in Gothenburg, Sweden. 26th National Conference on Breast Cancer; May 10, 1994; Palm Desert, California.

10. Miller A, Baines CJ, To T, Wall C. Canadian National Breast Screening Study: 1. Breast cancer detection and death rates among women aged 40–49 years. Can Med Assoc J 1992;147:1459–76.

11. Tarone RE. The excess of patients with advanced breast cancer in young women screened with mammography in the Canadian National Breast Screening Study. Cancer 1995;75:997–1003.

12. Shapiro S, Venet W, Strax P, Venet L. Periodic screening for breast cancer: the Health Insurance Plan project and its sequelae; 1963–1986. Baltimore (MD): Johns Hopkins University Press, 1988.

13. Curpen NB, Sickles EA, Sollitto RA, Ominsky SH, Galvin HB, Frankel SD. The comparative value of mammographic screening for women 40–49 years old versus women 50–64 years old. Am J Roentgenol 1995;164(5):1104–6.

Markov Models for Breast Tumor Progression: Estimates from Empirical Screening Data and Implications for Screening

Stephen W. Duffy, M.Sc.; N.E. Day; László Tabár, M.D.; and H.H. Chen

The potential for tumor progression, particularly while the disease is preclinical (i.e., asymptomatic) but screen-detectable, is of importance in explaining the varying efficacy of screening by age. Fundamental indicators of a tumor's opportunity to progress in the context of a screening program are the average duration of the preclinical screen-detectable period, known as the mean sojourn time (MST), and the sensitivity of the screening process. The former gives an upper bound to the average lead time attainable, and the latter is an indication of the effectiveness of the screening tool. These quantities have been the targets of numerous modeling methods in the past. Seminal research on modeling of disease screening was conducted by Zelen and Feinleib,[1] Prorok,[2] and Day and Walter.[3]

In this paper, we use Markov chain models, beginning with simple models to estimate the MST and sensitivity, and building up to complex models of tumor progression with respect to size and node status.[4–6] We estimate the parameters of these models using age-specific data, mainly from the Swedish Two-County Trial.[7] We further investigate the age-specific potential for dedifferentiation whereby in a tumor with an internally heterogeneous histological pattern, the more poorly differentiated component grows more rapidly than the better differentiated,[8] and a tumor with malignancy grade 1 or 2 at inception may become grade 3 if it is not detected and excised at an early phase of its development.

The implications of the results are considered in terms of mortality from breast cancer. With particular emphasis on the 40–49 age group, we address the questions of the existence and size of a benefit and the delay in time before such a benefit can be observed. The question of age at diagnosis is discussed.

Major findings (see Table 1) include the following:

- Progression in the preclinical phase is more rapid in women ages 40–49 than in women age 50 or older.

- There is a greater tendency to dedifferentiation in tumors in younger women. Also, the dedifferentiation tends to take place before progression with respect to size or regional lymph node invasion in women ages 40–49, contrary to the order of progression in women age 50 or older.

- Consequently, for a screening program to substantially reduce mortality from breast cancer in women ages 40–49, the interscreening interval should be around 1 year and certainly no more than 18 months.

- The modeling agrees with empirical results that indicate that the effect on mortality will not be seen until 8–10 years after the inception of screening unless the program is sufficiently sensitive to have an impact on the grade 3 or potential grade 3 tumors.

TABLE 1. Some Estimated Parameters of Progression and Screening by Age			
Quantity	Ages 40–49	Ages 50–59	Ages 60–69
MST	2.44	3.70	4.17
Sensitivity	83%	100%	100%
Annual dedifferentiation rate	47%	12%	15%
Annual node invasion rate	24%	19%	16%

- In a population ages 40–49 at the inception of a screening program, the results suggest that the major mortality benefit is observed in cancers diagnosed before age 50.

References

1. Zelen M, Feinleib M. On the theory of screening for chronic disease. Biometrika 1969;56:601–14.

2. Prorok PC. The theory of periodic screening II. Doubly bounded recurrence times and mean lead time and detection probability estimation. Adv Appl Probability 1976; 8:460–76.

3. Day NE, Walter SD. Simplified models of screening for chronic disease: estimation procedures from mass screening programs. Biometrics 1984;40:1–14.

4. Duffy SW, Chen HH, Tabar L, Day NE. Estimation of mean sojourn time in breast cancer screening using a Markov chain model of both entry to and exit from the preclinical detectable phase. Stat Med 1995;14:1531–43.

5. Chen HH, Duffy SW, Tabar L. A Markov chain method to estimate the tumor rate from preclinical to clinical phase, sensitivity and positive predictive value for mammography in breast cancer screening. The Statistician 1996;45:307–17.

6. Committee and Collaborators, Falun Meeting. Report of the meeting on mammographic screening for breast cancer in women aged 40–49, Falun, Sweden, March 1996. Int J Cancer (in press).

7. Tabar L, Fagerberg G, Chen HH, et al. Efficacy of breast cancer screening by age: New results from the Swedish two-county trial. Cancer 1995;75:2507–17.

8. Tabar L, Fagerberg G, Chen HH, Duffy SW, Gad A. Tumour development, histology and grade of breast cancers: Prognosis and progression. Int J Cancer 1996; 66:413–9.

Quantitative Interpretation of Age-Specific Mortality Reductions from Trials by Microsimulation

Harry J. de Koning, Ph.D., and Rob Boer, M.Sc.

Introduction

In 1995, it was still uncertain whether breast cancer screening for women under 50 years of age was effective in reducing breast cancer mortality. According to some investigators, the results from the five Swedish randomized breast cancer screening trials can be considered as the most conclusive evidence on the effect of mammographic screening.[1,2] The published 10–13 percent breast cancer mortality reduction rates for Swedish women under 50 years of age entered in a randomized study seemed encouraging. However, some women in this age group were also screened when they were 50-years-old and older. Part of the observed mortality reduction in these women is likely to have been a result of detecting the cancer earlier in later rounds when the women were 50 years old or older. Furthermore, any trial is specific in its design, quality, and background situation. Consequently, different screening trials will result in different breast cancer mortality reductions in the study group as compared with the control group, even for the same age categories (apart from random fluctuation). The results from all five Swedish randomized trials, specified per age category, again show that the estimates of reduction vary widely between the trials. Important characteristics of the trials, such as screening interval, attendance rate, followup period and age groups, should be considered.

By using one underlying model of the natural history of breast cancer and of the performance of mammographic screening, we have analyzed all five trials taking into account nine important characteristics within each trial.[3] The different policies for women under and above age 50 at entry are distinguished as are characteristics in screening policy regarding intervention in the control groups. The goal is to adjust for as many relevant characteristics in screening policy as possible in each trial that may have influenced the outcome and estimate the improvement in prognosis for screen-detected cases. We address the question: Which percentage of the observed mortality reduction for women ages 40-49 at entry into the trial might have been due to screening these women when they were 50 years or older. Specified methods are described elsewhere.[3,4]

Results

Table 1 shows the observed number of breast cancer deaths and relative risks (RR) in all study and control groups combined per age category and the expected numbers and RR obtained from the model based on the 1993 results. In all variants we used one same underlying model of the natural history of disease and performance of screening (test) for each of the five trials (i.e., sensitivity and duration distribution of preclinical, screen-detectable disease by age and state). The simulation of the specific Swedish trial designs then leads to different numbers of screen-detected cancers in the different states with a consequent reduced risk of dying from breast cancer. Also, one identical parameter for

TABLE 1. Woman-years and Observed Number of Breast Cancer Deaths in Study and Control Groups and Relative Risk (RR) of All 5 Swedish Trials Combined, per Age Category, Compared with Expected Ones with the Model with Different Assumptions on Improvement in Prognosis for Screen-Detected Cases for Women Ages 40–49

Woman-Years (x 1,000)*		Observed Breast Cancer Deaths*			Expected Breast Cancer Deaths		
Study group	Control group	Study group	Control group	RR	Study group	Control group	RR‡
(a) women ages 50–69 entering trial; assuming an observed overall 30-percent reduction (RR=0.70) in the risk of women ages 50–69 dying of breast cancer, for all trials, and fitting this same reduction to the model (RR=0.70) for consistency							
911	725	281	312	0.70†	275.3	317.7	0.70
(b) women ages 40–49 entering trial; assuming an improvement in breast cancer prognosis equal to the one estimated for women age 50 years or above as in line a							
428	350	84	75	0.90*	79.5	79.5	0.77
(c) women ages 40–49 entering trial; assuming no improvement in prognosis for cancers that are screen detected before 50 years of age and improvement in prognosis equal to the one estimated for women aged 50 years or above as in line a							
428	350	84	75	0.90*	86.7	72.3	0.93

* (follow-up model; end point 31 Dec. 1989); exclusion of cases 70–74 group estimated from numbers of women at entry

† average of published RRs (0.72 age group 50–59; 0.69 age group 60–69)

‡ calculated by Mantel-Haenszel method

improvement in prognosis was assumed for all trials. First, this state and age-specific parameter was estimated to make the results of the model for all trials together, consistent with the 1993 observed overall 30-percent reduction for women ages 50–69 (line a).

The Table also shows the results from the trials and the model for women ages 40–49 (or 45–49) at randomization (lines b-c). The observed reduction in this age group for all trials combined was 10 percent in 1993, a third of that in the age group above 50. In the analysis summarized in line b, all screen-detected cases, resulting from the design and the underlying model, have been given the same improvement in prognosis, equal to the one estimated for women above age 50 years (as in line a). With that assumption, we would have expected an overall 23-percent reduction in breast cancer mortality (RR= 0. 77) for all women ages 40–49 at entry, given the characteristics and the followup periods of all trials (line b). Assuming no improvement in prognosis for cancers that are screen-detected before 50 years of age and the same improvement as in line a for cancers that are screen-detected at 50 years and over, a 7-percent mortality reduction between invited and control groups would have been expected at the end of followup of these five trials (line c). In other words, in a so-called pessimistic variant

where there is no benefit in the model for women whose cancer was detected by screening before age 50 (compared to no screening), in this analysis we expected that the five trials would have shown a 7-percent reduction on 31 December 1989 for women ages 40–49 at the start of the trials. This expected 7-percent benefit, therefore, must have been derived from the model for women whose cancer was detected at age 50 years or more, where we did assume a reduction in risk of dying of breast cancer. For all trials together, we were able to make a good fit (agreement) between the observed reduction and the model.

Our analysis led to the assumption that the improvement in prognosis due to screening for women ages 40–49 is much smaller than for women above age 50. About 70 percent of the reduction observed in the trials' results in 1993 for women ages 40–49 at entry might be attributed to a reduction due to screening these women when they were 50-years-old or older. Certainly a number of details from the trials or the base line situation in the regions have not been available to us: regional epidemiology (population structure, incidence by age, stage distribution and treatment, survival), details about the influence of the type of randomization, and especially about the situation after the first screening of the control group. Other remaining factors might have influenced the outcome of screening. Further research should be initiated to quantify the quality of screening, especially considering younger women.
Such detailed information is not available in literature on all five trials. It seemed appropriate to base this analysis with respect to the natural history of breast cancer and age-specific sensitivity partly on Dutch data. With detailed information on detection rates, interval cancers, stage distributions, and background situation per Swedish trial, it would be possible to estimate whether the assumptions on natural history or sensitivity might have to be adjusted. The especially interesting question is whether it is indeed correct to attribute most of the published mortality reduction for women invited below 50 to screening above this age group. A standard 'meta-analysis' does not account for the underlying parameters causing a possible different RR in a particular screening situation.

New Results 1995–1996

True data on age at detection and mortality differences were not available until recently. Tabár and colleagues have shown that, in women ages 40–49 years at randomization, 36 percent of the breast cancers were indeed diagnosed at age 50 years and over.[5] They could not find less benefit in the women with screen-detected cancers before age 50 than for cancers detected at later stages, and apparently also detected more cancers before age 50 years than we modeled. Sensitivity might be higher and, for this trial, the benefit expected of screening young women larger than estimated. But to explore the precise implications likewise data from the other trials are crucial.

The newest results from Sweden, showing a 23-percent reduction in the younger age group[6] are again promising. The point estimate now resembles our variant in the Table (line b). But with longer followup, it is likely that Swedish trials will show a higher or significant breast cancer mortality reduction for the women ages 40–49 at entry, merely due to the beneficial effect of screening more women at age 50 and over in later rounds. Before we value this as an equally effective program in younger versus older women, it is still crucial to have a reasonable estimate with regard to the amount of reduction achieved for these women entering the trials at younger ages on the basis of screening in later rounds for all trials, as Tabár and colleagues did. The results for the age group 50–69 might have improved too with longer followup; still leading to a difference (although smaller) in improvement in prognosis for

younger and older women. The Göteborg trial results now influence the overall RR for young women much more, and no data on cancers detected, stages, etc. are available from literature.

For the conference, we will re-estimate the results of the modeling given these newest results, and including all new data available and most recent followup period. Sensitivity in young women will be varied. To illustrate the necessity and complexity of such analyses, if sensitivity indeed has to be assumed higher than formerly done in the model for young women, the improvement in prognosis per screen-detected case would probably still be lower at young ages, to result in the same point estimate as observed.[7] The quality of the screening (in the specific trial) is then the crucial point to have as reference for future programs. Radiation risk estimates have already been performed.[8] It is correct to state that the newest Swedish results finally give more strength to the possible benefit of screening in young women with high-quality screening. The extent of a mortality reduction remains to be analyzed as well as the balance between possible favorable and unfavorable effects of population-wide screening in this age group.[9]

References

1. Nyström L, Rutgvist LE, Wall S, Lindgren A, Lindqvist M, Rydén S, Andersson I, Bjurstam N, Fagerberg G, Frisell J, Tabár L, Larsson L-G. Breast cancer screening with mammography: overview of Swedish randomized trials. Lancet 1993;341:973–8.

2. Nyström L, Larsson LG. Breast cancer screening with mammography [letter]. Lancet 1993;341:1531–2.

3 . Koning HJ de, Boer R, Warmerdam PG, Beemsterboer PMM, Maas PJ van der. Quantitative interpretation of age-specific mortality reductions from the Swedish Breast Cancer-Screening Trials. J Natl Cancer Inst 1995;87:1217–23.

4. Oortmarssen GJ van, Habbema JDF, Maas PJ van der, Koning HJ de, Collette HJA, Verbeek ALM, Geerts AT, Lubbe KTN. A model for breast cancer screening. Cancer 1990;66:1601–12.

5. Tabár L, Duffy SW, Chen H-H. Re: Quantitative interpretation of age-specific mortality reductions from the Swedish Breast Cancer-Screening Trials. J Natl Cancer Inst 1996;88:52–3.

6. Nyström L. Update of the statistical analysis of screening women aged 40–49 in the combined Swedish trials. In: Congress Proceedings of the International Conference "The impact of breast cancer screening with mammography in women aged 40-49 years." March 21-22 1996, Falun, Sweden.

7. Koning HJ de, Boer R, Warmerdam PG, Beemsterboer PMM, Maas PJ van der. Response to Tabár Re: Quantitative interpretation of age-specific mortality reductions from the Swedish Breast Cancer-Screening Trials. J Natl Cancer Int 1996;88:54.

8. Beemsterboer PMM, Warmerdam PG, Boer R, Koning HJ de. Radiation risk of mammography related to benefit in screening programmes: a favorable balance? (submitted).

9. Koning HJ de. Current controversies in cancer. Is mass screening for breast cancer cost effective? Eur J Cancer 1996;32A:1835–44.

Problems With the Randomized Controlled Trials of Screening and Inappropriate Analysis of Breast Cancer Data

Daniel B. Kopans, M.D., F.A.C.R.

In 1993, the National Cancer Institute (NCI) ignored the nearly unanimous vote of its National Cancer Advisory Board as well as the objections of numerous breast cancer experts and numerous medical organizations and withdrew its support for screening women ages 40–49 for breast cancer. Despite the fact that, at that time, most of the trials were demonstrating a benefit for women in their forties, the NCI decision was based on an analysis that had required that the randomized, controlled trials of screening (RCT) demonstrate a *statistically significant* mortality reduction for the subgroup of women in the trials who were ages 40–49.[1] Since this new position (actually formulated and promulgated by only a few individuals at the NCI) contradicted both the NCI's own recent advice, established in 1989 with its signing of the Consensus Agreement, and the continued advice of many other health organizations, it caused a great deal of confusion among American women and their physicians. Compounding the confusion was the decision by the NCI, soon after the guideline change, that the Institute would no longer issue guidelines, essentially eliminating the possibility of any review. Not only was the NCI decision, and the process it had followed to arrive at that decision, criticized by numerous experts on breast cancer screening, but it was sharply criticized by a Congressional review that labeled it "misused science."[2]

The controversy arose as the result of retrospective subgroup analysis of data from trials that were not designed to permit such analysis. With the exception of the National Breast Screening Study of Canada (NBSS), *none* of the randomized, controlled trials (RCTs) of screening were designed to evaluate women ages 40–49 as a separate group. Even the NBSS had insufficient statistical power as well as major flaws in its execution.[3] Experts warn of the dangers of using data from trials with insufficient power,[4,5] but these warnings were ignored. The NCI and others have subsequently acknowledged the lack of power in the trials,[6,7] but they consistently failed to inform women and their physicians of its implications. The requirement for statistical significance in the unplanned subgroup analysis was specious because the trials had not been designed to permit this and a "significant" mortality reduction of 25 percent was impossible to achieve in the early years of followup. *None* of the trials, individually or even collectively, had sufficient numbers of women in this decade of life to permit an expected benefit of 25 percent to be statistically significant in the early years of followup. Almost 500,000 women would have been required to have an 80 percent power to show this benefit.[8] The trials had included approximately one-third the necessary number. Unless the women in the trials had differed significantly from the general population, given the rate of cancers developing and the rate of deaths from breast cancer, it was *mathematically impossible* for a statistically significant benefit of 25 percent to be demonstrated for women ages 40–49 in the early years of followup.

Nonetheless, when analyzed, *as they were designed*, the trials have for many years demonstrated a statistically significant benefit for screening *beginning by the age of 40*.[9] It is only the inappropriate use of retrospective, unplanned, sub-group analysis to advise women and their physicians that has caused the controversy.[10,11]

In addition to unplanned subgroup analyses that lacked statistical power, arguments against

screening women in their forties have often used data grouping that skews results, leading to faulty conclusions. Women ages 40–49 have been repeatedly compared with all women ages 50 and over. This artificial grouping, making the age of 50 the point of analysis, biases the interpretation and leads to the flawed conclusion that the age of 50 is a significant change point.[12]

Despite these facts, the argument should now be moot. The NCI requirement has been achieved. In spite of the weaknesses in the trials, with longer followup, the trials now demonstrate statistically significant benefit, even when women ages 40–49 are analyzed separately.[13,14] In an effort to maintain a scientifically untenable position, it has now been suggested that this benefit is due to women reaching the age of 50 during the trials and screening suddenly becoming effective. A computer model has been used to buttress this argument.[i] The model, however, was based on the results from a case-controlled study in Nijmegen that had particularly poor results for women ages 40–49 because of a long screening interval and high thresholds for intervention. Not only is an abrupt change at age 50 biologically not supportable, but RCT data cannot legitimately be analyzed by age at diagnosis. Age at diagnosis is a pseudo-variable influenced by the intervention. Its use will, *a priori*, bias an analysis against cancers detected among younger women in the screened groups. Nevertheless, even if the rules of RCT analysis are violated and age at diagnosis is used, it has been shown that the benefit, in the three trials that have evaluated the data by age at diagnosis (HIP, Gothenberg, and Kopparberg), was primarily for women whose cancers were diagnosed while they were still in their forties.

Breast cancer is not a trivial problem for women in their forties. In 1993, 16 percent of the cancers diagnosed were among women in their forties. Some have suggested that screening should concentrate on the 84 percent of cancers that occur in women who are not in their forties. This is superficially compelling, but the appearance is due to the manner in which the data are presented. Women in their forties are compared with all other women grouped together. These kinds of comparisons can be misleading. For example, the same argument could be suggested with regard to women in any decade of life. No decade of life accounts for more than 24 percent of all of the cancers diagnosed in any year. In 1993, women in their fifties actually only accounted for 17 percent of the cancers diagnosed. In fact, as a consequence of the large numbers of women in their forties, in 1995 and 1996, there were actually *more breast cancers diagnosed among women in their forties than in women in their fifties*.[16] Perhaps of greater importance is that more than 30 percent of the years of life lost to breast cancer are due to women diagnosed while in their forties.

Some analysts have raised concern over the “harms” from screening. These include anxiety from the process, anxiety from being recalled for additional evaluation, and trauma from biopsies with benign results (termed “unnecessary”). What has been ignored is that these “harms” are true for women at all ages and do not change abruptly at the age of 50. The age of 50 is actually an arbitrary point of analysis that has been chosen as a surrogate for menopause. Its use has skewed the interpretation of numerous factors associated with screening by singling out and comparing women ages 40–49 with *all* women age 50 and over. By choosing age 50 as the point of analysis, changes that actually occur steadily with increasing age are made to appear to change abruptly at the age of 50. There are *no parameters of screening* that change abruptly at the age of 50 or any other age. The breast does not turn to fat (become less dense), screening recall rates and recommendations for breast biopsy do not change, and the cancer detection rate increases steadily, mirroring the prior probability of breast cancer in the population with no abrupt change at age 50.[17]

Not only were the RCT too small to permit early analysis of women ages 40–49, but they were not properly designed or executed for these women. The quality of the mammography was not optimized, the time between screens in most programs was too long, single-view mammography, which has been shown to miss as many as 20 percent of cancers[18] was used in many of the trials, and high thresholds for intervention permitted cancers to pass through the screens. The NBSS has come under even harsher criticism. The quality of the mammograms in the NBSS was, according to its own reference physicist, below even the community practice of the time.[19] The poor quality of the mammography was confirmed by an objective review.[20,21] The randomization process used in the NBSS was not blinded and resulted in an excess of advanced cancers allocated to the screened group, biasing the trial against screening from the outset.[22,23]

Of necessity, the RCT actually underestimate the benefit of screening due to noncompliance and contamination. With the exception of the Canadian trial, which involved volunteers (a separate problem), the seven other trials first randomized a population and then invited them to be screened. Women allocated to be screened who refused the invitation (noncompliance) are still counted as having been screened, and if they die of breast cancer their deaths are attributed to the screened group. Similarly, women who had mammograms on their own outside of the screening program (as high as
35 percent), and whose lives were saved as a result, are still counted as unscreened controls. Thus, the trial data permit analysis of the benefit from an "invitation" to be screened, and not actually from having been screened. The true benefit of screening is likely greater than the trial results would indicate.[24]

It has been stated that screening is a public health issue that is not concerned about individuals.[25] This is incorrect. Breast cancer screening is primarily an issue for the individual woman and is only a public health issue in terms of access and reimbursement.[26] The costs of screening are significant and may, ultimately, determine public policy. Nevertheless, it is important to separate the medical and scientific analysis from the economic considerations when advising women. "Society" may decide that it is too expensive to provide screening for breast cancer, but women should be provided with all of the information so that they can decide for themselves whether to be screened. They should not be prevented from participating in the discussion of whether screening is "worthwhile" by being informed, incorrectly, that it has no benefit.

Mammographic screening is by no means the ultimate solution to the breast cancer problem, but it offers the chance to significantly reduce the number of deaths from these cancers. The available data suggest that screening should be *annual* for all women *beginning by the age of 40*.[27,–29] Since there is little if any risk to the breast from radiation for women ages 40 and over,[30] the only reason to recommend a longer time interval between screens is economic, and women should be so informed.

References

1. Fletcher SW, Black W, Harris R, Rimer BK, Shapiro S. Report of the International Workshop on Screening for Breast Cancer. J Natl Cancer Inst 1993;85:1644–56.

2. House Committee on Government Operations. Misused Science: The National Cancer Institute's elimination

of mammography guidelines for women in their forties. Union Calendar No. 480. House Report 103-863. October 20, 1994.

3. Miller AB, Baines CJ, To T, Wall C. Canadian National Breast Screening Study: 1. Breast cancer detection and death rates among women aged 40–49. Can Med Assoc J 1992;147;1459–76.

4. Lachin JM. Introduction to sample size determination and power analysis for clinical trials. Controlled Clin Trials 1981;2:93–113.

5. Moher D, Dulberg C, Wells GA. Statistical power, sample size, and their reporting in randomized controlled trials. JAMA 1994;272:122–4.

6. Eckhardt S, Badellino F, Murphy GP. UICC Meeting on breast-cancer screening in pre-menopausal women in developed countries. Int J Cancer 1994;56:1–5.

7. Joint Meeting (American Cancer Society and the National Cancer Institute) on the Feasibility of a Study of Screening Young Women for Breast Cancer. Atlanta, Georgia. April 20–21, 1994.

8. Kopans DB, Halpern E, Hulka CA. Statistical power in breast cancer screening trials and mortality reduction among women 40–49 with particular emphasis on the national breast screening study of Canada. Cancer 1994;74:1196–203.

9. Shapiro S. Screening: Assessment of current studies. Cancer 1994;74:231–8.

10. Kopans DB. Mammography screening and the controversy concerning women aged 40–49. Rad Clin N Am 1995;33:1273–90.

11. Sickles EA, Kopans DB. Deficiencies in the analysis of breast cancer screening data. J Natl Cancer Inst 1993;85:1621–4.

12. Kerlikowske K, Grady D, Barclay J, Sickles EA, Eaton A, Ernster V. Positive predictive value of screening mammography by age and family history of breast cancer. JAMA 1993;270:2444–50.

13. Smart CR, Hendrick RE, Rutledge JH, Smith RA. Benefit of mammography screening in women ages 40–49: current evidence from randomized, controlled trials. Cancer 1995;75:1619–26.

14. Duffy SW, Tabar L. Falun meeting on breast cancer screening with mammography in women aged 40–49 years: report of the organizing committee and collaborators. Int J Cancer (in press).

15. de Konig HJ, Boer R, Warmerdam PG, Beemsterboer PMM, van der Maas PJ. Quantitative Interpretation of age-specific mortality reductions from the Swedish breast cancer-screening trials. J Natl Cancer Inst 1995;87:1217–23.

16. American Cancer Society Facts and Figures 1995 and 1996.

17. Kopans DB, Moore RH, McCarthy KA, Hall DA, Hulka CA, Whitman GJ, Slanetz PJ, Halpern EF. The positive predictive value of breast biopsy performed as a result of mammography: there is no abrupt change at

age 50 years. Radiology 1996;200:357–60.

18. Wald NJ, Murphy P, Major P, Parkes C, Townsend J, Frost C. UKCCCR multicentre randomized controlled trial of one and two view mammography in breast cancer screening. BMJ 1995;311:1189–93.

19. Yaffe MJ. Correction: Canada study. Letter to the Editor. J Natl Cancer Inst 1993;85:94.

20. Baines CJ, Miller AB, Kopans DB, Moskowitz M, Sanders DE, Sickles EA, To T, Wall C. Canadian national breast screening study: assessment of technical quality by external review. AJR 1990;155:743–47.

21. Kopans DB. The Canadian screening program: A different perspective. AJR 1990;155:748–49.

22. Kopans DB, Feig SA. The Canadian national breast screening study: A critical review. AJR 1993;85:1621–60.

23. Sickles EA, Kopans DB. Deficiencies in the analysis of breast cancer screening data. J Natl Cancer Inst 1993;85:1621–4.

24. Feig S. Estimation of currently attainable benefit from mammographic screening of women aged 40–49 years. Cancer 1995;75:2412–19.

25. King J. Mammography screening for breast cancer. Letter to the editor. Cancer 1994;73:2003–4.

26. Kopans DB. Reply. Cancer 1994;73:2004–6.

27. Moskowitz M. Breast cancer: Age-specific growth rates and screening strategies. Radiology 1986;161:37–41.

28. Tabar L, Faberberg G, Day NE, Holmberg L. What is the optimum interval between screening examinations?—An analysis based on the latest results of the Swedish two-county breast cancer screening trial. Br J Cancer 1987;55:547–51.

29. Kerlikowske K, Grady D, Barclay J, Sickles EA, Ernster V. Effect of age, breast density, and family history on the sensitivity of first screening mammography. JAMA 1996;276:33–8.

30. Mettler FA, Upton AC, Kelsey CA, Rosenberg RD, Linver MN. Benefits versus risks from mammography: a critical assessment. Cancer 1996;77:903–9.

Results from the National Breast and Cervical Cancer Early Detection Program,1991–1995

Nancy C. Lee, M.D.

Objective

To describe results from an ongoing nationwide mammography screening program for low-income, uninsured or underinsured women that is part of the National Breast and Cervical Cancer Early Detection Program (NBCCEDP), administered by the Centers for Disease Control and Prevention.[1,2]

Design

Longitudinal surveillance.

Setting

Hundreds of mammography and clinical facilities throughout the United States, including community health centers, health department clinics, private practitioners' offices, university-based facilities, and mobile mammography units.

Participants

Low-income uninsured or underinsured women who had at least one mammogram provided by the NBCCEDP from July 1991 through June 1995. The numbers of women included in this report were: 88,493 women ages 40–49 years; 72,001 women ages 50–59 years; 44,434 women ages 60–69 years; and 20,645 women age 70 years and older.

Main Outcome Measures

Abnormal mammography* and breast cancer rates, stage distribution of breast cancers, and positive predictive value of abnormal mammograms and biopsies, by age group, for first and subsequent rounds of mammography.

Results

Abnormal mammograms were 5.3 percent of mammograms in the first round and 3.6 percent in subsequent rounds, both proportions declining from the youngest to the oldest age groups (first round: 5.8 percent, 5.6 percent, 4.6 percent, 3.7 percent for women ages 40–49, 50–59, 60–69, and age 70 years

---

* Abnormal mammogram results are those categorized by the American College of Radiology Breast Imaging and Reporting and Data System as *suspicious abnormality*, *highly suggestive of malignancy*, or *assessment incomplete*.[3]

and older, respectively; subsequent rounds: 4.5 percent, 3.5 percent, 3.2 percent, 3.0 percent, respectively.) Abnormal mammograms in women ages 40–49 years were associated with a greater percentage of ultrasound use and needle aspiration than those in older women, while the percentage having a surgical biopsy increased with age. During the first screening round, 1,141 breast cancers were diagnosed among women with an abnormal mammogram; during the subsequent rounds, an additional 109 cancers were diagnosed. Breast cancer rates per 1,000 mammograms were 5.1 for the first round and 2.0 for subsequent rounds; the rates increased with age (first round: 3.6, 5.8, 5.8, 7.3 per 1,000 mammograms for women ages 40–49, 50–59, 60–69, and age 70 years and older, respectively; subsequent rounds: 1.2, 1.8, 2.6, 4.1 per 1,000 mammograms, respectively.) Early-stage cancers (carcinoma *in situ* or Stage 1) accounted for 54.0 percent of first-round cancers and 80.7 percent of subsequent round cancers. In both first and subsequent rounds, the proportions of cancers diagnosed at an early stage among women ages 40–49 were less than those for older women (first round: 50.2 percent, 51.7 percent, 59.4 percent, 59.3 percent for women ages 40–49, 50–59, 60–69, age 70 years and older, respectively; subsequent rounds: 66.7 percent, 78.9 percent, 88.2 percent, 84.2 percent, respectively.) Positive predictive values (PPV) of abnormal mammograms (PPV-AMs) declined from 9.5 cancers per 100 abnormal mammograms in the first round to 5.6 cancers per 100 in the subsequent rounds; for both first and subsequent rounds, the PPV-AMs increased with age. Although the magnitudes were greater, similar patterns were observed for the PPV of biopsies: overall, 25.4 cancers were diagnosed per 100 biopsies in the first round and 20.2 cancers were diagnosed per 100 in the subsequent rounds.

Conclusions

A large, nationwide breast cancer early detection program conducted through hundreds of diverse facilities provided results that indicated that women ages 40–49 years had higher rates of abnormalities on mammograms, but lower detection rates of breast cancer. Women diagnosed with cancer in this age group were less likely to have early stage cancer. The positive predictive values of abnormal mammograms and biopsies were less favorable for women ages 40–49 years than for older women. These results should be useful to clinicians, researchers, and public health personnel in counseling patients, planning new studies, and improving efforts to control breast cancer.

References

1. Henson RM, Wyatt SW, Lee NC. The National Breast and Cervical Cancer Early Detection Program: A comprehensive public health response to two major health issues for women. J Public Health Management Practice 1996;2:36-47.

2. May DS, Lee NC, Nadel MR, Henson RM, Miller DS. The National Breast and Cervical Cancer Early Detection Program: Report on the first four years of mammography provided to medically underserved women. Submitted to JAMA.

3. American College of Radiology. Breast imaging reporting and data system. 2nd ed. Reston, VA: American College of Radiology; 1995.

Screening Outcomes: Clinical Experience with Service Screening Using Modern Mammography

Edward A. Sickles, M.D., F.A.C.R.

1. Is there a reduction in mortality from breast cancer due to screening women ages 40–49 with mammography, with or without physical examination? How large is the benefit? How does this change with age?

The best evidence on reduction in mortality from breast cancer must come from the several randomized controlled trials (RCTs) that already have been conducted. Service screening is performed on entire populations of women, either by invitation or by self-selection, without data collection from randomly selected control groups of non-screened women. Therefore, service screening programs do not generate outcomes data that are sufficiently rigorous to provide convincing evidence on mortality reduction. However, there still is considerable value in the outcomes data from modern service screening programs, because (1) the data from existing RCTs indicate the presence of a mortality reduction, but with only marginal statistical significance; (2) known deficiencies in design and execution of the RCTs may have reduced the extent of observed mortality reduction[1,2] and therefore affected calculations of statistical significance; (3) there have been numerous advances since the conduct of the RCTs in mammographic equipment, technical imaging factors, quality assurance procedures, education of personnel, and mammographic interpretation performance,[1] such that the mammography of the 1990s is demonstrably better than that done when the RCTs were conducted; and (4) the use of surrogate markers of screening benefit[2] permits the demonstration of how the improved mammography of the 1990s can be expected to produce a greater degree of mortality reduction than that already demonstrated in the RCTs, thereby increasing the likelihood that modern mammography truly benefits screened women.

Having thus established the rationale for assessing the benefits of modern service screening mammography, the outcomes data from these programs are presented in the context of Question #3 ("Are There Other Benefits?"), using surrogate markers for mortality reduction.

2. What are the risks of screening women ages 40–49 associated with mammography and with physical examination? How large are the risks? How do they change with age?

Discussion of the nature and magnitude of the risks of screening with mammography and with physical examination will be discussed by others. However, the outcomes data from modern service screening programs do provide relevant information, presented in this section, on how these risks change with age. The risks of screening should be considered separately for three different populations of screened women.

The first population involves all screened women, who are subject to the risks of the screening examination itself (mammography alone, or mammography plus physical examination). Outcomes data from modern service screening programs do not provide useful insights into how these risks change with age.

The second population involves women recalled for additional non-interventional evaluation after abnormal screening examinations. Most such women will be found to have no clinically significant abnormalities; they will be advised to return for routine examination at the appropriate screening

interval. These women experience several risks of false-positive examination (anxiety, inconvenience, discomfort, cost). There also will be some women who eventually are found to have breast cancer, and because the incidence of breast cancer increases with advancing age, fewer women ages 40–49 (when compared to older women) will have true-positive examinations. Outcomes data from modern service screening programs demonstrate that women ages 40–49 are recalled at approximately the same frequency as women screened in later decades of their lives.[3,4] When the outcomes data are examined by 5-year age groupings (and even by 1 year at a time), the same results are found. The above-described age-dependent effect on true-positive examinations (prior probability of having breast cancer) is of very small magnitude because there are many more false-positive than true-positive examinations. Therefore, among women recalled for additional non-interventional evaluation after an abnormal screening examination, the risks of screening are essentially age-independent.

The third population involves women recalled for interventional evaluation (fine needle aspiration biopsy, core biopsy, or surgical biopsy) after abnormal diagnostic examinations. Most of these women will be found to have benign lesions; these women experience several risks of false-positive examination (anxiety, inconvenience, discomfort, scarring, cost) that are of greater magnitude than the risks discussed above. There also will be some women who eventually are found to have breast cancer, and because the incidence of breast cancer increases with advancing age, fewer women ages 40–49 (when compared to older women) will have true-positive examinations. Outcomes data from modern service screening programs demonstrate that women ages 40–49 undergo biopsy at approximately the same frequency as women screened in later decades of their lives.[3,4] When the outcomes data are examined by 5-year age groupings (and even by 1 year at a time), the same results are found. The above-described age-dependent effect on true-positive examinations (prior probability of having breast cancer) is of relatively small magnitude because there are more false-positive than true-positive examinations. Therefore, among women recalled for interventional evaluation after abnormal diagnostic examinations, the risks of screening are essentially age-independent.

One other point merits consideration. In the United States over the past 5 years, there has been a dramatic increase in the number of patients with screening-detected lesions who undergo biopsy by percutaneous sampling rather than by surgical excision. Compared with surgical biopsy, percutaneous sampling is equally accurate but results in much less discomfort, essentially no scarring, and is done at half the cost. When percutaneously biopsied lesions are found to be benign, in most cases surgical biopsy is averted, thereby resulting in substantially reduced risk (morbidity). The trend toward percutaneous rather than surgical biopsy of screening-detected lesions will continue to progress because of its inherent advantages, probably at an accelerated rate as we proceed further into managed-care medicine.

3. Are there other benefits? If so, what are they? How do they change with age?

The most powerful surrogate markers for screening benefit (i.e., those most likely to predict reduced mortality) are tumor size and axillary lymph node status. Cancer stage, which is derived primarily from these two markers, is the penultimate surrogate marker for screening benefit; this marker in fact is so valuable clinically that it is widely used in formulating the treatment plans for breast cancer patients.

There is considerable evidence on the tumor size, lymph node status, and stage of cancers detected during modern service screening mammography. These prognostic factors appear to be at least as favorable for women ages 40–49 as they are for women ages 50–64 (among whom screening is widely accepted as being effective in reducing mortality). In the UCSF service screening program, when screening-detected cancers in women ages 40–49 were compared with those in women ages 50–64, median tumor size was 10 mm versus 11 mm, the lymph node metastasis rate was 12 percent versus 12 percent, and the rate of stage 2 or higher cancers was 19 percent versus 26 percent.[4] Similar results were reported for the province-wide Screening Mammography Program of British Columbia (Canada),[5] and for the county-wide service screening program in Uppsala (Sweden).[3] These surrogate-marker data strongly suggest that if modern service screening is effective in reducing mortality among women ages 50–64, then it should also be effective among women ages 40–49.

There also is substantial evidence that the optimal screening interval for women ages 40–49 is 1 year, rather than the 2-year interval used in most of the RCTs. Analysis of results from both the Kopparberg County portion of the Swedish Two-County RCT,[6] the Uppsala service screening program,[2] and the Cincinnati Breast Cancer Detection Demonstration Project[7] (another, albeit older, service screening program) indicates that the lead time from screening women in their forties is substantially less than that from screening older women. Finally, data from the UCSF service screening program demonstrate a substantial decline in sensitivity for screening women ages 40–49 when the screening interval is increased from 1 year to 2 years (83.6 percent versus 71.4 percent).[8] The UCSF data further suggest that these findings are due to more rapid tumor growth in younger women rather than the somewhat increased breast density found in the breasts of younger women, as does mathematical modeling of the Kopparberg[9] and entire two-county[10] RCT data.

One further useful piece of evidence can be derived from the UCSF service screening program. By comparing the outcomes from initial versus subsequent screening examinations, it is demonstrated that the recall rate (frequency of abnormal screening interpretation) is substantially higher yet the surrogate markers (tumor size, lymph node status, cancer stage) appear to be less favorable for initial screening examinations than for subsequent examinations.[11] These results apply equally to women ages 40–49 and to older women. Note that ongoing service screening will involve many subsequent screening examinations but only one initial examination. Thus, outcomes data based either entirely or predominantly on initial screening will tend to underestimate some of the benefit and overestimate some of the risk of ongoing screening. For example, a previously published report of UCSF data on the positive predictive value (PPV) of screening by age showed a significantly lower PPV for screening women ages 40–49 versus ages 50–59 for initial screening examinations, but no difference at all between these age groups for subsequent screening.[12]

There are other benefits of screening women ages 40–49 years apart from those indicated by the surrogate-marker evidence cited above. These range from the reassurance gained from knowledge that a screening examination was normal to the greater likelihood of being eligible for breast

conservation therapy when cancer is detected by screening versus usual care. However, the outcomes data from modern service screening programs do not provide evidence to document such benefits, so that discussion of these benefits is beyond the scope of this presentation.

4. What is known about how the benefits and risks of breast cancer screening differ based on known risk factors for breast cancer?

The RCTs were not designed to provide separate data on subpopulations of women at higher than average risk for breast cancer, and therefore no evidence on mortality reduction can be expected. However, outcomes data from the UCSF service screening program, using surrogate markers, does provide some indirect evidence for women ages 40–49 years. (1) The PPV of screening is higher in women who have a strong or very strong family history of breast cancer than in the remainder of screened women in this age group.[12] However, it is likely that this simply is due to the increased incidence of breast cancer in these high-risk women (greater prior probability of cancer), rather than being due to an improved ability of screening to detect cancer in high-risk women. In addition, the degree of difference in PPV between high- and average-risk women is only slightly higher when comparing women ages 47–49 years with women ages 50–52 years, paralleling the only slightly increased incidence of breast cancer in the slightly older age group. (2) Similarly, as one would expect because of the greater prior probability of cancer, the biopsy yield of cancer is higher in high-risk women ages 40–49 than in the remainder of screened women in this age group. Also, there are no abrupt changes in biopsy yield for high- versus average-risk women in comparing age groups just below 50 years with age groups just above 50 years. There are essentially no differences in the size, lymph node status, and stage of screening-detected breast cancers when comparing high-risk women ages 40–49 years with the remainder of screened women in this age group. Had screening among women ages 40–49 years been limited to the 12 percent at high risk by family history, this strategy would have detected only 19 percent of the extant cancers.[4] The overall conclusion to be drawn from this experience is that for the age range 40–49 years, modern service screening mammography appears to detect breast cancer equally effectively in women with differing risk for breast cancer, that the increased incidence of breast cancer in high-risk women will improve the cost-effectiveness of screening these high-risk women, but that the more cost-effective strategy of screening only high-risk women will relinquish to usual-care detection more than 80 percent of the cancers in the entire population.

5. What are the directions for future research?

There have been numerous advances in conventional mammography over the past 10 years, involving equipment, technical imaging factors, quality assurance procedures, education of personnel, and mammographic interpretation performance. Continued advances are expected as we enter the 21st century. There also is promising and very important research involving digital mammography, high-resolution breast ultrasound, magnetic resonance imaging, and isotope-scanning of the breast. Among these imaging techniques, digital mammography may provide increased sensitivity and/or specificity when used for breast cancer screening. All techniques may permit increased sensitivity and/or specificity in the “diagnostic” setting (i.e., providing non-interventional evaluation of screening-detected abnormalities).

In contrast to breast imaging, which has undergone (and continues to undergo) many improvements, very little change has occurred in the practice of breast physical examination, other than the realization that it appears to be more accurate when performed with diligence by specially trained practitioners. Unfortunately, there currently is little enthusiasm either within or external to the

medical community to improve the current state of breast physical examination in the United States. Two approaches that are likely to reap considerable benefit are (1) the recruitment, training, and deployment of large numbers of paramedical personnel to perform breast physical examination in screening centers, and (2) Federal legislation mandating quality assurance practices for breast physical examination, to parallel the provisions of the Mammography Quality Standards Act of 1992 (which has resulted in considerably improved delivery of high-quality mammography services).

The National Cancer Institute has funded a multisite Breast Cancer Surveillance Consortium, which is currently collecting outcomes data from more than 1 million women on many aspects of modern breast cancer screening practice. This research will provide valuable direction into methods of improving breast cancer screening in the United States. However, there is urgent need to go beyond this effort by creating a national cancer registry, to permit collection of meaningful outcomes data for all American women. To be truly effective, such a cancer registry must permit low-cost data linkage by individual breast cancer screening practices, so that complete rather than partial outcomes data are available at the service-provider level for the purpose of continuous quality improvement.

References

1. Sickles EA, Kopans DB. Deficiencies in the analysis of breast cancer screening data (editorial). J Natl Cancer Inst 1993;85:1621–4.

2. Feig SA. Determination of mammographic screening intervals with surrogate markers for women aged 40-49 years (editorial). Radiology 1994;193:311–4.

3. Thurfjell EL, Lindgren JA. Population-based mammography screening in Swedish clinical practice: prevalence and incidence screening in Uppsala county. Radiology 1994;193:351–7.

4. Curpen BN, Sickles EA, Sollitto RA, Ominsky SH, Galvin HB, Frankel SD. The comparative value of mammographic screening for women 40–49 years old versus women 50-64 years old. Am J Roentgenol 1995;164:1099–1103.

5. Warren Burhenne LJ, Burhenne HJ, Kan L. Quality-oriented mass mammography screening. Radiology 1995;194:185–8.

6. Tabar L, Fagerberg G, Day NE, et al. What is the optimum interval between screening examinations? An analysis based on the latest results of the Swedish two-county breast cancer screening trial. Br J Cancer 1987;55:547–51.
7. Moskowitz M. Breast cancer: age-specific growth rates and screening strategies. Radiology 1986;161:37–41.

8. Kerlikowske K, Grady D, Barclay J, Sickles EA, Ernster V. Effect of age, breast density, and family history on the sensitivity of first screening mammography. JAMA 1996;276:33–8.

9. Tabar L, Fagerberg G, Chen HH, Duffy SW, Gad A. Tumour development, histology and grade of breast cancers: prognosis and progression. Int J Cancer 1996;66:413–9.

10. Tabar L, Fagerberg G, Chen HH, et al. Efficacy of breast cancer screening by age. Cancer 1995;75:2507–17.

11. Frankel SD, Sickles EA, Curpen BN, Sollitto RA, Ominsky SH, Galvin HB. Initial versus subsequent screening mammography: comparison of findings and their prognostic significance. Am J Roentgenol 1995;164:1107–9.

12. Kerlikowske K, Grady D, Barclay J, Sickles EA, Eaton A, Ernster V. Positive predictive value of screening mammography by age and family history of breast cancer. JAMA 1993;270:2444–50.

Outcomes of Modern Screening Mammography

Karla M. Kerlikowske, M.D.

Randomized controlled screening mammography trials have not clearly demonstrated a reduction in breast cancer mortality for women ages 40–49, at least not for the first 7–9 years after the initiation of screening.[1] Screening mammography may not be effective for women ages 40–49, in part because mammography is less sensitive in younger women. Proponents of screening mammography in women ages 40–49 criticize the technology used in the randomized controlled trials of screening mammography, implying that modern screening mammography is better able to detect breast cancer in younger women. However, modern mammography has yet to demonstrate an improved sensitivity among younger women. Studies of modern screening mammography[2–6] report overall sensitivities of screening mammography (71.1–91.5 percent) similar to those published for clinical trials.[7] Two studies report the sensitivity of mammography by age and show that sensitivity is still lower for women less than 50 years old (63 percent and 80 percent) compared with women age 50 and older (89 percent and 94 percent).[4,5] A recent study that evaluated the sensitivity of mammography by decade of age showed that the sensitivity of mammography to detect invasive breast cancer was lower among women ages 40–49 compared with women age 50 and older (75 percent versus 92 percent; see Table 1) and suggested that the lower sensitivity of mammography in younger women may to be due to more rapid tumor growth in these women.[8] Of note, modern screening mammography has improved in its ability to detect invasive cancer in women age 50 and older and ductal carcinoma *in situ* (DCIS) in all women.

TABLE 1. Sensitivity of Modern First Screening Mammography[8]

Sensitivity	Age (y)		
	40–49	50–59	60–69
Invasive Cancer Only (%)	75.0	92.3	93.2
(95% CI)	(52.9–89.4)	(78.0–98.0)	(80.3–98.2)
DCIS	100	100	100
All Breast Cancer (%)	86.7	93.6	94.1
(95% CI)	(72.5–94.5)	(81.4–98.3)	(82.8–98.5)

Even though there is not definitive proof that screening mammography reduces breast cancer mortality in younger women and the ability to detect invasive cancer is less than that for older women, why not do it anyway? The main reasons not to use unproven screening tests are (1) the burden of unnecessary workups of false-positive examinations with associated morbidity, anxiety, and cost; (2) the potential to detect lesions that may be clinically insignificant yet are treated; and (3) the false reassurance resulting from having a normal screening examination.

Workups for False-positive Examinations

Nationwide, about 11 percent of all screening examinations are read as abnormal (range=3–57 percent), with the average positive predictive value (PPV—proportion of women with abnormal screening examinations who have breast cancer) of mammography for women age 50 and older. This is around twofold that of women below age 50 (4.7 versus 2.0).[9] Even in institutions like the University of California, San Francisco (UCSF), with well-trained, full-time mammographers, about 7 percent of first screening mammography examinations are read as abnormal and the PPV of mammography is low.[10] For example, the PPV of first screening mammography at UCSF for women ages 40–44 and those 45–49 is about 5 percent. This means that among 100 average-risk women ages 40–49 with an abnormal first screening examination, about 95 do not have cancer (Table 2) and must undergo further evaluation that may include clinical breast examination, additional mammography, ultrasound, needle aspiration, or excisional biopsy. On average, approximately two additional diagnostic tests are performed per abnormal screening examination. Because many mammographic abnormalities are nonpalpable, needle localization biopsy is often required. Although risk is low, there are complications associated with biopsies, such as hematomas, infection, and scarring; from wire localization itself, complications include vasovagal reactions (7 percent) and, rarely, prolonged bleeding (1 percent) and extreme pain (1 percent). Assuming the higher level of mammography performance obtained at UCSF, our data suggest that if 1,000 women ages 40–49 undergo screening mammography annually for 10 years, approximately 352 will have an abnormal finding requiring some additional procedure (including 81 biopsies); 20 will have cancer, 11 of which will be invasive cancer and 9 DCIS. In comparison, if 1,000 women ages 50–59 undergo screening mammography annually for 10 years, approximately 293 women will have an abnormal finding requiring some additional procedure (including 82 biopsies); 42 will have cancer, 30 of which will be invasive and 12 DCIS. Thus, women ages 40–49 will undergo the same number of biopsies to diagnose half as many breast cancers compared with women ages 50–59, and will have 2.5 times as many diagnostic procedures (33 versus 14) for every DCIS or invasive cancer diagnosed or 3 times as many procedures (60 versus 20) for every invasive cancer diagnosed. The lower yield of cancer per breast biopsy and higher number of diagnostic tests per cancer detected in women ages 40–49 are because of the lower incidence of breast cancer in these women. The number of cancers diagnosed per 1,000 first screening examinations increases with age from 3 in women ages 40–49 to 12 in women ages 60–69 (Table 2). It is important to emphasize that these numbers are based on a 7-percent abnormal rate for first screening and 2 percent for subsequent screening, which is low compared with the national average of 11 percent for all types of screening exams. Thus, the numbers presented are a conservative estimate of the number of procedures conducted per cancer detected.

Finding Clinically Insignificant Lesions

DCIS is a proliferation of cells with malignant features that is confined within the mammary ducts. Of breast cancers detected by screening mammography in average-risk women ages 40–49, approximately 44 percent are DCIS compared with 20–30 percent of those detected in women age 50 and older.[10] Given that the natural history of DCIS is unknown (in particular, the natural history of small mammographically detected lesions), the current clinical dilemma lies in not being able to distinguish which lesions will progress to invasive cancer. Thus, screening mammography may be benefiting some women through early detection of potentially fatal breast cancers, while it is potentially harming other women through detection of clinically insignificant lesions that, for lack of good prognostic indicators, are almost always treated surgically.

TABLE 2. Positive Predictive Value of First Screening Mammography and Yield of Breast Biopsy*			
	AGE (y)		
	40–49	50–59	60–69
Normal Exams (%)	12,312 (93.6)	6,648 (93.2)	4,168 (92.0)
Abnormal Exams (%)	848 (6.4)	489 (6.8)	362 (8.0)
Breast Cancers/1,000 Exams	3	6	12
(95% CI)	(2, 4)	(4, 8)	(9, 15)
PPV Mammography			
All Breast Cancer (%)	4.6	9.0	16.7
(95% CI)	(3.2–6.0)	(6.5–11.5)	(13.0–20.4)
Invasive Cancer Only (%)	2.6	6.3	12.2
(95% CI)	(1.5–3.7)	(4.1–8.5)	(8.8–15.6)
Breast Cancer/Biopsy			
All Breast Cancer (%)	20	32	42
(95% CI)	(14–26)	(24–40)	(34–51)
Invasive Cancer Only (%)	11	22	34
(95% CI)	(7–15)	(15–39)	(26–42)

*Data from University of California Mobile Mammography Screening Program, 1985–1996

False Negatives

Of 100 women ages 40–49 with invasive breast cancer, about 75 cases will be detected by mammography and 25 will go undetected, compared with 92 detected and 8 undetected in women ages 50–59.[8] This means that, potentially, 25 women ages 40–49 with invasive breast cancer will be told their screening examination is normal and may be falsely reassured that they do not have breast cancer. As a result, they do not seek medical attention for breast symptoms. However, for women who do not have breast cancer, they may be reassured by having a normal screening examination that they do not have breast cancer. The annual risk of breast cancer for a 40-year-old woman is about 1 in 625; having a normal screening examination decreases her risk to about 1 in 2,500. Although the very low risk of breast cancer after a normal screening examination may reassure women that they do not have breast cancer, the risk of breast cancer *before* mammography is already quite low. Thus, screening mammography is not justified solely to reassure women that breast cancer is not present.

In summary, there are associated risks with undergoing screening mammography, including additional diagnostic evaluations and the associated morbidity and anxiety, the potential for detecting clinically insignificant breast lesions, and the false reassurance resulting from having a normal

examination. Before mass screening is recommended to healthy persons, the benefits of the intervention should be proven and should clearly outweigh the risks. Rather than receive recommendations for or against screening mammography when the benefits are uncertain yet the risks known, women should be informed of the risks, potential benefits, and limitations of the test. They will then be able to make informed decisions based on their personal risk status and on the associated risks and benefits of screening mammography.

References

1. Kerlikowske K, Grady D, Rubin S, Sandrock C, Ernster V. Efficacy of screening mammography: a meta-analysis. JAMA 1995;273:149–54.

2. Bird RE. Low-cost screening mammography: report on finances and review of 21,716 consecutive cases. Radiology 1989;171:87–90.

3. Linver MN, Paster SB, Rosenberg RD, Key CR, Stidley CA, King WV. Improvement in mammography interpretation skills in a community radiology practice after dedicated teaching courses: 2-year medical audit of 38,633 cases. Radiology 1992;184:39–43.

4. Burhenne HJ, Burhenne LW, Goldberg F, et al. Interval breast cancers in the screening mammography program of British Columbia: analysis and classification. Am J Roentgenol 1994;162:1067–71.

5. Robertson CL. A private breast imaging practice: medical audit of 25,788 screening and 1,077 diagnostic examinations. Radiology 1993;187:75–9.

6. Sienko DG, Hahn RA, Mills EM, et al. Mammography use and outcomes in a community: the greater Lansing area mammography study. Cancer 1993;71:1801–9.

7. Fletcher SW, Black W, Harris R, Rimer B, Shapiro S. Report of the international workshop on screening for breast cancer. J Natl Cancer Inst 1993;85:1644–56.

8. Kerlikowske K, Grady D, Barclay J, Sickles EA, Ernster V. Effect of age, breast density, and family history on the sensitivity of first screening mammography. JAMA 1996;276:33.

9. Brown ML, Houn F, Sickles EA, Kessler LG. Screening mammography in community practice: positive predictive value of abnormal findings and yield of follow-up diagnostic procedures. Am J Roentgenol 1995;165:1373–7.

10. Kerlikowske K, Grady D, Barclay J, Sickles EA, Eaton A, Ernster V. Positive predictive value of screening mammography by age and family history of breast cancer. JAMA 1993;270:2444–50.

Mammography Outcomes in a Practice Setting by Age: Prognostic Factors, Sensitivity, and Positive Biopsy Rate

Michael N. Linver, M.D., F.A.C.R.

Introduction and Purpose

The value of regular screening mammography in reducing breast cancer mortality has been amply demonstrated in multiple randomized controlled trials (RCT's) for women 50 and older. However, owing to limitations of study design, equivalent evidence is not available from RCT's for women 40–49. A similar mortality reduction has been seen in this age group in most RCT's, but due to inadequate statistical power, point estimates have not achieved statistical significance.[1] It is only through meta- analysis of these RCT's that a statistically significant difference in mortality between women screened and not screened with mammography has been demonstrated.[1] A single, definitive randomized trial in the United States to test screening efficacy in the 40–49 group, large enough to give statistical significance to the results, not only would be difficult (at least 1.5 million women would need to be enrolled) but also would not yield meaningful results for another 10–15 years.[2] A trial requiring fewer women has been proposed in Europe but has not yet begun.

Given these difficulties with RCT analysis, many have suggested and employed surrogate endpoints to assess screening efficacy.[2–5] We have undertaken such an analysis in a private practice setting.

Materials and Methods

Our group comprises 12 general radiologists, 2 of whom have a special interest in mammography. All 12 radiologists interpret mammograms, each interpreting between 1,000 and 8,000 mammograms a year. Altogether, we interpret approximately 40,000 mammograms yearly, 90 percent of which are screening studies, utilizing 4 private outpatient offices, 2 community hospitals, and 2 mobile vans. Through a computerized reporting system we designed and have utilized since 1988, we performed an audit of more than 104,000 mammograms performed on women over age 40 between February 1988 and January 1993. Surrogate endpoints chosen to evaluate screening efficacy were tumor size and axillary lymph node status in screening-detected cancers in women ages 40–49, and in the group age 50 and over. These prognostic factors are the biological underpinnings that distinguish those women whose prognosis is more favorable in the RCT's. In addition, because we have access to computer linkage with a Statewide tumor registry, which allows successful computer data matches in 95 percent of cases, we have evaluated the accuracy of mammography in each group by tracking all false negatives and comparing resultant sensitivity values. Furthermore, we evaluated mammography efficiency in detecting cancers by calculating Positive Biopsy Rates (PBR's) (biopsies positive for cancer/all biopsies performed based on mammographic recommendation for biopsy) in each age group. (*P*-values for data sets obtained were calculated with Stat.Xact-Turbo software [Cytel Corporation, C. Mehta and N. Patel, 1992].)

Results

We successfully diagnosed 186 cancers in the 40–49 group and 695 in the over 50 group. Approximately two-thirds in each group were screening-detected (see Table 1). When only the screening-detected cancers were evaluated, the cancer detection rate was 3.8 per 1,000 screening cases for the 40–49 group and 7.5 for the over 50 group. These rates are proportional to those expected in these age groups, based on SEER data.[6]

TABLE 1. Cancers Detected

	40–49 Group	Over 50 Group
Total Mammographic Exams	35,480	68,891
Screening Mammographic Exams	31,758	61,657
All Cancers Found Mammographically	186	695
Asymptomatic (screening-detected) Cancers Found Mammographically	121	464
Cancer Detection Rate*	3.8	7.5

*Number of cancers detected per 1,000 screening examinations.

Review of prognostic factors for screening-detected cancers showed that 75 percent were minimal cancers (ductal carcinoma *in situ* or invasive cancers 1 cm or less) in the 40–49 group, compared with 60 percent in the over 50 group (see Table 2). Median size of the invasive cancers was 1.0 cm in the 40–49 age group and 1.1 cm in the over 50 group. Axillary lymph node positivity was 14.8 percent in the 40–49 group and 11.9 percent in the over 50 group.

TABLE 2. Screening-Detected Cancers: Size and Nodal Status

	40–49 Group	Over 50 Group
Screening-Detected Cancers, Total	121	464
DCIS	54 (45%)	111 (24%)
Invasive Cancers	67 (55%)	353 (76%)
Minimal Cancers (DCIS or œ 1 cm)	75%	60%
Median Size (invasive cancers only)	1.0 cm	1.1 cm
Axillary Lymph Node Positivity	14.8%	11.9%

Axillary lymph node status of small screening-detected invasive cancers 1 cm or less in size yielded similarly low node positivity values of 3 percent for the 40–49 group and 5 percent for the over 50 group (see Table 3).

TABLE 3. Axillary Lymph Node Positivity—Invasive Cancers œ 1 cm (screening-detected cancers only)

	40–49 Group	Over 50 Group
Cancers œ 1 cm	27	125
Positive Axillary Lymph Nodes	1 (3%)	6 (5%)

Using the 1-year definition of false negative (i.e., detection of cancer within 1 year of mammographic examination with normal findings), we calculated sensitivity values of 85.3 percent for the 40–49 group and 87.7 percent for the over 50 group (see Table 4). These values were found

not to be significantly different (P=.36).

TABLE 4. Sensitivity and Positive Biopsy Rate

	40–49 Group	Over 50 Group
Total Mammographic Exams	35,480	68,891
Biopsies Done Based on Mammographic Findings	689	1,986
Cancers Found at Biopsy (and correctly identified mammographically)	186	695
False Negative Cases	32	97
Overall Sensitivity	85.3%	87.7%
Positive Biopsy Rate	27%	35%

PBR's were 27 percent in the 40–49 group and 35 percent in the over 50 group (Table 4). This difference is statistically significant (P=.0001), but within the range of target values endorsed in the AHCPR *Clinical Practice Guideline on Quality Determinants of Mammography*.[7]

Discussion

Day and others have argued persuasively that intermediate measures are useful for evaluation of a screening program, serving as proxies for conventional endpoints such as death from breast cancer.[4,5] The surrogate endpoints evaluating screening efficacy in the 40–49 group versus the over 50 group chosen here, tumor size and axillary lymph node positivity, have both been shown to correlate inversely with survival[8,9]: When tumor size is small and node positivity is absent, survival is much greater in all age groups over 40. Our findings reflect favorable measures for both parameters in women 40–49 (Table 2) and show no significant difference between the 40–49 group and the over 50 group, paralleling closely those in other recent studies.[8–11] These data would imply that, as demonstrated by Tabar[8] and Therfjell and Lindgren,[9] women in both age groups have an equally high likelihood of long survival when a small tumor is detected by screening. The differences in these same factors explain the differences in mortality found in the RCT's.

We found an even more impressive prognostic indicator in the extremely low axillary lymph node positivity (3–5 percent) in both groups when evaluating screening-detected invasive cancers 1 cm or smaller (Table 3). Our findings were virtually identical to those of Curpen and Sickles,[10] and support the hypothesis that advancing the time of diagnosis at any age reduces the likelihood for axillary lymph node metastasis, thus improving prognosis.

One could anticipate that this evidence would further translate into a reduction in mortality for all women screened at age 40 and older, although the many biases intrinsic in the use of survival data warrant caution.[3] Nevertheless, long-term survival provides confidence that a benchmark of cure has been achieved. The fact that our findings also parallel those reported by Curpen and Sickles[10] would seem to support their reproducibility outside the academic setting.

Furthermore, our detection of smaller, node-negative breast cancers was accomplished with a high degree of accuracy, regardless of patient age: A sensitivity in the 85 percent range was achieved in both age groups (Table 4). This finding would appear to refute the contention made by some that the value of screening under age 50 is compromised by markedly lower sensitivity.[12] Certainly, our data suggest that concern about the risk of subjecting women 40–49 to a procedure with a low

likelihood of detecting a tumor at an earlier stage, with a more favorable prognosis, appears to be unfounded.

Note is made of our finding a lower PBR of 27 percent in the 40–49 age group, compared with the 35 percent in the over 50 group. The 8 percent margin found here is not a remarkable difference and is primarily due to the demonstrated lower incidence of breast cancer (3.8/1000) in the 40–49 group compared with the over 50 group (7.5/1000) in our study (Table 1) and in those of others.[5,10] Although the greater number of biopsies yielding benign findings in the 40–49 group could well be construed as a risk of screening, the 27 percent PBR we achieved here is still very acceptable, given the high rate of small, node-negative tumors we detected through screening. This is a trade-off we have accepted in our practice and within our community, in view of the derived measurable benefits demonstrated.

Future research to offer still more support for the conclusions reached here should center on efforts to encourage others in academia and private practice to perform and publish similar audits of their screening mammography practices. These efforts could be substantially strengthened by passage of national legislation to protect audit data from discovery. Such protection does not presently exist in most states.[13] These audit data are especially important in light of the absence of statistically significant RCT data in women 40–49. Even if such RCT data were available, it is no less important to evaluate the performance of mammography in the community setting.

Conclusion

We find modern screening mammography in the private practice setting to be as successful in detecting breast cancers with favorable prognostic factors in women age 40–49 as in women over 50. Our findings parallel the recent favorable results of others who have similarly evaluated screening efficacy via surrogate endpoints.[5–8] Furthermore, we find these results to be achievable without the risk of generating a comparatively unacceptable large number of biopsies and, at the same time, without subjecting women age 40–49 to a procedure that has any substantially lower sensitivity than it has in women over age 50.

References

1. Smart CR, Hendrick RE, Rutledge JH III, Smith RA. Benefit of mammography screening in women ages 40–49 years: current evidence from randomized controlled trials. Cancer 1995;75:1619–26. [Erratum: Cancer 1995;75:2788].

2. Eckhardt S, Badellino F, Murphy GP. UICC meeting on breast cancer screening in pre-menopausal women in developed countries. Int J Cancer 1994;56:1–5.

3. Feig, SA. Determination of mammographic screening intervals with surrogate measures for women aged 40–49 years. Radiology 1994;193:311–4.

4. 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.

5. Day NE. Quantitative approach to the evaluation of screening programs. World J Surg 1989;13:3–8.

6. American Cancer Society. Cancer facts and figures—1995. Atlanta: American Cancer Society, 1995.

7. Bassett LW, Hendrick RE, Bassford TL, et al. Quality determinants of mammography. Clinical practice guideline no. 13. AHCPR publication no. 95-0632. Rockville, MD: U.S. Department of Health and Human Services, Public Health Service, Agency for Health Care Policy and Research. October 1994.

8. Tabar L, Duffy SW, Warren Burhenne L. New Swedish breast cancer detection results for women aged 40–49. Cancer 1993;72:1437–48.

9. Therfjell EL, Lindgren JAJ. Breast cancer survival rates in women younger and in those older than 50 years: effect of mammography screening. Radiology 1996;201:421–6.

10. Curpen BN, Sickles EA, Solitto RA, Ominsky SH, Galvin HB, Frankel SD. The comparative value of mammographic screening for women 40–49 years old versus 50–64 years old. Am J Roentgenol 1995; 164:1099–103.

11. Warren Burhenne LJ, Burhenne HJ, Kan L. Quality-oriented mass mammography screening. Radiology 1995;194:185–8.

12. Fletcher SW, Black W, Harris R, Rimer BK, Shapiro S. Report of the International Workshop on Screening for Breast Cancer. J Natl Cancer Inst 1992;85:1644–56.

13. Linver MN, Rosenberg RD, Smith RA. Mammography outcomes analysis: potential panacea or Pandora's box? (commentary). Am J Roentgenol 1996;167:373–5.

Radiation Risk

Stephen A. Feig, M.D., and R. Edward Hendrick, Ph.D.

Estimation of the risk of radiation-induced breast cancer is a consideration in determining the advisability of mammographic screening for women of any age group and may be especially important for women ages 40–49 years. Due to their relatively lower breast cancer incidence it is particularly important that lives saved vs deaths caused and years of life expectancy gained/lost through screening be quantitatively compared for these younger women.

Risk Assessment

Although no women have ever been shown to have developed breast cancer as a result of mammography, not even from multiple examinations received over many years at mean glandular doses considerably higher than the current one of less than 5 mGy (500 mrad), the possibility of such risk exists because excess breast cancers have been observed among populations receiving much higher doses of 250 mGy - 20 Gy (25-2,000 rads). These include Japanese A-bomb survivors,[1] North American tuberculosis sanitoria patients from Massachusetts[2] and Canada[3] who underwent multiple chest fluoroscopies, women from New York State[4] and Sweden[5] treated with radiation therapy for benign breast conditions such as postpartum mastitis, and women who had been treated in California with radiation therapy for Hodgkin's Disease.

Estimating the risk of breast cancer from low-dose radiation is complex. However, relatively similar estimates have been made by various committees over the past 20 years, most notably by the National Cancer Institute (NCI) Ad Hoc Working Group on the risks associated with mammography and mass screening for the detection of breast cancer (1977)[7] by the Committee on the Biological Effects of Ionizing Radiation (BEIR III) of the National Academy of Sciences (1980),[8] the National Institutes of Health Ad Hoc Group to Develop Radiolepidemiological Tables (1985)[9] and the National Academy of Sciences - National Research Council Committee on the Biological Effects of Ionizing Radiation (BEIR V) in 1990.[10] The most recent risk estimate was made by the Radiation Effects Research Foundation (RERF) in 1994.[1] Each committee has had to base their estimate not only on the data available at that time from all epidemiologic studies (with the exception of the RERF estimate based only on the Japanese A-bomb survivors, the largest of all exposed study populations) but also on a selection of other assessment options such as dose-response models, length of latent period, duration of radiation effect, age-at-exposure-related radiation sensitivity, and additive versus relative risk models.

Dose–Response Models

Because radiation-induced and spontaneously occurring breast cancers cannot be distinguished histologically,[11,12] the presence of radiation-induced tumors can only be established statistically if a significant number of excess cancers is observed in an exposed population. This type of inference becomes harder and harder to establish as lower and lower doses are considered since the number of exposed women required to demonstrate risk is related to the inverse square of dose. If there is any risk to the breast from doses in the mammographic range or even from doses of 100mGy (10 rad) or less, the magnitude of the risk may be estimated by means of dose-response curves which describe the possible relationship between radiation dose and radiogenic cancer incidence. Because the linear

dose response model in which risk/rad remains constant down to zero dose represents the upper limit of risk, it is the model most often employed to estimate risk at low doses. Lower risk estimates are obtained with either quadratic or linear-quadratic dose-response relationships.

Latent Period and Duration

The latent period refers to the minimal length of time between exposure and earliest demonstration of excess cancers in the population. Most reports have assumed latent periods of at least 10–15 years and a lifetime persistence of radiation effect in the exposed population.

Age at Exposure

All but one of the studies found decreased risk with increasing age at exposure. Indeed, the BEIR V report concluded: "there is little evidence of any increased risk to women exposed after age 40".[10]

Additive and Relative Risk Models

These represent two different methods of estimating excess risk (defined as either excess breast cancer incidence or mortality) following radiation. Additive (absolute) risk estimates are given as a number of excess cancers/million women/year/cGy (rad). Relative risk estimates are given as the percentage increase in the natural breast cancer incidence/year/cGy (rad). BEIR V used a time-dependent relative risk model in which relative risk varied over time during the followup, reaching a peak at 15–20 years after exposure and then declining.[10]

Quantifying Benefits and Risks

Using the 1985 NIH relative risk estimate, Feig and Ehrlich found that a single screen of women at ages 40–44 and 45–49 with a dose of 2.5mGy and 20 percent mortality reduction would result in benefit/risk ratios of 35 and 90 years of life expectancy gained/loss respectively.[13]

Using the 1990 BEIR V relative risk estimate, Feig et al.[14] calculated that a single mammographic screening of women at age 45 with a dose of 2.5 mGy and mortality reductions of 20 percent and 40 percent respectively would avert 30 and 60 deaths per death caused respectively. Assuming that some radiogenic cancers would be detected by subsequent screening, the benefit risk ratios from the single screen would be 37.5 and 100 respectively at the same levels of benefit.

Based on the 1994 RERF relative risk estimate, Mettler et al developed benefit/risk ratio tables comparing fatal cancers of breast cancer prevented/caused by mammographic screening.[15] A mortality reduction of 15 percent for screening women ages 40–49 and 25 percent for screening women ages 50–75 was assumed along with the dose of 2.8 mGy. As an example, if a women began annual mammography at age 40, mammographic examination at age 44 would provide 850 times more benefit than the potential harm from all of her mammographic examinations combined. Since annual screening with modem mammography should result in even greater mortality reduction, benefit/risk ratios could be even higher than those suggested in their Study.[16]

More recently, it has been suggested that the mean glandular dose from mammography could be 4mGy per two-view examination due to a larger estimate for compressed breast thickness (5.0–5.7cm vs. 4.2cm)[17] and use of higher optical densities (1.4–1.8 vs 1.3) to provide even better exposure of mammographic images. However, even at this higher dose, favorable benefit/risk ratios should be

achieved with any reasonable level of benefit from screening women ages 40–49. Moreover, higher optical densities have been shown to result in earlier detection of cancer[18] so that there should be a concomitant improvement in benefit.

Radiation Risk and Other Risk Factors

Possible effect of other risk factors on radiation risk is incompletely known and for some risk factors may be extremely difficult to evaluate. Although American women have a higher breast cancer incidence than Japanese women, probably due to diet and other environmental factors, absolute breast cancer risk from radiation is similar when both populations are compared, but relative risk factors are markedly different.[19]

Older age is a major risk factor for breast cancer, yet there is an inverse relationship between radiation sensitivity and age and exposure.

It has been suggested but not demonstrated that ataxia-telangiectasia gene heterozygotes may have increased sensitivity to radiation carcinogenesis,[20] but the study had methodologic flaws[21–24] and these individuals comprise less than 1 percent of the U.S. population.

Inherited mutations in the BRCA 1 and BRCA 2 genes may be involved in 5–10 percent of breast cancers. Due to their very high baseline breast cancer incidence and the fact that they represent a relatively small proportion of the general population, meaningful studies of radiation sensitivity in these women might not be feasible.

Subgroup analysis of radiation-sensitivity high-risk women should not become a distraction from the overriding conclusion that for the general population of women in the breast cancer age group, the theoretical radiation risk from screening mammography is negligible compared with the known benefit.

References

1. Tokunaga M, Land CE, Tokuoka S, Nishimori I, Soda M, Akiba S. Incidence of female breast cancer among atomic bomb survivors, 1950-1985. Radiation Research 1994; 138:209–23.

2. Hrubec Z, Boice JD, Monson RR, Rosenstein R. Breast cancer after multiple chest fluoroscopies: Second follow-up of Massachusetts women with tuberculosis. Cancer Res 1989;49:229–34.

3. Miller AB, Howe GR, Sherman GJ, et al. Mortality from breast cancer after irradiation during fluoroscopic examination in patients being treated for tuberculosis. N Eng J Med 1989;321:1285–9.

4. Shore RE, Hildreth N, Woodard ED, Dvoretsky P, Hempehnann L, Pasternack B. Breast cancer among women given x-ray therapy for acute post partum mastitis. J Natl Cancer Inst 1986;77:689–96.

5. Mattson A, Bengt-Inge R, Hall P, Wilking N, Rutquist LE. Radiation-induced breast cancer: Long-term follow-up of radiation therapy for benign breast disease. J Natl Cancer Inst 1993;85:1679–85.

6. Hancock SL, Tucker AM, Hoppe RT. Breast cancer after treatment of Hodgkin's disease. J Natl Cancer Inst 1993;85:25–31.

7. Upton AC, Beebe GW, Brown JM, Quimby EH, Shellabarger C. Report of the NCI Ad Hoc Working Group on risks associated with mammography in the mass screening for the detection of breast cancer. J Natl Cancer Inst 1977;59:481–93.

8. BEIR III Committee on the Biological Effects of Ionizing Radiation. The effects on populations of exposure to low levels of ionizing radiation. Washington DC: National Academy of Sciences, 1980.

9. National Institutes of Health Ad Hoc Working Group to Develop Radioepidemiological Tables. Report of the National Institutes of Health Ad Hoc Working Group to Develop Radioepidemiological Tables. NIH Publication No. 85-2748. Bethesda MD: National Institutes of Health, National Cancer Institute, 1985.

10. BEIR III Committee on the Biological Effects of Ionizing Radiation. Health effects of exposure to low levels of ionizing radiation. Washington DC: National Academy Press, 1990.

11. Dvoretsky PM, Woodard E, Bonfiglio TA, et al. The pathology of breast cancer in women irradiated for acute postpartum mastitis. Cancer 1980;46:2257–62.

12. Tokuoka S, Asano M, Tsutomu Y et al. Histologic review of breast cancer cases in survivors of atomic bombs in Hiroshima and Nagasaki, Japan. Cancer 184;54:849–54.

13. Feig SA, Ehrlich SM. Estimation of radiation risk from screening mammography: Recent trends and comparison with expected benefits. Radiology 1990;174:638–47.

14. Feig SA, Dodd GD, Hendrick RE. Mammography risks and benefits. In: Radiation Protection in Medicine, Proceedings of the Twenty-eighth Annual Meeting of the National Council on Radiation Protection and Measurements, Proceedings No. 14. Bethesda MD: National Council on Radiation Protection and Measurements 1993;240–53.

15. Mettler FA, Upton AC, Kelsey CA, et al. Benefits versus risks from mammography: a critical reassessment. Cancer 1996;77:903–9.

16. Feig SA. Assessment of radiation risk from screening mammography. Cancer 1996;77:818–22.
17. Geise RA, Palchevsky A. Composition of mammographic phantom materials. Radiology 1996;198:347–50.

18. Young, KC, Wallis MG, Ramsdale ML. Manimographic film density and detection of small breast cancers. Clin Radiol 1994;49:461–5.

19. Land CE. Studies of cancer and radiation dose among atomic bomb survivors: the example of breast cancer. JAMA 1995;274:402–7.

20. Swift M, Morrell D, Massey RB, et al. Incidence of cancer in 161 families affected by ataxia-telangiectasia. N Eng J Med 1991;325:1831–6.

21. Boice JD Jr, Miller RW. Risk of breast cancer in ataxia-telangiectasia (letter to the editor). N Engl J Med 1991;326:1357–8.

22. Wagner LK. Risk of breast cancer in ataxia-telangiectasia (letter to the editor). N Engl J Med 1991;326:1358.

23. Hall EJ, Geard CR, Brenner DJ. Risk of breast cancer in ataxia-telangiectasia (letter to the editor). N Engl J Med 1991;326:1358–60.

24. Land CE. Risk of breast cancer in ataxia-telangiectasia (letter to the editor). N Eng J Med 1991;326:1359–60.

Mammography Versus Clinical Examination of the Breasts

Cornelia J. Baines, M.D., F.A.C.E.

In the context of contemporary North American society, the real issue is not "screening mammography versus clinical examination," it is "screening mammography with or without clinical examination." In the context of other societies where breast cancer is not a major priority and where funding for and expertise in screening mammography are scarce, clinical examination of the breasts as a single screening modality deserves consideration.

The North American fixation on technology as the solution for most problems, combined with the population's generally inflated view of the risks of getting breast cancer, of dying from breast cancer, and of benefiting from mammographic screening,[1,2] has resulted in clinical examination of the breast being given short shrift. Evaluation of its efficacy is hampered by the constrained opportunity for valid comparisons of combined screening with mammography and clinical breast examination (MA+CBE) versus CBE alone in the age group 40–49.

What can be examined are:

- Cancer detection rates achieved when CBE is the only screening modality.
- The Canadian National Breast Screening Study (NBSS) program sensitivities for CBE alone compared with MA+CBE.
- The mode of cancer detection (both MA and CBE positive, MA-only positive, and CBE-only positive) for screen-detected cancers in six screening studies using MA+CBE.
- Nodal status at time of cancer detection comparing two-modality with single-modality screening.
- Survival at 10 years post entry into a screening program by mode of detection.
- Deaths at 10 years post entry comparing preliminary NBSS data for MA+CBE and CBE in women ages 40–49.

The Health Insurance Plan (HIP) of Greater New York Study,[3] the Utrecht-based DOM Project,[4] the Edinburgh trial, [5] and the NBSS[6,7] all yield information about CBE. While CBE made a major contribution to breast cancer mortality reduction in the HIP, it is also true that contemporary mammography has improved so much that conclusions relevant in the 1970's are unlikely to be valid currently. Both the Utrecht and the Edinburgh study attributed little benefit to the CBE component of their programs. The CBE technique used in Utrecht (examination of four quadrants with cupped hand, personal communication) renders it very different from that employed in the NBSS.[8] Only in the Edinburgh trial at screening rounds 2, 4, and 6, and at the first screening round for women ages 40–49 and all screening rounds for women ages 50–59 assigned to CBE-only in the NBSS, does one have the opportunity to observe detection rates for single modality screening by CBE. They were highest in the NBSS.

NBSS program sensitivities for invasive cancers (using the detection method) reveal that highest and lowest sensitivities were attained in women ages 50–59—83 percent sensitivity for those

randomized to MA+CBE and 61 percent for those randomized to CBE. For women ages 40–49 randomized to single CBE, sensitivity was 67 percent compared with 77 percent for those getting annual MA+CBE.

In studies that have used two-modality screening,[4–7,9,10] the percentage of screen-detected cancers detected by CBE alone has ranged from 3 to 45 percent, whereas the range for all CBE detection (alone or combined with mammography) is 44–78 percent.

Examination of nodal status at first screen in the NBSS in women ages 40–49 reveals that there is a slight excess of both negative and positive nodal status in women assigned to MA+CBE. Thus far it is clear from available comparisons that two-modality screening detects more breast cancers than either modality alone, that yield from CBE alone varies among studies, and that MA-only screening detects more cancers and smaller cancers than CBE only. It also appears that the sensitivity of CBE is higher in younger than in older women. In contrast, the sensitivity of mammography is higher in older than in younger women. If two modalities are to be used, it may be most propitious to do so in younger women.

Analysis of survival status associated with mode of detection of screen-detected cancers is hampered by two factors. First, available Breast Cancer Detection Demonstration Project (BCDDP) data[10] for women ages 40–49 on entry into the study are limited to women diagnosed between ages 40 and 49, a method not used in other studies; the alternative was to use data for all screen-detected cancers in women ages 34–74. Secondly, the data for the HIP study[9] include all screened women ages 40–64. Survival improves for each category of detection in order of comparison, HIP to BCDDP to NBSS, even though the specificity of the latter's data to women ages 40–49 makes the comparison somewhat disadvantageous to the NBSS. However, 10-year survival data for screen cancers in HIP, BCDDP, and NBSS detected by MA alone are 77, 85, and 92 percent respectively; for CBE-only detection, survival is 59, 76, and 83 percent respectively, with all NBSS data being preliminary.

Finally, available (*and preliminary*) NBSS data on deaths within 10 years due to breast cancer detected at the first screen and in the first 12 months after the first screen, reveal 21 deaths in the MA+CBE arm and 19 in the CBE arm. The MA+CBE arm also includes one additional death and the CBE arm two additional deaths in patients with breast cancer who died of verified other causes. Of the 19 deaths due to breast cancer in the CBE arm, one occurred 10 years and 16 days, and the other 10 years and 3 months after date of entry, with no screen-1 or interval-1 deaths recorded thereafter. The only further death in the MA+CBE arm occurred in a screen-1 detection at 11 years 5 months post entry. To my knowledge, these are the only available data demonstrating mortality outcomes from two equivalent groups of women ages 40–49, the one receiving two-modality and the other single-modality screening. The slight excess in mortality in the screened group is no greater than has been shown in other studies.[11] Despite excellent 10-year survival after mammography-only detection, the NBSS does not provide compelling arguments to dismiss CBE.

It already has been observed that mammography misses at least 10 percent of breast cancers, that CBE independently detects breast cancer, that combined CBE and MA is superior to either modality alone, and that biopsy of a palpable dominant mass should not be deferred because the mammogram is normal.[12] It has already been demonstrated that medical school curricula could be revised to enhance CBE competence among medical students[13] and educational programs can effectively improve CBE competence among health professionals.[14] Women ages 40–49 would benefit from competent CBE if for no other reason than that the sensitivity of mammography declines with decreasing age.

References

1. Black WC, Nease Jr RF, Tosteson ANA. Perceptions of breast cancer risk and screening effectiveness in women younger than 50 years of age. J Natl Cancer Inst 1995;87:720–31.

2. Baines CJ. Women and breast cancer: Is it really possible for the public to be well informed? Can Med Assoc J 1992;146:2148–9.

3. Shapiro S, Venet W, Strax P, Venet L, Roeser R. Ten- to fourteen-year effect of screening on breast cancer mortality. J Natl Cancer Inst 1982;69:349–55.

4. DeWaard F, Collette HJA, Rombach JJ, Banders-van Halewijn EA, Honing C. The DOM project for the early detection of breast cancer, Utrecht, The Netherlands. J Chron Dis 1984:37;1–44.

5. Roberts MM, Alexander FE, Anderson TJ, Chetty U, Donnan PT, Forrest P, et al. Edinburgh trial of screening for breast cancer: Mortality at seven years. Lancet 1990;335:241–6.

6. Miller AB, Baines CJ, To T, Wall C. Canadian National Breast Screening Study: I. Breast cancer detection and death rates among women aged 40 to 49 years. Can Med Assoc J 1992;147:1459–76. (Published erratum appears in Can Med Assoc J 1993;148:718.)

7. Miller AB, Baines CJ, To T, Wall C. Canadian National Breast Screening Study: II. Breast cancer detection and death rates among women aged 50 to 59 years. Can Med Assoc J 1992;147:1477–88. (Published erratum appears in Can Med Assoc J 1993;148:718.)

8. Baines CJ, Miller AB, Bassett AA. Physical examination: Its role as a single screening modality in the Canadian National Breast Screening Study. Cancer 1989;63:1816–22.

9. Shapiro S, Venet W, Strax P, Venet L. Current results of the breast cancer screening randomized trial: The Health Insurance Plan (HIP) of Greater New York Study. In: Day NE, Miller AB, eds. Screening for Breast Cancer. Toronto: Hans Huber, 1988,(p.)3–15.

10. Seidman H, Gelb SK, Silverberg E, LaVerda N, Lubera JA. Survival experience in the Breast Cancer Detection Demonstration Project. CA 1987;37:258–90.

11. 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;30:187–209.

12. Winchester DP. Physical examination of the breast. Cancer 1992;69:1947–9.

13. Campbell HS, McBean M, Mandin H, Bryant H. Teaching medical students how to perform a clinical breast examination. Acad Med 1994;69:993–5.

14. Campbell HS, Pilgrin CA, Fletcher SW, Morgan TM, Lin S. Improving physicians' and nurses' clinical breast examination: a randomized controlled trial. Am J Prev Med 1991;7:1–8.

The Psychosocial Consequences of Mammography

Barbara K. Rimer, Dr.P.H.

Introduction

The consequences of mammography must be considered in an assessment of the risks and benefits of mammography. Studies have examined outcomes of the mammography experience such as anxiety, distress, depression, excessive fear of cancer, subsequent practice of BSE and adherence to recommended schedules for mammography or other followup.[1–3] Some attention has been paid to the psychological sequelae of abnormal mammograms. This is an important consideration, because some data suggest there is a higher rate of false positives in women aged 40–49.[4] Since from 6 to 20 percent of women who undergo mammography may require followup of an abnormality, an assessment of the consequences of this process is essential.[5] One concern is that the experience of an abnormal mammogram could act as a negative reinforcer that would deter women from subsequent mammograms. This presentation will consider the psychosocial consequences of both mammograms and abnormal mammograms.

The literature on the psychosocial consequences of abnormal exams is extremely limited. A search of cancer literature in MEDLINE and PsychInfo identified fewer than 30 discrete articles, some of which were tangential to the topic. For a field as large as the study of breast cancer screening, there has been surprisingly little study of the psychological consequences of mammography compared with the psychosocial barriers to mammography. Moreover, most of the studies have been conducted in Europe, and the generalizeability of results to the United States is not known.

Among the most rigorous of the reports were several in which questionnaires were completed among women who experienced false-positive mammograms. Lidbrink and colleagues found a significant short-term emotional reaction in 45 women who were recalled because of inconclusive findings on mammograms.[6] Bull and Campbell sent questionnaires to 750 women prior to breast cancer screening and then subsequently to women with normal findings and those who required other procedures.[7] There was no increase in general levels of depression or anxiety in any of the groups; there was a significant increase in the over-practice of breast self-examination (BSE) among women who required special assessments.

Lerman and colleagues evaluated women's psychological responses to abnormal mammograms and the effect on mammography adherence. [8,9] The sample of women from an IPA-model HMO included 1,221 women who had normal mammograms 3 months earlier, 119 women with low-suspicion mammograms, and 68 women with high-suspicion mammograms but not breast cancer. The authors assessed psychological responses and subsequent adherence to mammography. There was a relationship between the degree of mammogram suspicion and the strength of adverse outcome. Women with suspicious abnormal mammograms demonstrated significantly elevated levels of distress, and their mammography-related anxiety and breast cancer worries interfered with their moods and functioning. In the high-suspicion group, 47 percent of women had mammography-related anxiety, and 41 percent had worries about breast cancer; such worries affected the moods (26 percent) and daily functioning
(17 percent) of these women. Women with high and low levels of impairment were less likely to practice BSE than those with moderate impairment. Approximately 75 percent of women with

abnormal mammograms obtained their next mammograms on schedule.

These data suggest that even when the results of an abnormal mammogram are shown not to be cancer, some women experience negative sequelae. Nevertheless, this study was conducted among women ages 50–74, and it is not clear to what extent the results would be similar among women ages 40–49.

Psychosocial Consequences of Mammography

A small body of research examines the psychological consequences of mammography, primarily through questionnaires completed by patients.

Sutton and colleagues surveyed more than 2,000 breast cancer screening attenders and nonattenders aged 50–64 in England at three points in time.[10] Anxiety was highest at baseline, but the women were not overly anxious. Fine, Rimer and Watts[11] interviewed 250 women immediately after they had mammograms: 60 percent of women were anxious about having a mammogram, and 20 percent were extremely anxious. Some of this anxiety seemed to be due to a lack of information about what to expect. One study[12] with a small sample (53 women) indicated that women with a high familial risk of breast cancer had higher levels of distress and avoidant and intrusive thoughts after mammography than did normal risk women. The impact of family or other risk factors on response to mammography should be investigated further, because these women are likely to be advised to start mammography at a younger age.

Perceived Risks and Other Nonpsychological Sequelae Among Women in Their Forties

There is little research on women's perceptions of risks or negative sequelae related to mammography. In a survey conducted among 300 women who were patients at two North Carolina clinics, only 35 percent mentioned any negative consequences of mammography, and the most frequently mentioned were pain and radiation.[13] These also have been mentioned as barriers to mammography in other studies.[14] When false positives and negatives were defined, about 72 percent of women said they would be somewhat or very concerned about a false positive, and 77 percent would be somewhat/very concerned about a false-negative result.[13] More research is needed to learn how best to help women weigh the risks and benefits so as to make informed decisions about mammography.

Interventions To Reduce Negative Psychosocial Consequences and/or Improve Coping

There has been little research to test the impact of interventions to reduce anxiety and distress and to improve coping after an abnormal result. In one of the few studies in this area, Lerman and colleagues sent women in the experimental condition a booklet designed to improve adherence to the subsequent mammogram following an abnormal test.[15] The brief psychoeducational booklet resulted in a 13-percent increase in adherence.

Summary

The research base on the psychosocial consequences of mammography, in general, and abnormal mammograms, in particular, is extremely limited. The investigations have used different age groups, timeframes, measurements, and outcomes. Moreover, different levels of support have been provided to help women cope with the abnormal experience, thus serving as a potential confounder. Women's

reactions also may be affected by the manner in which the results are communicated. Thus, it is difficult to reach clear conclusions about the impact of mammography or abnormal mammograms on such outcomes as anxiety, distress, or adherence to recommended breast screening. Among some women, there does seem to be short-term distress, and at least one study shows that level of distress is related to the index of mammogram suspicion. There is a need for more research that is rigorous and includes sufficient numbers of women ages 40–49. It would be useful to determine how long the negative effects persist and also whether the impact is exacerbated in women with a previous abnormal exam. Moreover, the impact of brief psychoeducational interventions designed to help women cope with the abnormal experience should be investigated. It would be helpful to determine whether women at high risk can be identified and provided with intervention in a proactive manner. Finally, it is not known to what extent negative psychosocial sequelae of mammography might affect followup with recommendations for additional tests or delay in seeking care for potential symptoms.[2] Noncompliance with followup recommendations continues to be a problem,[5] and it would be critical to assess to what extent noncompliance is affected by psychosocial factors related to the experience of mammography.

At present, there are more questions than answers. One of the more intriguing questions is why there has been so little inquiry in an area that is of such vital concern.

References

1. Ellman R, Angeli N, Christians A, Moss S, Chamberlain J, Maguire P. Psychiatric morbidity associated with screening for breast cancer. Br J Cancer 1989;60:781–.

2. Lerman C, Rimer BK, Engstrom PR. Cancer risk notification: psychosocial and ethical implications. J Clin Oncol 1991;9:1275–82.

3. Paskett E, Rimer BK. Psychosocial effects of abnormal pap tests and mammograms: a review. J Women's Health 1995;4(1):73–82.

4. Fletcher SW, Black W, Harris R, Rimer BK, Shapiro S. Report of the International Workshop on Screening for Breast Cancer. J Natl Cancer Inst 1993;85(20):1644–56.

5. McCarthy BD, Yood MU, Boohaker EA, Ward RE, Rebner M, Johnson CC. Inadequate follow-up of abnormal mammograms. Am J Prev Med 1996;12: 282–8.
6. Lidbrink E, Levi L, Pettersson I, Rosendahl I, Rutqvist LE, de la Torre B, Wasserman J, Weige M. Single-view screening mammography: psychological, endocrine and immunological effects of recalling for a complete three-view examination. Eur J of Cancer 1995;31A:932–3.

7. Bull AR, Campbell MJ. Assessment of the psychological impact of a breast screening programme. Br J Radiol 1991;64:510–5.

8. Lerman C, Trock B, Rimer BK, Jepson C, Brody D, Boyce A. Psychological side effects of breast cancer screening. Health Psychol 1991;10:259–67.

9. Lerman C, Trock B, Rimer BK, Boyce A, Jepson C, Engstrom PF. Psychological and behavioral implications of abnormal mammograms. Ann Intern Med 1991;114:657–61.

10. Sutton S, Saidi G, Bickler G, Hunter J. Does routine screening for breast cancer raise anxiety? Results from a three wave prospective study in England. J Epidemiol Commun Health 1995;49: 413–8.

11. Fine MK, Rimer BK, Watts P. Women's responses to the mammography experience. J Am Board Fam Pract 1993;6:546–55.

12. Valdimarsdottir HB, Bovbjerg DH, Kash KM, Holland JC, Osborne MP, Miller DG. Psychological distress in women with a familial risk of breast cancer. Psycho-Oncol 1995;4:133–41.

13. Morton E, Tambor E, Rimer BK, Tessaro I, Farrell D, Siegler IC. The impact of the National Cancer Institute's revised mammography screening guidelines on women age 40–49. Women's Health Issues 1996;6(5):246–54.

14. Rimer et al. 1989.

15. Lerman C, Ross E, Boyce A, Gorchov P, McLaughlin R, Rimer BK, Engstrom P. The impact of mailed psycho educational materials to women with abnormal mammograms. Am J Public Health 1991;82(5):1–2.

16. Gram IT, Glenker SE. Cancer anxiety and attitudes toward mammography among screening attenders, nonattenders, and women never invited. Am J Public Health 1992;82:249–51.

Variation of Benefits and Harms of Breast Cancer Screening With Age

Russell P. Harris, M.D., M.P.H.

Getting the Question Right

The issue at this conference is what level of recommendation to make to women of different ages about breast cancer screening. I want to emphasize the phrase "what level of recommendation." Some may think the answer is a simple "yes" or "no"—either we recommend or we don_t. I will argue, along with groups like the U.S. Preventive Services Task Force, that the strength of the recommendation should depend on the evidence and that the evidence should be about two issues: the *benefits* of screening and the *harms* of screening, that is, the answers to conference questions 1–3. The issue, then, is not whether there is some small benefit demonstrated for screening for women in their forties. The issue is larger than a *p* value. The issue is where the balance lies between benefits and harms.

I will further argue that, in cases where the balance between benefits and harms is not clear, as I believe is the case with breast cancer screening for women in their forties, we should not recommend screening, but rather *discussion of screening,* so that individual women may make the informed decision that is best for them.

Question 1: Mortality Benefits of Screening

Screening seeks to decrease the risk of dying of breast cancer, not the risk of getting it. The specific risk a woman is trying to reduce by being screened for the next 10 years is the risk of eventually dying of cancer diagnosed in those next 10 years. These risks for women of different ages, calculated from NCI Surveillance, Epidemiology, and End Results (SEER) program data before widespread screening, are given in the second column of Table 1. Not surprisingly, the risk increases with age.

TABLE 1. Benefits of Screening

Age	Risk per 1,000 Women[a]	Relative Risk Reduction	Absolute Risk Reduction[b]
40	7.8	13[c]	1.0
		23[d]	1.8
50	12.9	15	1.9
		30	3.9
60	19.5	30	5.9
Æ70	25.3	30[e]	7.6

[a] Risk of dying in next 15–20 years of breast cancer diagnosed in next 10 years, from SEER data, 1973–80 and 1989–91.

[b] Number of lives ultimately extended per 1,000 by screening over the next 10 years.

[c] From Swedish meta-analysis.

[d] From Edinburgh trial, beginning with age 45 years.

[e] Extrapolated from 60–69 age group.

When the relative risk reduction from the randomized trials (Table 1, column 3) is factored in, we can calculate the absolute risk reduction (Table 1, column 4)—the number of women per 1,000 whose lives would ultimately be extended by screening over the 10 years. Note we are assuming here that the relative risk reduction from the randomized trials of women in their forties is real. Again, not surprisingly, the benefit (as indicated by the absolute reduction in risk) increases with age.

Questions 2 and 3: Nonmortality Effects of Screening

Questions 2 and 3 deal with nonmortality effects of screening, whether positive or negative. Like Professor Rimer, I wish we had more data on this issue. Some things, however, are apparent from simply looking at the screening "cascade," that is, the expected sequence of events following screening. I have shown this cascade for a single screening of a hypothetical population of 10,000 women of two different age groups in Figures 1 and 2. The numbers in these diagrams were taken from reviews of mammogram performance and from a recent national survey of mammography facilities. I have conservatively used a mammogram "positivity" rate of 5 percent—less than the 10–11 percent figure found in the national survey. (Using the higher rate would double the number of false-positives.) Radiologists I have spoken to have not challenged these numbers.

But what does the diagram tell us about nonmortality benefits or harms? Let_s start with the women who have a negative mammogram. Most of these women are truly negative—they do not have breast cancer. A few, however, truly have cancer but are screen-negative—falsely negative. A research priority is to find out whether some of these women have been injured by false reassurance. It seems possible that some will ignore early symptoms of breast cancer because they have been reassured by the negative mammogram. We don_t know.

The true negatives would seem to be in a position to benefit—I call this the "peace of mind" group. But if you look carefully at the probability of having cancer before screening versus after being screen- negative, the difference (for women in their forties: from .0016 before screening to .0004 after a negative screen) doesn_t seem large enough to make a truly objective woman change from worrying to relaxing. The woman was at low risk before screening and is still at low (but not zero) risk after being screen-negative. Not much real benefit there.

How about those women who were screen-positive? The great majority are falsely positive. And the cascade shown (Figures 1 and 2) is for a single screen. The cumulative probability of having at least one false-positive over 10 years of screening is unknown and should be a research priority. This probability could easily be as high as 30 percent (or more) of women.

The false-positive women then undergo a "workup," which may be fine needle aspirate, ultrasound, or magnification views. Some will come to biopsy. These are the women that Professor Rimer discussed so well—some burdened with anxiety not only before biopsy but even months after being told there is no cancer. There are certainly potential harms there.

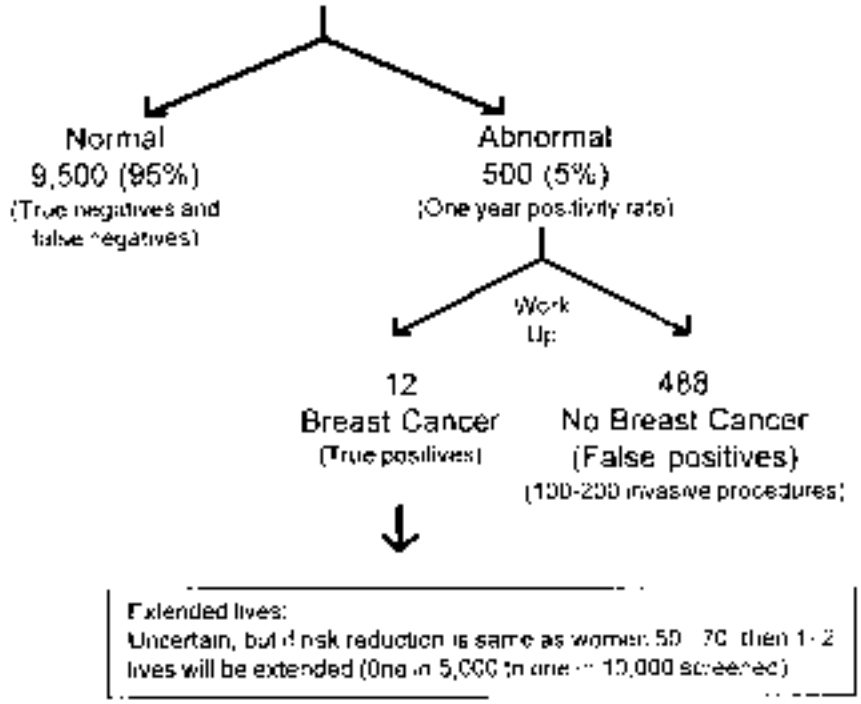

FIGURE 1. 10,000 Women Ages 40–49 Screened Annually

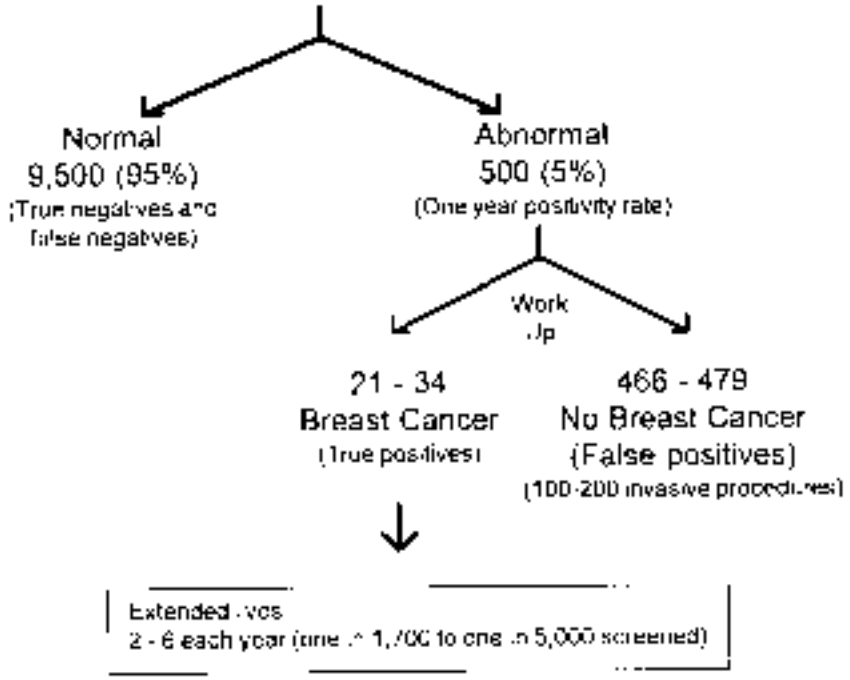

FIGURE 2. 10,000 Women Ages 50–75 Screened Annually

The next group to consider is the "true-positives"—those women who screened positive and were found to have breast cancer. These are the very women we usually think have been helped by screening. Yet the facts are that not all women whose cancer was found by screening have benefited from that finding. About 50 percent of these women would not have died from breast cancer, even if they had waited until an obvious lump appeared to detect it. These cancers are clearly slower growing and relatively more treatable. Screening has not altered the natural history, because the natural history is excellent. The perception of many of these women, quite understandably, is that their lives were "saved" by screening.

Another group of "true-positive" women, unfortunately, will die of breast cancer regardless of when it is found. They have cancers that are highly malignant, metastasizing from an early, undetectable stage. Again, the natural history of the disease has not been altered by screening.

It is difficult to find benefit in situations in which a woman has been diagnosed with breast cancer before she would otherwise have been diagnosed, and yet the natural history of the disease has not been changed. The woman has simply been made to live longer with knowledge of the diagnosis. One could argue that these women have been harmed, not helped, by screening.

Finally, there are some women in this "true-positive" group who have been harmed, because they have been given a diagnosis that would never have become clinically apparent—so-called "pseudodisease." Some women with ductal carcinoma *in situ* (DCIS) are probably in this category. The natural history of DCIS is unknown. Some, but likely not all, of these lesions will progress to invasive carcinoma. And when progression occurs, it may take many years (thus allowing opportunities for detection at a later age). Understanding the natural history of DCIS and determining the characteristics of those lesions that will become clinically important as opposed to those that are actually "pseudodisease" should be a research priority. If, as we suspect, as many as 50 percent of these lesions are clinically unimportant, then the potential for harming women by unnecessary treatment could be an important factor for women to consider in deciding about screening.

Variation of Harms by Age

But thus far we have discussed harms that apply both to women in their forties and to older women. How do these effects vary by age? Our best guess is shown in Table 2. The group that I want to emphasize, however, is the false-positives, by far the largest number of women who could potentially be harmed by screening. As noted above, cumulatively over 10 years this group may include as many as 30 percent of all women. Any harm to this large number of women would need to be considered carefully in making screening recommendations.

A critical question, then, is whether the probability of a woman becoming a false-positive varies by age. An important determinant of the probability of having a false-positive is the initial "positivity rate" of screening—that is, the percentage of women screened who required some further workup. There is conflicting evidence about whether this percentage varies with age. In some studies, especially those of academic practices, the positivity rate appears fairly constant. In studies of community practices, younger women have higher positivity rates (and thus more false-positives) than older women. The issue is important and should be a research priority. Even with the same positivity rate, however, the fact that the incidence of breast cancer is higher in older women means that more of the positives in younger women will be falsely positive.

TABLE 2. Harms of Screening		
Type of Finding	Harm	Relationship with Age
False-Negative	False Reassurance	40>50/60[a]
False-Positive	Psychological trauma	40>50/60
True Positive	Living longer with knowledge of disease	40u50/60 (?)
No change in Natural History		
Pseudodisease	Labeling-psychological effects	40u50/60 (?)
	Unnecessary treatment	

[a] 40 = women in their forties; 50/60 = women in their fifties and sixties

But another factor makes it very likely that women in their forties would have a larger—even a much larger—probability of a false-positive than older women. This other factor is the frequency of screening. From the trials of women over 50, it appears that a large percentage of the benefit of annual screening can be obtained by screening biennially. For women in their forties, however, it is clear that if screening works at all, it must be done annually. Although there is need for research in this area, it seems likely that screening twice as frequently would produce a higher cumulative rate of false-positive findings than screening annually.

The bottom line is that breast cancer screening is not the final answer to the problem of breast cancer in any age group. It certainly has benefits, however, among women ages 50–70, and perhaps some benefits among women ages 40–49. I suspect, like others, that these benefits begin at a low level in the early forties, then gradually increase with age (Table 1). Harms, on the other hand, flow primarily from the large number of false-positives, which are probably higher among women in their forties than in older women (Table 2).

Restating the Problem–and One Last Harm

The problem, then, can be immediately appreciated. As they grow older, even well-informed women will naturally differ in their perceptions of the age at which the increasing probability of benefit outweighs the decreasing probability of harm. And policymakers will naturally differ in their evaluation of the age, on the population level, at which the increasing benefits of screening begin to outweigh the decreasing harms. Perhaps the disagreement should tell us something. We differ not because we disagree about what the evidence is, but rather because our values differ. There is no consensus about screening for women in their forties, nor should there be. This is a “close call.” In such situations, women should be helped to make their own decisions.

One further population-level harm of aggressive screening in borderline situations should be mentioned. Several research studies, including ones that we conducted, have shown that women of all ages, but especially women in their forties, far overestimate both their risk of breast cancer and the potential benefits of screening. But should we be surprised about these erroneous perceptions in a climate of strong recommendations in close-call situations? Strong recommendations send the message that we have an extremely powerful tool and that using it is a totally positive experience. Perhaps in borderline situations we should give less advice—and more information.

A more measured approach is needed, particularly for women ages 40–49. The recommendation for these women should be that they be informed that there are pros and cons to being screened, and

that reasonable women will disagree. They should be encouraged to discuss screening with their physicians, to clarify their own values, and to make an informed choice for themselves. Then we should get to work on the real issue: how to efficiently and effectively reach women with this discussion.

Bibliography

1. Harris R, Leininger L. Clinical strategies for breast cancer screening: weighing and using the evidence. Ann Intern Med 1995;122:539–47.

2. Black WC, Welch HG. Advances in diagnostic imaging and overestimations of disease prevalence and the benefits of therapy. N Engl J Med 1993;328:1237–43.

3. Redelmeier DA, Rozin P, Kahneman D. Understanding patients_ decisions: cognitive and emotional perspectives. JAMA 1993;270:72–6.

4. Black WC, Nease RF, Tosteson ANA. Perceptions of breast cancer risk and screening effectiveness in women younger than 50 years of age. J Natl Cancer Inst 1995;87:720–31.

5. Brown ML, Houn F, Sickles EA, Kessler LG. Screening mammography in community practice: positive predictive value of abnormal findings and yield of follow-up diagnostic procedures. Am J Roentgenol 1995;165:1373–7.

6. Kerlikowske K, Grady D, Barclay J, Sickles EA, Eaton A, Ernster V. Positive predictive value of screening mammography by age and family history of breast cancer. JAMA 1993;270:2444–50.

7. Ciatto S, Cecchini S, Del Turco MR, Grazzini G, Iossa A, Bartoli D. Referral policy and positive predictive value of call for surgical biopsy in the Florence Breast Cancer Screening Program. J Clin Epidemiol 1990;43:419–23.

8. Baker LH. Breast Cancer Detection Demonstration Project: five-year report. Cancer 1982;32:194–225.

Screening for Breast Cancer in Younger Women Ages 40–49

Helena R. Chang, M.D., Ph.D., and Kirby I. Bland, M.D.

While early detection of breast cancer through screening among women age 50 years or older is considered to be the gold standard in the nation, screening in younger women has been repeatedly scrutinized. The main issues are (1) the uncertainty of the effectiveness of screening mammography in younger women ages 40–49 and (2) a potential risk of radiation-induced breast cancer among women starting mammography screening at age 40.

A large series reported by the Breast Cancer Detection Demonstration Project (BCDDP) showed that the breast cancer detection rates by screening mammography were similar between women ages 40–49 years and older women 90 percent versus 92 percent. The distribution of tumor sizes was essentially identical between the two groups as well. Finally, the difference between case fatality rates at 14 years of followup was not statistically significant between the younger and older women. Since BCDDP is not a randomized study, the reduction of breast cancer mortality by screening mammography was calculated based on the comparison with the expected mortality obtained from the National Cancer Institute Surveillance, Epidemiology, and End Results (SEER) experience. An estimated lower breast cancer mortality rate of 11 percent was reported in younger women as a result of the screening.

Several randomized studies have examined the survival benefit derived from breast cancer screening among young women. Based on a meta-analysis of six trials conducted prior to the Canadian National Breast Screening Study (NBSS), a reduction of 15 percent was seen in breast cancer mortality in women ages 40–49 years at entry to the screening. With the inclusion of the Canadian results, the reduction of breast cancer mortality rate by screening drops to 7 percent.

The Health Insurance Plan (HIP) of Greater New York was the only randomized study conducted in the United States. After 18 years of followup, a 25-percent reduction in breast cancer mortality was observed in the screened young women. Similarly, a 13-percent reduction in the breast cancer mortality rate was observed in women ages 40–49 years, after a followup of 8 years in the Swedish randomized trials. No evidence of any detrimental effect of screening was found in any age group.

The Canadian NBSS, reported in 1992, was aimed to study the detection and death rates by breast cancer screening among women ages 40–49 years. This study claimed an increased risk of breast cancer death associated with breast cancer screening. The NBSS study suffered from several major flaws:
(1) inclusion of a greater number of individuals with positive physical examinations in the screening group, (2) insufficient statistic power as well as length of followup, and (3) questionable quality control of the screening mammography employed in the study. Even in the presence of these unfavorable factors, the NBSS study indicated that the highest survival rate was observed in young women with breast cancer detected only by mammography.

In conclusion, young women ages 40–49 years benefit from periodic screening mammographies, and there is little risk of breast cancer associated with screening mammographies of this age group. It should be emphasized that young women account for one-third of all breast cancer victims in the United States. It is our responsibility to establish guidelines for screening among these women. In the future, research should address areas other than survival, such as (1) Is there a shift of early cancer detection by screening in young women ages 40–49 years? (2) Is there an increasing utilization of breast conservation surgery in young women because of the smaller tumors detected at an asymptomatic stage? (3) How can we improve the sensitivity and specificity of the screening tools? and (4) Can we develop a customized screening program based on the improved knowledge of risk assessment for breast cancer?

References

1. Feig SA. Assessment of radiation risk from screening mammography. Cancer 1996;77:818–22.

2. Mettler FA. Upton AC, Kelsey CA, Ashby RN, Rosenberg RD, Linver MN: Benefits versus risks from mammography. A critical reassessment. Cancer 1996;77:903–9.

3. Lein BC, Alex WR, Zebly M, Pezzi CM. Results of needle localized breast biopsy in women under age 50. Am J Surg 1996;171:356–9.

4. Smart CR, Hendrick RE, Rutledge JH III, Smith RA. Benefit of mammography screening in women aged 40–49: Current evidence from randomized controlled trials. Cancer 1995; 75:1619–26.

5. Mettlin C, Smart CR. Breast cancer detection guidelines for women aged 40 to 49 years: rationale for the American Cancer Society reaffirmation of recommendations. CA Cancer J Clin 1994;44:248–55.

6. Harris R. Breast cancer among women in their forties toward a reasonable research agenda. J Natl Cancer Inst 1994; 86: 410–12.

7. Smart CR, Hartmann WH, Beahs OH, Garfinkel L. Insights into breast cancer screening of younger women. Cancer Suppl 1993;72:1449–56.

8. Woolf SH. United States Prevention Services Task Force recommendations on breast cancer screening. Cancer Suppl 1992;69:1913–18.

9. Kessler LG. The relationship between age and incidence of breast cancer. Cancer Suppl 1992;69:1896–903.

10. Miller AB, Baines CJ, To T, Wall C. Canadian National Breast Screening Study: Breast cancer detection and death rates among women aged 40 to 49 years. Can Med Assoc J 1992;147:1459–76.

Detection and Treatment Trends: A Clinical Experience

Blake Cady, M.D.

The size of all primary invasive breast cancer has been tracked at the Deaconess Hospital in Boston, Massachusetts, since the founding of the pathology department in 1929.[1] These data offer an intriguing historical perspective on what is happening with invasive breast cancer in the United States today. Between 1929 and 1948 there was no change whatsoever in the mean maximum diameter of all invasive breast cancer. From 1949 through 1968 there was a small but statistically significant decrease in the mean maximum diameter of all invasive breast cancers, which averaged about 3 percent every 5 years, probably because of public education efforts at early detection of cancer. Beginning in 1969—my definition of the beginning of the mammographic era—the mean maximum diameter of all invasive breast cancer declined by 10 percent every 5 years, and between 1989 and 1993 was only 2.1 cm.[1] The median maximum diameter declined similarly during that period, and during the years 1989–1993 was only 1.5 cm. Because of biases that might be present in clinical presentation at our tertiary care hospital, similar analysis was conducted at a community hospital in Cambridge, Massachusetts (Mt. Auburn Hospital), for patients examined since their Tumor Registry began 15 years ago. In the years 1989 to 1993, those community hospital data were essentially identical (mean maximum diameter 2.0 cm; median maximum diameter 1.7 cm).[1] The proportion of all invasive breast cancer that were T1a and T1b (1 cm or smaller in diameter) was now 29 percent and 28 percent in the two hospitals, respectively. Of most interest is the fact that the sharp downward slope of decreasing mean and median diameter of all invasive breast cancer shows no signs of abating at the present time. This indicates that within a decade the median maximum diameter of all invasive breast cancer in the United States may be only 1 cm and therefore that 50 percent of all invasive breast cancers will be of T1a and T1b size.

In addition, the data during the 25 years ending in 1993 demonstrate an increasing proportion of patients presenting with small duct carcinoma *in situ* (DCIS). Whereas DCIS patients made up only 3 percent of cases in the early 1970's, they have made up 14 percent of all patients in our hospital in the past 5 years, and 26 percent of all cases 1 cm or smaller in diameter.[1] That increasing incidence of DCIS will undoubtedly continue, so that it can be projected that between 25 and 33 percent of all breast cancer within the next decade will consist of DCIS, the proportion frequently described in mammographic screening data.

In the years 1989–1993, 44 percent of all breast cancer appearing at the Deaconess Hospital consisted of DCIS or T1a and T1b invasive breast cancers. Again, the sharply rising slope of the incidence of small breast cancers, both noninvasive and invasive, shows no current sign of leveling off. Projecting these curves of size and invasiveness would predict that, within a decade, between two-thirds and three-quarters of all breast cancers will be either DCIS or T1a and T1b invasive cancers. This decrease in size and increased DCIS in both a tertiary care and a community hospital in a sophisticated medical community with a relatively high socioeconomic base demonstrate the impact of mammographic screening on a proportion of our current patient population, and on the entire population if universal screening were achieved.

Such striking changes in the presentation of breast cancer truly predict a new era in breast cancer in this country due to mammographic screening.[1] All of this has occurred at a time when women's health surveys indicate that fewer than one-half of appropriately aged woman have routine yearly mammographic screening and one-third of women still have not had a mammogram. In that

women's health survey, women with the highest proportion of routine mammographic screening were those between the ages of 40 and 49. This undoubtedly reflects anxiety about breast cancer in the population at large but also the fact that women in their forties with children or an expanded idea of life duration are most interested in protecting their families and their health.

Accompanying the decline in mean and median diameter invasive breast cancer is a continuing decrease in the proportion of invasive breast cancer patients who have axillary lymph node metastases. In the 5 years 1989–1993, only 31 percent of all of our invasive breast cancer patients who had axillary dissection had axillary lymph node metastases, and only 10 percent of patients with positive nodes had more than three nodes positive.[1] Furthermore, of all patients with T1a and T1b invasive breast cancers who had an axillary dissection, only 10 percent had positive nodes. Of patients with T1a and T1b breast cancers with positive nodes, 70 percent had only one or two positive nodes, and of those, nearly 50 percent were only micrometastases. Illustrating the conundrum about the present status of lymph node dissection and lymph node analysis was the fact that during the 5 years 1989–1993, while Deaconess Hospital data showed a decline in T1a and T1b positive node rates to 10 percent, Mt. Auburn Hospital data revealed a doubling of the positive node rate in these small cancers to over 20 percent.[1] This was subsequently discovered to be the result of a new definition of axillary lymph node examination by pathologists at Mt. Auburn Hospital who had a special academic interest,[2] in contrast to the adherence to the traditional single histologic section of lymph nodes practiced at Deaconess Hospital through the entire 65 years of the pathology department up to 1993. By the use of multiple sectioning of the lymph nodes (up to 15 sections per node) with haematoxylon and eosin staining, and the use of histochemical markers, the incidence of "positive" lymph nodes at Mt. Auburn Hospital doubled, but these newly discovered additional node metastases consisted almost entirely of micrometastases.[2] At Mt. Auburn Hospital during those years, every patient who had a positive node, whether a single micrometastasis or several macrometastases, was reported as having a "positive" axilla. The clinical report for the patient record did not separately define micrometastatic from macrometastatic disease. Thus, many patients with small cancers received adjuvant chemotherapy or Tamoxifen on the basis of a single or double micrometastasis discovered by extensive (nonroutine) examination of axillary lymph nodes.[3] It can be seen by this description that the TNM system of nodal staging of the current American Joint Committee on Cancer (AJCC) staging system for breast cancer may be completely altered by a change in the standards of lymph node examination. This alteration is a clear-cut example of the "Will Rogers" effect of stage shifting as a result of more extensive diagnostic maneuvers and detection of previously obscure components of disease.[4]

Primary tumor features have been found to correlate with the risk of lymph node metastases and thus enable selection of many patients who do not need axillary dissection for prognostic purposes.[5–12]

Because of the many small (T1a and T1b) invasive breast cancers and a high proportion of well-differentiated or moderately differentiated lesions,[13] the risk of in-breast local recurrences after excision and radiotherapy may be very low, and the increased risk of local recurrence without adjuvant radiotherapy may still be so low as to not justify the high cost of radiotherapy ($15,000–$20,000 for 6000 cGy) because no difference in survival occurs.[14,15] Recent development of the Van Nuys Prognostic Index for DCIS also indicates that a large proportion of patients with mammographically discovered noninvasive cancer do not need radiotherapy.[16]

Thus, under the impact of mammographic screening, a large proportion of invasive breast cancers in the near future will not require axillary dissection,[17] and a large proportion of both DCIS and T1a

and T1b invasive cancers may not require adjuvant radiation therapy[18] with the result of marked simplification of therapy, reduction in morbidity, and reduction in cost with no compromise in the excellent prognosis of mammographically discovered breast cancer.[19,20]

References

1. Cady B, Stone MD, Schuler JG, Thakur R, Wanner MA, Lavin PT. The new era in breast cancer invasion, size, and nodal involvement dramatically decreasing as a result of mammographic screening. Arch Surg 1996;131:301–8.

2. Lester SC, Rushing L. The effect of standard methods of processing lymph nodes on the number and size of metastases detected in women with breast cancer. Unpublished manuscript.

3. Henderson IC. Adjuvant systemic therapy for early breast cancer. Cancer 1994;74(Suppl):401–9.

4. Giuliano AE, Dale PS, Turner RR, Morton DL, Evans SW, Krasne DL. Improved axillary staging of breast cancer with sentinel lymphadenectomy. Ann Surg 1995;222:394–401.

5. Ahlgren J, Stahl O, Westman G, Arnesson LG. Prediction of axillary lymph node metastases in a screened breast cancer population. Acta Oncol 1994;33:603–8.

6. Barth A, Craig PH, Silverstein MJ. Predictors of axillary lymph node metastases in patients with T1 breast carcinoma. Cancer (in press).

7. Cady B. The need to reexamine axillary lymph node dissection in invasive breast cancer. Editorial. Cancer 1994;73(3):505–8.

8. Chadha M, Chabon AB, Friedmann P, Vikram B. Predictors of axillary lymph node metastases in patients with T1 breast cancer. Cancer 1994;73:350–3.

9. Lin PP, Allison DC, Wainstock J, Miller KD, Dooley WC, Friedman N, Baker RR. Impact of axillary lymph node dissection on the therapy of breast cancer patients. J Clin Oncol 1993;11:1536–44.

10. Mansour EG, Ravdin PM, Dressler L. Prognostic factors in early breast carcinoma. Cancer 1994;74:381–00.
11. Seidman DJ, Schnaper LA, Aisner SC. Relationship of the size of the invasive component of the primary breast carcinoma to axillary lymph node metastasis. Cancer 1995;75:65–71.

12. Silverstein MJ, Gierson ED, Waisman JR, Senofsky GM, Colburn WJ, Gamagami P. Axillary lymph node dissection for T1a breast carcinoma. Cancer 1994;73:664–7.

13. Arnesson LG, Smeds S, Fagerberg G. Recurrence-free survival in patients with small breast cancers. An analysis of cancers 10mm or less detected clinically and by screening. Eur J Surg 1994;160:271–6.

14. Leitner SP, Swern AS, Weinberger D, Duncan LJ, Hutter RVP. Predictors of recurrence for patients with small (One centimeter or less) localized breast cancer (T1a,b N0 M0). Cancer 1995;76(11):2266–74.

15. Schnitt SJ, Hayman J, Gelman R, Eberlein TJ, Love SM, Mayzel K, Osteen RT, Nixon AJ, Pierce S, Connolly JL, Cohen P, Schneider L, Silver B, Recht A, Harris JR. A prospective study of conservative surgery alone in the treatment of selected patients with stage 1 breast cancer. Cancer 1996;77(6):1094–100.

16. Silverstein MJ, Lagios MD, Craig PH, Waisman JR, Lewinsky BS, Colburn WJ, Poller DN. A prognostic index for ductal carcinoma *in situ* of the breast. Cancer 1996;77:2267–74.

17. Cady B. Is axillary node dissection necessary in routine management of breast cancer? No. In: DeVita VT, Hellman S, Rosenberg SA, eds. Important advances in oncology. Philadelphia: Lippincott-Raven Publishers, 1996. p. 251–65.

18. Cady B, Stone MD, Wayne J. New therapeutic possibilities in primary invasive breast cancer. Ann Surg 1993;218(3):338–49.

19. Hellman S. Karnofsky memorial lecture. Natural history of small breast cancers. J Clin Oncol 1994;12(10):2229–34.

20. Tabar L, Fagerberg G, Chen HH, Phil M, Duffy SW, Smart CR, Gad A, Smith RA. Efficacy of breast cancer screening by age. New results from the Swedish Two-County Trial. Cancer 1995;75:2507–17.

Increases in Ductal Carcinoma *In Situ* in Relation to Mammography: A Dilemma

Virginia L. Ernster, Ph.D.

The widespread adoption of screening mammography has led to a dramatic increase in detected cases of ductal carcinoma *in situ* (DCIS) of the breast. DCIS is usually referred to as "preinvasive" or "noninvasive" cancer because it is confined to the milk ducts of the breast and has not spread to the surrounding breast tissue. Although DCIS lesions are visible on mammograms in asymptomatic women, they are rarely clinically palpable; before the advent of mammography they were often only detected incidental to a biopsy for benign breast disease. In 1993, there were an estimated 23,275 newly diagnosed cases of DCIS in the United States (compared with 4,901 in 1983), of which 4,676 were in women ages 40–49 (compared with 742 in 1983).[1]

The recent "epidemic" of reported cases of DCIS presents women and their physicians with a dilemma. Probably only a fraction of DCIS cases will actually go on to invasive breast cancer and become clinically important. However, since current medical knowledge does not permit us to identify which women with DCIS will progress to invasive breast cancer and which will not, at present almost all women with a DCIS diagnosis are treated surgically. The hope is that by detecting malignant changes as early as possible, we are saving lives. The concern is that we may be detecting changes that for many women would never become life-threatening or even clinically apparent and that, in the process, we are overtreating women. This situation is similar to the current dilemma posed by PSA screening for prostate cancer; while debate continues as to whether that test reduces risk of prostate cancer death, it is known that PSA screening picks up many occult cancers that are clinically unimportant but for which thousands of men have had their prostates removed, resulting in impotence and incontinence for many individuals.

Other presentations address the fact that most abnormal mammography results are *false* positives. Here we are discussing the problem of *true-but-possibly-clinically-insignificant* positives, with particular reference to DCIS. For every individual woman contemplating screening, the willingness to risk a false positive or a positive result that may be clinically insignificant will differ, and it is therefore important that women know the probabilities of such outcomes in order to make their own informed decisions.

The DCIS Epidemic and its Relation to Changes in Mammography Prevalence

Based on data from NCI's Surveillance, Epidemiology, and End Results (SEER) program, age-adjusted incidence rates for DCIS in the United States increased 549 percent over the 21-year period 1973–1993, with most of the increase occurring during the 1980s. (For comparison, the increase in incidence rates for invasive breast cancer over the period 1973–1993 was 31.9 percent.) During 1983–1993, increases in DCIS incidence rates in the United States were dramatic for women in the age group 40 and older but much more modest for women younger than 40, who are much less likely than older women to undergo screening mammography (see figure below). For all age groups combined, DCIS accounted for 2.8 percent of newly diagnosed breast cancers in the United States in 1973, 3.8 percent in 1983, and 12.5 percent in 1993. Among women ages 40–49 years, DCIS accounted for 3.7 percent of all breast cancers in 1973, 4.2 percent in 1983, and 14.7 percent in 1993.[1] Mammography screening programs, which focus on asymptomatic women, typically report much higher proportions of DCIS among all breast cancers detected than is observed in data from

tumor registries, which include symptomatic women as well. For example, of breast cancers detected among women ages 40 years and older presenting to the UCSF Mobile Mammography Screening Program during 1985–1996 who had no report of a palpable mass, 27.6 percent were DCIS. Among breast cancer cases detected among women ages 40–49 years in that program, 40.9 percent were DCIS.

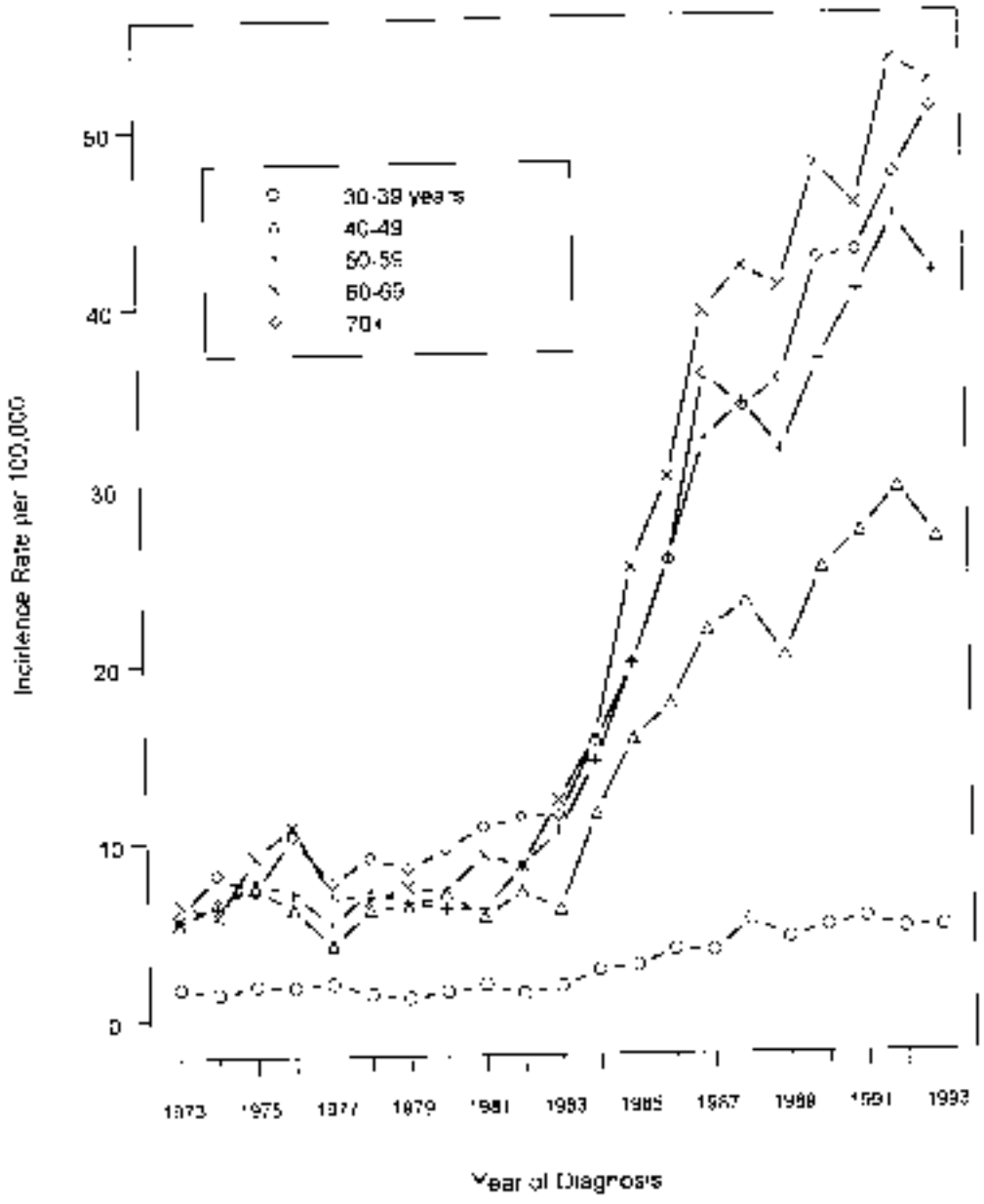

FIGURE 1. Trends in DCIS Incidence Rates by Age, 1973–1993

There were 134 mammography machines in the United States in 1982 and an estimated 10,000 by 1990.[3] Meanwhile, mammography prevalence increased markedly; the proportion of U.S. women reporting recent mammography doubled between 1987 and 1992.[2] Most but not all of the increase in invasive breast cancer incidence during the 1980s has been attributed to increased

detection through screening,[5] and probably nearly all of the excess of DCIS cases in the 1980s compared with earlier years can be too.

Time Trends in DCIS Treatment and Current Estimates of Numbers of Women Treated by Mastectomy versus Lumpectomy

Almost all DCIS is treated surgically, either by mastectomy or by lumpectomy with or without radiation; according to SEER data for 1993, only 1.7 percent of patients with DCIS did not have surgery. As shown in Table 1, for the period 1983–1993, the proportion of cases treated by mastectomy declined

TABLE 1. Estimated Numbers of DCIS Cases, Percent Treated by Mastectomy, and Estimated Numbers of Mastectomies for DCIS in All U.S. Women and in Women Ages 40–49, 1983–1993

	Estimated Number of DCIS Cases		% Cases Treated by Mastectomy		Estimated Number of Mastectomies	
Year	All Women	Ages 40–49	All Women	Ages 40–49	All Women	Ages 40–49
1983	4,901	742	71.0	75.8	3,479	563
1984	7,069	1,433	66.6	67.7	4,706	971
1985	9,897	1,991	59.5	57.5	5,887	1,144
1986	12,279	2,283	56.1	57.0	6,890	1,300
1987	16,034	3,000	59.3	62.8	9,515	1,884
1988	17,196	3,345	57.8	56.3	9,934	1,882
1989	16,584	3,086	56.3	53.0	9,334	1,635
1990	19,890	3,970	53.7	51.0	10,682	2,025
1991	20,735	4,325	47.8	44.2	9,908	1,912
1992	23,438	4,973	43.8	45.3	10,265	2,250
1993	23,275	4,676	39.7	40.4	9,245	1,890
Total	171,298	33,824			89,845	17,456

substantially, from 71 percent to 39.7 percent for women of all ages combined and from 75.8 percent to 40.4 percent for women ages 40–49 years. There is marked variation in DCIS treatment patterns across SEER areas; in 1993, 58.8 percent of cases were treated by mastectomy in Utah compared with only 28.0 percent in Connecticut. Extrapolating from SEER incidence rates and treatment patterns to the general U.S. population, an estimated 9,245 mastectomies were performed for DCIS in the United States in 1993, of which 1,890 were in women ages 40–49 years; there were an additional estimated 516 mastectomies for DCIS in women under 40 years of age. Between 1983 and 1993, an estimated 17,456 breasts were removed for DCIS in U.S. women ages 40–49 years (89,845 for women of all ages combined), and presumably most of those cases were detected by mammography.

Directions for Future Research

It is agreed that most of the increase in reported cases of DCIS results from better detection of the disease through mammography rather than a true excess of new cases. Given the numbers of women diagnosed with DCIS in recent years, the need for understanding the relationship of mammographically detected DCIS to invasive and potentially life-threatening breast cancer is urgent.

DCIS shares at least some risk factors and genetic changes in common with invasive breast cancer, which suggests etiologic similarities and supports the position that at least some DCIS cases are precursors to invasive disease. Other evidence suggests that many, if not most, cases of DCIS are not clinically significant; in most autopsy series examined, occult DCIS is not uncommon in women who died of causes other than breast cancer, and small historical series of women with DCIS who received no treatment beyond diagnostic biopsy show that most did not develop clinically apparent invasive breast cancer, at least in the first decade or so following biopsy.[4] Thus, biologic and epidemiologic studies to identify prognostic markers and risk factors associated with progression, focusing on specific histologic types of DCIS and perhaps correlated with breast imaging studies, are to be encouraged.

Better information about the appropriate treatment of DCIS is also needed to reduce the confusion and uncertainty many women and their physicians currently experience in the face of a DCIS diagnosis. Good evidence about whether detecting and treating breast cancer at the DCIS stage confers a survival advantage is lacking, although designing such a study would be ethically and logistically difficult. At least one observational study of minimal treatment (local excision) for what is considered to be favorable prognosis DCIS is about to begin, which should provide useful information for that subset of DCIS. For the present, informed decision-making about screening mammography should include the likelihood of being diagnosed with DCIS, with an explanation that only some DCIS cases may be clinically significant, as well as the likelihood of having breast surgery as a result of DCIS detection.

References

1. Surveillance, Epidemiology, and End Results (SEER) Program public use CD-ROM (1973–1993). Bethesda (MD): National Cancer Institute, DCPC, Surveillance Program, Cancer Statistics Branch, 1996.

2. Trends in cancer screening—United States, 1987 and 1992. MMWR Morb Mortal Wkly Rep 1996; 45:57–61.

3. Brown ML, Kessler LG, Rueter FG. Is the supply of mammography machines outstripping need and demand? An economic analysis. Ann Intern Med 1990;113:547–52.

4. Ernster VL. Epidemiology and natural history of ductal carcinoma *in situ*. In: Silverstein MJ, editor. Ductal carcinoma *in situ* of the breast: a diagnostic and therapeutic dilemma. Baltimore: Williams and Wilkins. In press.

5. Wun LM, Feuer EJ, Miller BA. Are increases in mammographic screening still a valid explanation for trends in breast cancer incidence in the United States? Cancer Causes Control 1995;6:135–44.

6. Choi WS, Parker BA, Pierce JP, Greenberg ER. Regional differences in the incidence and treatment of carcinoma *in situ* of the breast. Cancer Epidemiol Biomarkers Prev 1996;5:317–20.

7. Ernster VL, Barclay J, Kerlikowske K, Grady D, Henderson IC. Incidence and treatment for ductal carcinoma *in situ* of the breast. JAMA 1996;275:913–8.

8. Eusebi V, Foschini MP, Cook MG, et al. Long-term follow-up of *in situ* carcinoma of the breast with special emphasis on clinging carcinoma. Semin Diagn Pathol 1989;6:165–73.

9. Hetelekidis S, Schnitt SJ, Morrow M, Harris JR. Management of ductal carcinoma *in situ*. CA 1995;45:244–53.

10. Morris CR, Wright WE, eds. Breast cancer in California. Sacramento (CA): California Department of Health Services, Cancer Surveillance Section, 1996.

11. Nielsen M, Thomsen JL, Primdahl S, Dyreborg U, Andersen JA. Breast cancer and atypia among young and middle-aged women: a study of 110 medicolegal autopsies. Br J Cancer 1987;56:814–9.

12. Page DL, Dupont WD, Rogers LW, Jensen RA, Schuyler PA. Continued local recurrence of carcinoma 15–25 years after a diagnosis of low grade ductal carcinoma *in situ* of the breast treated only by biopsy. Cancer 1995;76:1197–1200.

13. Steele J, Jessup JM, Winchester DP, Menck HR, Murphy GP. National cancer data base: annual review of patient care 1995, 1995.

14. Winchester DP, Menck HR, Osteen RT, Kraybill W. Treatment trends for ductal carcinoma *in situ* of the breast. Annals Surg Oncol 1995;2:207–13.