Adventures in Statistics

Attributing reductions in breast cancer mortality to treatment vs screening: A Bayesian CISNET model

Donald A. Berry
dberry@mdanderson.org
Adventures in Statistics

• Clinical trials
  • Part 1: Bayesian bandits to platforms (Sept 28)
  • Part 2: Embracing adaptive Bayesian clinical trials (Oct 26)
  • Part 3: Regulators & CID, Complex Innovative Designs (today)

• Other adventures
  • Attributing ancestry: The corny but $B$ story of genetics of maize ... & Bayes
  • DNA profiling
  • The science of doping ... or lack thereof
  • BRCAPRO: Risk model for carrying a deleterious mutation of BRCA1 or BRCA2 based on family history ... a marriage of Mendel and Bayes
  • Attributing reductions in breast cancer mortality to treatment vs screening: A Bayesian CISNET model (early ABC—Approximate Bayesian Computation)
  • Multiplicities, $p$-values, & observational studies
OUTLINE

Setting the stage for CISNET
• Screening mammography, a hot potato politically
• Misunderstandings of “early detection”
• Lead-time and length biases

CISNET to the rescue!
• Models D, E, G, M, R, S, W
• Model M & “Approximate Bayesian Computation”

• January’s Hero: Esther Tomljanovich
• February’s Hero: Mary Claire King
• March’s Hero: Fran Visco, President of National Breast Cancer Coalition
Senator Tom Harkin’s words

• “Mammography is a useful tool for early detection. The earlier detected, the better your prognosis is going to be.” [For emphasis: shrugging with arms up]

• I once wrote an editorial for the British Journal of Cancer entitled, “The Screening Mammography Paradox: Better when Found, Perhaps Better Not to Find”

• Lead-time and length bias
Sojourn times

Length Bias

First screen  Second screen  Truncation point
Sojourn times
Mean 1 year
Mean lengths
33 screen detected: 1.45 yrs
21 interval cancers: 0.29 yrs
Mean survival (assuming ...)
54 if no screening: 20 yrs
33 screen detected: 29 yrs
21 interval cancers: 5.8 yrs

Length Bias
Length Bias

• Waiting for a bus (waiting time bias)
• Falling stars (falling star bias)
• Potato chips (potato chip bias)
Why is screening mammography so compelling?

• Finding early must be good (Senator Harkin!)
• Notion of “cure” & “We got it all”
• Clinical practice and lead-time and length biases
• Reality check: Randomized trials
A recent study that tried to assess the usefulness of mammography among 80- and 90-year-olds found that very few women in this age group, 22 percent, underwent regular screenings for breast cancer, but that those who did were more likely to find the cancer early enough to avoid a mastectomy and survive at least five years [88% vs 82%].
The *Journal of Clinical Oncology* erred in publishing this article. It was a disservice to women, young as well as old. Such publication reveals a seriously defective editorial process at the journal.

**Donald A. Berry**
Division of Quantitative Sciences, The University of Texas M.D. Anderson Cancer Center, Houston, TX

**Cornelia J. Baines**
Department of Public Health Sciences, University of Toronto, Toronto, Ontario, Canada

**Michael Baum**
Departments of Surgery and Medical Humanities, University College, London, United Kingdom

**Kay Dickersin**
Department of Epidemiology; and US Cochrane Center, Johns Hopkins University, Baltimore, MD
A single blood test claims to detect dozens of cancers. Skeptics wouldn’t bet your life on it.

A counterintuitive problem with some cancer screening tests has been identified: Early detection is not always the best thing because some tumors never become harmful.

Five months ago, Grail Inc. began marketing the first-ever prescription test intended to detect more than 50 types of cancer — most of which have no recommended screening test — with a mere blood sample.
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?

2) What Are the Risks Associated With Screening Women Ages 40-49 With Mammography and With or Without Physical Examination? How Large Are the Risks? How Do They Change With Age?

3) Are There Other Benefits? If So, What Are They? How Do They Change With Age?

4) What Is Known About How the Benefits and Risks of Breast Cancer Screening Differ Based on Known Risk Factors for Breast Cancer?

5) What Are the Directions for Future Research?
CONCLUSIONS: The Panel concludes that the data currently available do not warrant a universal recommendation for mammography for all women in their forties. Each woman should decide for herself whether to undergo mammography.
What were the “data” in 1997?
Total about 5 days
Aside: ERSPC Schröder NEJM 2012
A way to understand risks is to relate them with risks that are familiar. For example, the estimated average of 5 days of life lost if a woman in her early forties delays mammography for 10 years is similar to that for not wearing a seat belt over 20 years of typical automobile travel, of riding a bicycle for 15 hours without a helmet (or 50 hours if wearing a helmet), and of gaining two ounces of body weight (and keeping it on) (41).
In spite of the “data,” and the expert panel’s conclusion, the U.S. Senate ...

• Voted 98-0 dictating that mammography would be effective (!) for women in their 40s

• Withheld the NCI’s budget until the NCI agreed to recommend screening for women in their 40s
Between 1997 and 2009

• 2001: Cochrane review led to political flare-ups
• 2002: Swedish trial updates
• 2005: CISNET models published in NEJM
• 2006: U.K. “Age” trial announced
• 2007: ACP’s guidelines (essentially 1997!)
• But mostly an armistice from 2002 to 2009
Berry JNCI 1998 ... with updates
USPSTF 2009

“Recommendations: The USPSTF recommends against routine screening mammography in women aged 40 to 49 years. The decision to start regular, biennial screening mammography before the age of 50 years should be an individual one and take into account patient context, including the patient’s values regarding specific benefits and harms. (Grade C recommendation).”
USPSTF 2009

“These findings were combined in an updated meta-analysis ... RR for breast cancer death of 0.85 (CI, 0.75 to 0.96; 8 trials) ... number needed to invite for screening of 1904 (CI, 929 to 6378) to prevent 1 breast cancer death in women aged 39 to 49 years.

“... 6 trials among women aged 50 to 59 years ... RR ... of 0.86 (CI, 0.75 to 0.99; number needed to invite, 1339 [CI, 322 to 7455]) ....”

My complaint: “Not adhering to the planned schedule of mammograms in the control group is a major flaw in the conduct and ensuing analysis of four Swedish trials.”

Nyström 2002 (Swedish trials metaanalysis):
- Evaluation analysis: 0.80 (CI: 0.63 to 1.01)
- Follow-up analysis: 0.91 (CI: 0.76 to 1.09)
Reactions

• Obama critics charged “rationing healthcare” and “death squads”

• *The Cancer Letter*: Political pressure, name-calling and accusations over mammography grew so strong that two days after the recommendation was published, HHS Secretary Kathleen Sebelius distanced the Obama administration from the non-partisan group of public health experts who comprise the USPSTF.
More reactions ...

- USPSTF Downgrades PSA Screening from "I" to "D" — As In "Don't Do It"
- Delays in Taking a Vote Suggest Political Meddling with USPSTF
Media Reactions


- *LA Times*: Mammogram guidelines spark heated debate. A government panel's recommendation that women under 50 do not need regular mammograms is attacked by oncologists, gynecologists and cancer groups

- *Boston Globe*: Breast screening advice upended
Brawling Over Mammography

A scientific study of the benefits and harms of screening women in their 40s got buried by politics

The Obama Administration is a self-described champion of science. But it was put on the spot last fall when it received a scientific report that questioned the value of routine mammograms. For breast cancer, the American College of Radiology, Daniel Kopans of Massachusetts General Hospital in Boston, was less restrained. He circulated a critique saying that the sponsor of the new report, the U.S. Preventive Services Task Force, failed to consider new evidence of the benefits of mammography.

Researchers who had worked on the USPSTF guidelines were disappointed that their analysis was being dismissed out of hand. “Politics got in the way of the science and the best public health practice,” says Jeanne Mandelblatt, an M.D.-epidemiologist at Georgetown University in Washington, D.C., and first author of an analysis for USPSTF by six groups that compared models to find the best screening strategy. “It was very unfortunate,” adds Heidi Nelson, an M.D.-epidemiologist at the Oregon Health & Science University in Portland, who led a separate team that gathered evidence for USPSTF.

Karel Krulikowska, a breast cancer specialist at the University of Wisconsin, says there’s still no consensus on the best way to screen women for breast cancer. The recent American Cancer Society guidelines suggest getting mammograms every 5 years starting at age 40, whereas USPSTF recommends screening women every 2 years starting at age 40. Researchers in the other studies have their own views on the optimum strategy.
Otis Brawley, American Cancer Society’s CMO

ACS will continue to recommend yearly mammograms for women in their 40s.

Dr. Brawley said the task force “is telling women that mammography saves lives -- just not enough of them to recommend that all women over 40 get screened.”
“Number needed to invite” for screening to extend one life
CANCER INTERVENTION AND SURVEILLANCE MODELING NETWORK
Contributions of CISNET modeling to Task Force recommendations

Based on consistency with randomized trials in overall benefit, and by age:

• Biennial vs annual screening
• Women in 40s
• False positives/biopsies
• Context of chemo- and hormonal therapy
Effects of Mammography Screening Under Different Screening Schedules

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

Initiating biennial screening at age 40 years (vs. 50 years) reduced mortality by an additional 3% (range, 1% to 6%), consumed more resources, and yielded more false-positive results.

Data Sources: National data on age-specific incidence, competing risks, and mortality.

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

Results of Base-Case Analysis: The 6 models produced consistent rankings of screening strategies. Screening biennially maintained an average of 81% (range across strategies and models, 67% to 99%) of the benefit of annual screening with almost half the number of mammograms.

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

Primary Funding Source: National Cancer Institute.
Effect of Screening and Adjuvant Therapy on Mortality from Breast Cancer

Donald A. Berry, Ph.D., Kathleen A. Cronin, Ph.D., Sylvia K. Plevritis, Ph.D.,
Dennis G. Fryback, Ph.D., Lauren Clarke, M.S., Marvin Zelen, Ph.D.,
Jeanne S. Mandelblatt, Ph.D., Andrei Y. Yakovlev, Ph.D., J. Dik F. Habbema, Ph.D.,
and Eric J. Feuer, Ph.D., for the Cancer Intervention and Surveillance
Modeling Network (CISNET) Collaborators*

ABSTRACT

BACKGROUND
We used modeling techniques to assess the relative and absolute contributions of screening mammography and adjuvant treatment to the reduction in breast-cancer mortality from 1976 to 1995.

RESULTS
The proportion of the total reduction in the rate of death from breast cancer attributed to screening varied in the seven models from 28 to 65 percent (median, 46 percent).
United States Breast Cancer Mortality
(Age Standardized to US Population in 2000)

24%
CISNET Population Models

Common inputs

- Background trends
- Screening behavior
- Diffusion of new treatments
- Other common inputs

Simulation or analytical models

7 different breast cancer models

BC incidence & mortality
CISNET from NEJM

A

Women 40-79

Percentage Undergoing Screening Mammography

Year


Never Irregular Every 2 yr Annual

B

Node-positive BC

Percentage Receiving Treatment

Year


Multiagent chemotherapy only

Both

Tamoxifen only
CISNET from NEJM
## Apportioning Effects of Interventions

<table>
<thead>
<tr>
<th>%Total Effect</th>
<th>Factorial Runs</th>
<th>Mean</th>
<th>Std Dev</th>
</tr>
</thead>
<tbody>
<tr>
<td>39.2% Screening</td>
<td>7.11</td>
<td>3.01</td>
<td></td>
</tr>
<tr>
<td>37.3% Tamoxifen</td>
<td>6.77</td>
<td>2.98</td>
<td></td>
</tr>
<tr>
<td>24.5% Chemotherapy</td>
<td>4.44</td>
<td>2.38</td>
<td></td>
</tr>
<tr>
<td>−0.7% Screening x Tamoxifen</td>
<td>−0.12</td>
<td>0.87</td>
<td></td>
</tr>
<tr>
<td>−0.2% Screening x Chemotherapy</td>
<td>−0.03</td>
<td>0.87</td>
<td></td>
</tr>
<tr>
<td>−0.2% Tamoxifen x Chemotherapy</td>
<td>−0.04</td>
<td>0.68</td>
<td></td>
</tr>
<tr>
<td>0.0% Screening x Tam x Chemo</td>
<td>−0.01</td>
<td>0.41</td>
<td></td>
</tr>
<tr>
<td>100.0% Total Effect</td>
<td>18.13</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Percent reductions in BC mortality due to adjuvant Rx and screening

<table>
<thead>
<tr>
<th>Model</th>
<th>Tamoxifen</th>
<th>Chemotherapy</th>
<th>Both therapies</th>
<th>Screening</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td>D</td>
<td>6.1</td>
<td>6.1</td>
<td>12.0 (35%)</td>
<td>22.7 (65%)</td>
<td>32.9</td>
</tr>
<tr>
<td>E</td>
<td>12.0</td>
<td>9.6</td>
<td>20.9 (58%)</td>
<td>15.3 (42%)</td>
<td>30.9</td>
</tr>
<tr>
<td>G</td>
<td>7.7</td>
<td>7.0</td>
<td>14.6 (54%)</td>
<td>12.4 (46%)</td>
<td>24.9</td>
</tr>
<tr>
<td>M</td>
<td>10.7</td>
<td>9.5</td>
<td>19.5 (65%)</td>
<td>10.6 (35%)</td>
<td>27.5</td>
</tr>
<tr>
<td>R</td>
<td>NA</td>
<td>NA</td>
<td>19.0 (72%)</td>
<td>7.5 (28%)</td>
<td>25.6</td>
</tr>
<tr>
<td>S</td>
<td>8.9</td>
<td>6.9</td>
<td>14.9 (47%)</td>
<td>16.9 (53%)</td>
<td>29.9</td>
</tr>
<tr>
<td>W</td>
<td>12.5</td>
<td>8.9</td>
<td>20.8 (51%)</td>
<td>20.3 (49%)</td>
<td>38.3</td>
</tr>
</tbody>
</table>
Reductions (%) in BC mortality due to adjuvant Rx & screening
Percent reductions in BC mortality due to adjuvant Rx and screening
Conclusions

Screening lowers BC mortality
Population treatment benefit similar to clinical trials
Little evidence for synergy
Some model differences, reflecting modeling uncertainty
Overall robustness across models
NY Times Editorial

“What seems most important is that each team found at least some benefit from mammograms. The likelihood that they are beneficial seems a lot more solid today than it did four years ago, although the size of the benefit remains in dispute”
(Bayesian) Model M

1. Parameters
   a. By stage/age/ER status
      i. Baseline survival
      i. Chemotherapy effectiveness
      i. Tamoxifen effectiveness
      i. “Beyond stage shift”
   a. Age-period-cohort model

1. Prior distributions

1. Posteriors by rejection sampling:
   Fit to true mortality/incidence
Original M Modelers

Lurdas Inoue
Mark Munsell
John Venier
Greg Ball
Yu Shen
Donald Berry

Current M Modelers

Xuelin Huang
Yisheng Li
Juhee Song
Donald Berry
Simulated BC mortality, 66 acceptances
Tamoxifen Effect (marginal)

prior: $m = 0.28$, $s = 0.15$
posterior: $m = 0.39$, $s = 0.13$
Chemotherapy Effect (marginal)

prior: m = 0.14, s = 0.16
posterior: m = 0.13, s = 0.12
Effect of mammography, belong stage shift, stages 1 & 2

prior: m = 0.4 , s = 0.23
posterior: m = 0.27 , s = 0.17
Effect of mammography, belong stage shift, stages 3 & 4

prior: m = 0.25, s = 0.14
posterior: m = 0.24, s = 0.13
Parameter for baseline survival distribution

prior: $m = 0.9$, $s = 0.06$
posterior: $m = 0.87$, $s = 0.04$
Age-period-cohort model vs flat incidence

prior: $m = 0.5, s = 0.29$
posterior: $m = 0.62, s = 0.28$
## Apportioning Effects of Interventions

<table>
<thead>
<tr>
<th>%Total Effect</th>
<th>Factorial Runs</th>
<th>Mean</th>
<th>Std Dev</th>
</tr>
</thead>
<tbody>
<tr>
<td>39.2%</td>
<td>Screening</td>
<td>7.11</td>
<td>3.01</td>
</tr>
<tr>
<td>37.3%</td>
<td>Tamoxifen</td>
<td>6.77</td>
<td>2.98</td>
</tr>
<tr>
<td>24.5%</td>
<td>Chemotherapy</td>
<td>4.44</td>
<td>2.38</td>
</tr>
<tr>
<td>-0.7%</td>
<td>Screening x Tamoxifen</td>
<td>-0.12</td>
<td>0.87</td>
</tr>
<tr>
<td>-0.2%</td>
<td>Screening x Chemotherapy</td>
<td>-0.03</td>
<td>0.87</td>
</tr>
<tr>
<td>-0.2%</td>
<td>Tamoxifen x Chemotherapy</td>
<td>-0.04</td>
<td>0.68</td>
</tr>
<tr>
<td>0.0%</td>
<td>Screening x Tam x Chemo</td>
<td>-0.01</td>
<td>0.41</td>
</tr>
<tr>
<td>100.0%</td>
<td>Total Effect</td>
<td>18.13</td>
<td></td>
</tr>
</tbody>
</table>
Reductions (%) in BC mortality due to adjuvant Rx & screening

All 7 models
Accepted Model M simulations
% Reduction in Breast Cancer Mortality

Model M

Due to Treatment

Due to Screening

R, M, E, W, S, D
Effect of Screening Mammography on Breast-Cancer Mortality in Norway

Mette Kalager, M.D., Marvin Zelen, Ph.D., Frøydis Langmark, M.D., and Hans-Olov Adami, M.D., Ph.D.

ABSTRACT

BACKGROUND
A challenge in quantifying the effect of screening mammography on breast-cancer mortality is to provide valid comparison groups. The use of historical control subjects does not take into account chronologic trends associated with advances in breast-cancer awareness and treatment.

METHODS
The Norwegian breast-cancer screening program was started in 1996 and expanded geographically during the subsequent 9 years. Women between the ages of 50 and 69 years were offered screening mammography every 2 years. We compared the incidence and prevalence of breast cancer among women 50 years of age or older in Norway from 1973 through 2000, with similar-aged women in the United States, the United Kingdom, France, Australia, and Sweden.
Accepted Model M simulations
Additional Conclusions from Model M

- Bayesian approach is ideally suited for assessing uncertainty, including modeling uncertainty
- Bayesian approach encompasses other six models
- Having probability distributions of parameters is great for making predictions
Reality Check
Computer-Assisted Detection and Screening Mammography: Where’s the Beef?

Consider the growth of a breast cancer from the first tumor cell: after perhaps 27 doublings, the tumor becomes detectable on a mammogram. After approximately 29 doublings—but with substantial variability—the tumor becomes symptomatic or detectable other than by a mammogram. So the “lead time” provided by screening mammography is on the order of two doubling times. (Sometimes there is no lead time, as when cancers found in screening programs are interval cancers—detected between mammograms.) Regardless of the lead time, the window for finding cancers via mammographic screening is small in comparison with the life of the tumor.

Moreover, improving sensitivity may preferentially find less aggressive tumors, or it might find more of those tumors that would have otherwise revealed themselves as interval tumors. Neither type of increment in sensitivity could have much of an impact on the overall survival benefit of screening mammography.
When does the tumor become metastatic?
Nobody knows!

Mammographically detectable
Clinically detectable

Tumor doubling times

If here then screening is too late
If here then screening irrelevant

Mortality benefit of early detection only if here
7%
Some Challenges

• Which tumors become metastatic? (As of today, 15% of breast cancers are lethal ... which 15%?)
• Which are cured by surgery?
• Which “cancers” can we ignore (don’t need surgery)?
• Modeling neoadjuvant approaches
• What role for prevention?
• Estimating overdiagnosis
OUTLINE

Setting the stage for CISNET
• Screening mammography, a hot potato politically
• Misunderstandings of “early detection”
• Lead-time and length biases

CISNET to the rescue!
• Models D, E, G, M, R, S, W
• Model M & “Approximate Bayesian Computation”
Adventures in Statistics

• Clinical trials
  • Part 1: Bayesian bandits to platforms (Sept 28)
  • Part 2: Embracing adaptive Bayesian clinical trials (Oct 26)
  • Part 3: Regulators & CID, Complex Innovative Designs (today)

• Other adventures
  • Attributing ancestry: The corny but $B story of genetics of maize ... & Bayes
  • DNA profiling
  • The science of doping ... or lack thereof
  • BRCAPRO: Risk model for carrying a deleterious mutation of BRCA1 or BRCA2 based on family history ... a marriage of Mendel and Bayes
  • Attributing reductions in breast cancer mortality to treatment vs screening: A Bayesian CISNET model (early ABC—Approximate Bayesian Computation)

• Multiplicities, $p$-values, & observational studies
  April 26, 2022