

Overdiagnosis of Breast Cancer
Associated with Screening
Mammography

Andy Coldman

NAACCR 2014

Ottawa

Mammography Screening



Overdiagnosis

Background

Almost more than any other screening, mammography screening for breast cancer, has attracted much recent attention.

Why?

What has changed

- Evidence based reviews suggest lower levels of effectiveness than was generally accepted a decade ago
- Overdiagnosis, once not considered to be a problem in breast screening, is now a major concern
- To screen or not to screen has shifted from the question “Is it effective” to “What is the balance of harms and benefits”
- An effective lobby of breast screening skeptics has developed which have published studies purporting to demonstrate that screening is both less effective and more harmful than the conventional view

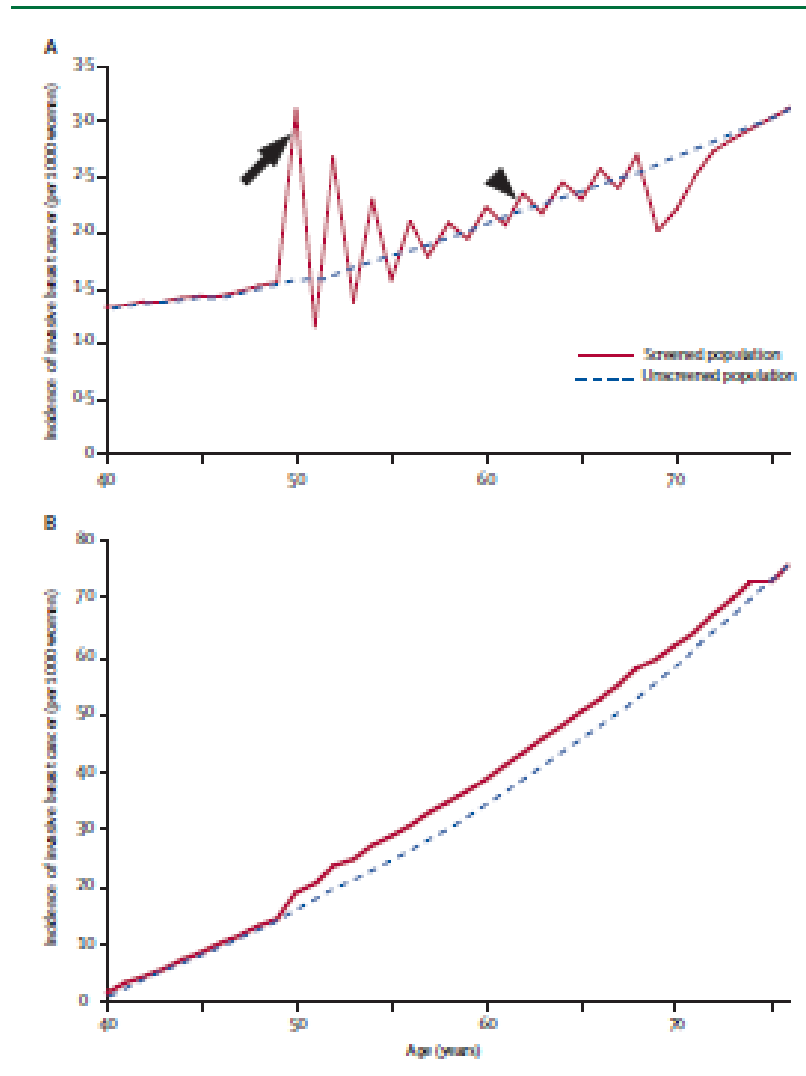
Overdiagnosis of Breast Cancer

Overdiagnosis is certainly possible as several autopsy studies have demonstrated the frequent presence of breast malignancy in women with no diagnosis prior to death

I will provide an overview of some of the studies which have examined overdiagnosis. I will confine attention to those using evidence from RCT's or Population studies. I will not review studies which have used sophisticated stochastic models.

Hypothetical Effect of Biennial Mammography Screening between Ages 50-69 on Incidence – No Overdiagnosis

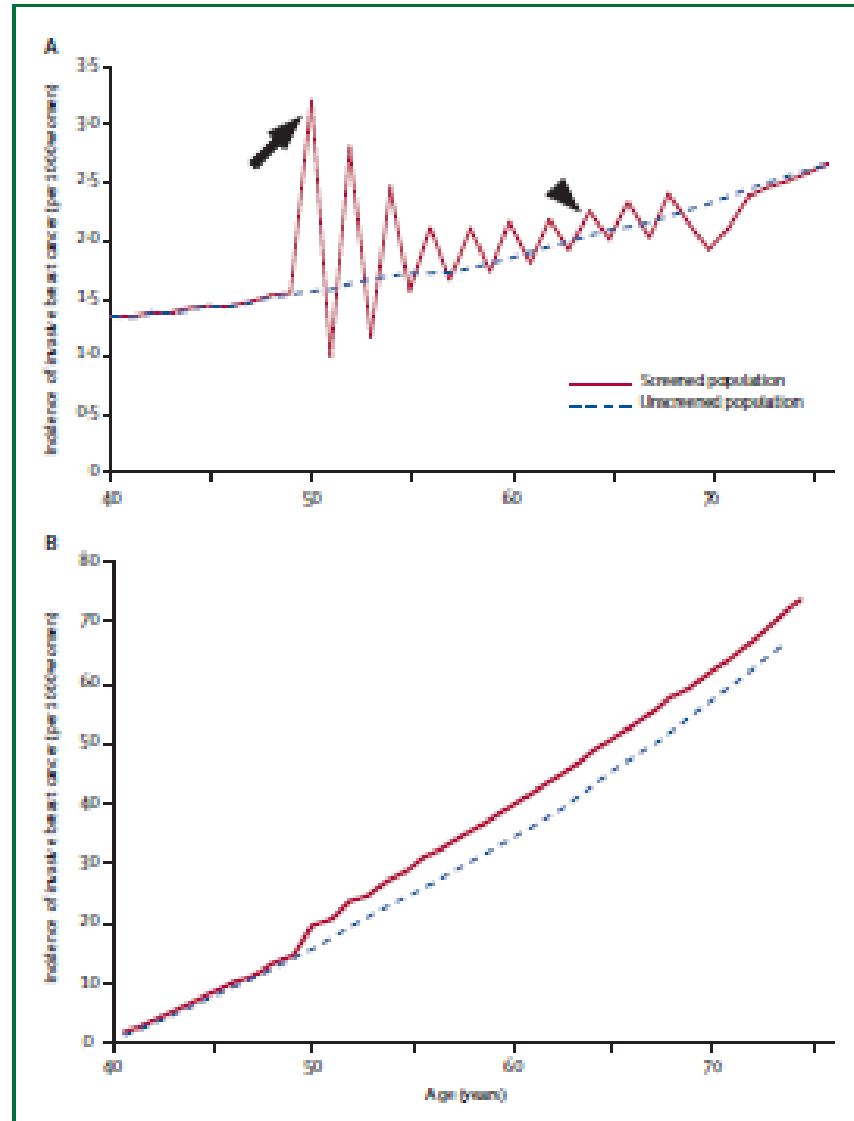
Incidence Rates



Cumulative Incidence Rates

Hypothetical Effect of Biennial Mammography Screening between Ages 50-69 on Incidence – Overdiagnosis Present

Incidence Rates



Cumulative Incidence Rates

Usual Metric for Expressing Overdiagnosis



$$\left[\begin{array}{c} \text{Screen Detected} \\ \text{Cases +} \\ \text{Incident Cases} \\ \text{In *Screened* Group} \end{array} \right] - \left[\begin{array}{c} \text{Incident Cases} \\ \text{In *Unscreened*} \\ \text{Group} \end{array} \right]$$

Incident Cases In *Unscreened* Group

Alternate Metric for Expressing Overdiagnosis



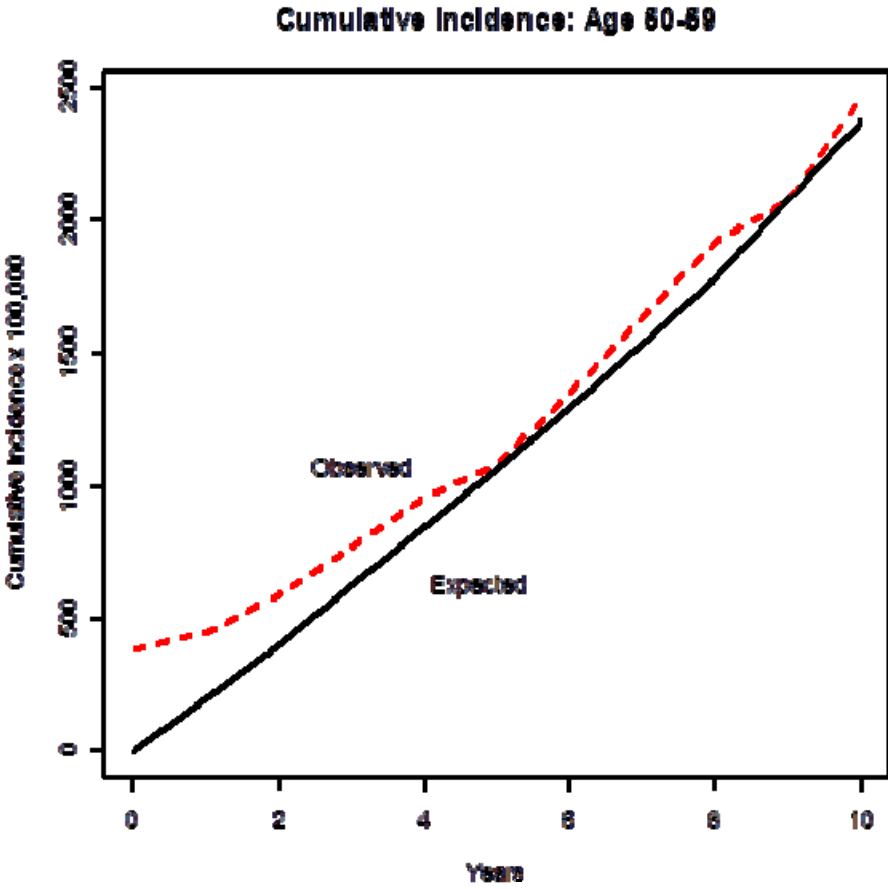
$$\left[\begin{array}{c} \text{Screen Detected} \\ \text{Cases +} \\ \text{Incident Cases} \\ \text{In *Screened* Group} \end{array} \right] - \left[\begin{array}{c} \text{Incident Cases} \\ \text{In *Unscreened*} \\ \text{Group} \end{array} \right]$$

Screen Detected Cases

RCT Measurement of Overdiagnosis

- RCT's provide the most valid and easily comprehended measurement of overdiagnosis
- Most importantly sufficient time must be given for catch-up, that is, we must allow sufficient elapsed time (lead time) to reach the clinical detection time for any cancers which were detected by screening. This means that the cohorts must be followed long after screening stops (this is longer than the follow-up required for measuring mortality effects).
- Fortunately several randomized trials of breast cancer screening have been conducted

Follow-up until Excess Risk Stabilizes



Screening of Controls

In most RCT's no screening was formally offered at the conclusion of screening in the trial but in some trials a screen was offered to the control group at the end of screening in the intervention group.

Screening in both arms serves to catch-up the control group so that prolonged follow-up of the two groups is no longer required to estimate overdiagnosis.

Results from Summary of Randomized Trials
[Moss et al, Breast Cancer research, 2005, p230]

Study	Cumulative Incidence of Breast Cancer per 1,000 WY				Screening Available at Study End
	Invasive Only		Invasive +DCIS		
	Control	Screen - Control	Control	Screen - Control	
Two County	2.54	-0.13	2.66	-0.02	Yes
Stockholm	0.85	-0.04	0.91	-0.03	Yes
Goteborg	1.73	-0.17	1.90	-0.11	Yes
Malmo 1	2.12	0.50	2.39	0.73	No
NBSS1	1.68	0.12	1.77	0.25	No
NBSS2	2.38	0.04	2.44	0.26	No

Comments on The RCT's

- RCT's with control screening do not provide estimates of overdiagnosis for screening versus none but for more (incident) screens (both groups have a prevalence screen)
- Without control screening long follow-up (>10 years) after cessation of screening is necessary to allow for lead-time effects but will inevitably lead to lower estimates of overdiagnosis since the denominator (cancers diagnosed) increases with time
- Use of screen detected cases as a denominator removes the preceding effect but increases the estimated overdiagnosis
- Long follow-up leads to increased variability but reduced bias in the estimation of overdiagnosis
- There is no model which relates the amount of overdiagnosis to screening patterns or patient characteristics

Population Based Studies

Most estimates of overdiagnosis associated with mammography come from studies based upon populations where mammography has been implemented.

These have used a variety of methods in which the observed rate of breast cancer is compared to some predicted rate.

The two issues which all population studies must address:

- How to estimate the counterfactual risk
- How to compensate for lead time

Some Estimates of Overdiagnosis from Population Studies

I will provide a short commentary on some studies:

US – Bleyer and Welch, NEJM 2012;367:1998-2005

Australia – Morrell et al, Cancer Causes & Control, 2010;21:275-82

Canada – Coldman & Phillips, CMAJ 2013;185: 492-8.

Europe - Puliti et al, J Medical Screening 2012;19(suppl):42-56

Bleyer and Welch,
NEJM 2012;367:1998-2005

ORIGINAL ARTICLE

Effect of Three Decades of Screening
Mammography on Breast-Cancer Incidence

Archie Bleyer, M.D., and H. Gilbert Welch, M.D., M.P.H.

ABSTRACT

Bleyer and Welch, USA

Approach Taken

- This study used SEER data to examine changes in breast cancer incidence in women 40+ between 1976-78 and 2006-08 (30 years).
- Assumed risk of breast cancer in women 40+ did not change (base case) with annual increases of 0.25% and 0.5% also considered
- Examined trends in localized and distant disease
- Used direct standardization of rates to 2000 population

Bleyer and Welch,

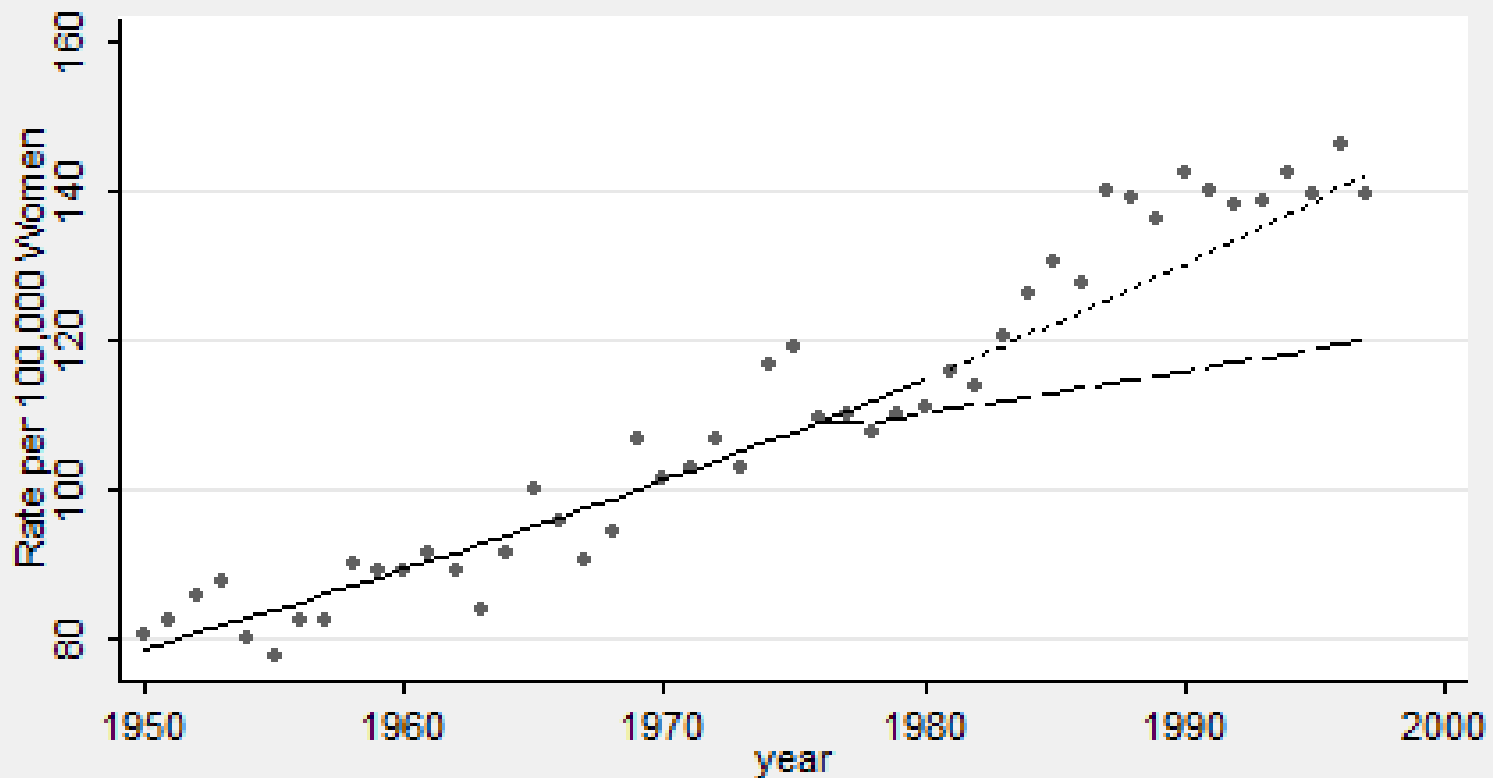
Main Finding

- Concluded that 31% of breast cancers diagnosed were overdiagnoses (base case – constant risk)

Comments

- Did not adjust for lead time
- Standardized incidence rates for unclear reasons
- Base case assumed constant breast cancer risk throughout the period

CISNET Counterfactual Incidence



- * CT State Tumor Registry Age-Adjusted Breast Cancer Incidence
- Fitted Annual Growth 1950-1980
- Fitted Annual Growth Extrapolated through 2000
- - - Bleyer & Welch 'Extreme' Growth from 1976-1978 Baseline

Effect of Differing Assumptions of Annual Increase in Breast Cancer Risk - SEER

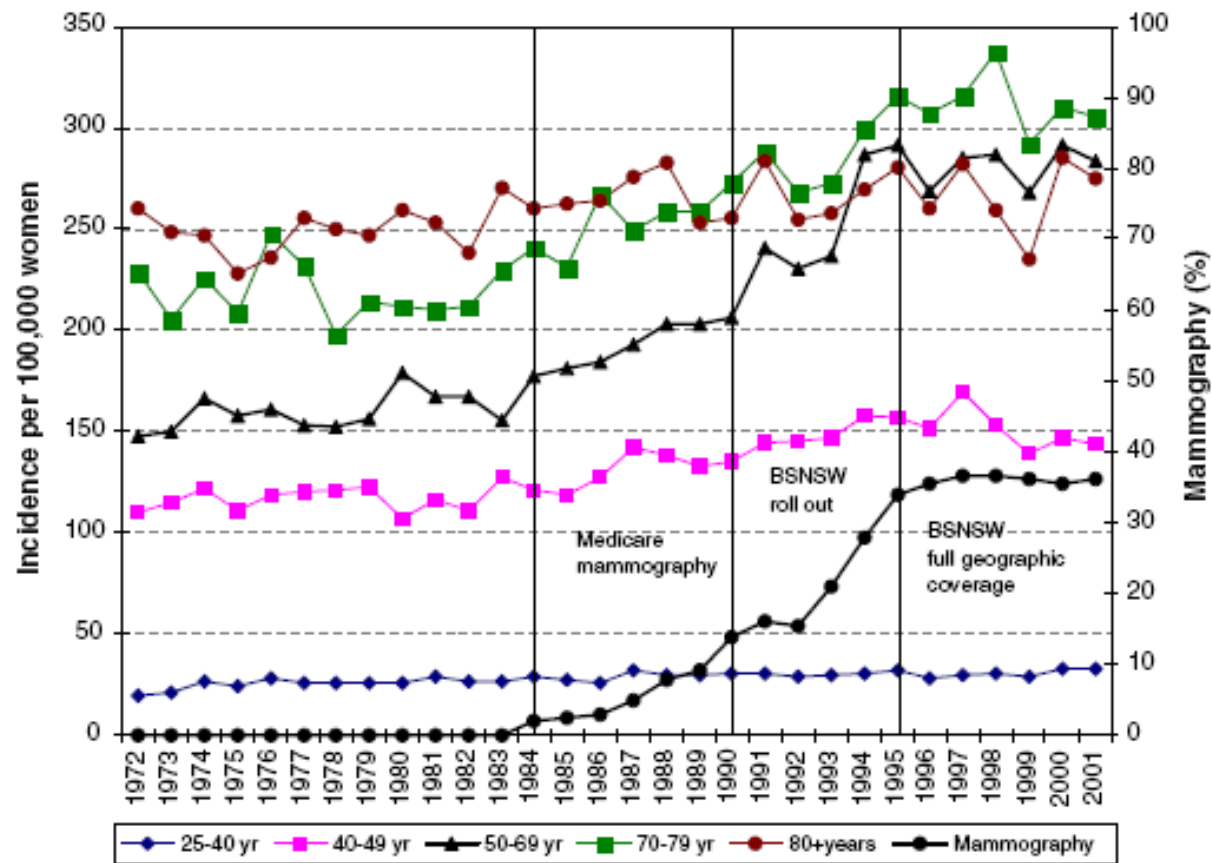
Source	Annual % Increase in Incidence prior to Initiation of Screening	Projected % Increase over 30 Years	Ratio of Observed to Projected Rate at end of 30 Year Period
0	0	0	1.53
B&W best guess	0.25	7.5	1.42
B&W Extreme	0.5	15.6	1.32
Europe	1.0	33.5	1.15
Conneticut	1.3	45.4	1.05

Morrell et al, Australia

Approach Taken

- This study used data from the New South Wales cancer registry to examine changes in breast cancer incidence in women 50-74 between 1972 and 2001.
- Used population data on risk factors (HRT use, obesity and nulliparity) to estimate changes in breast cancer risk over study period
- Also extrapolated incidence rates using two different approaches
- Adjusted for lead time using two values: 2.5 and 5 years.
- Estimated age-specific (5 year categories) and cumulative 50-69

Trends in Breast Cancer Incidence in New South Wales 1972-2001 (Morrell, Cancer C&C, 2010, p275)



Morrell et al,

Main Finding

- Estimated that overdiagnosis was between 30-42% depending on estimated counterfactual incidence
- Estimated overdiagnosis decreased with age: 53-35% at ages 50-54 to 21-15% at ages 65-69

Comments

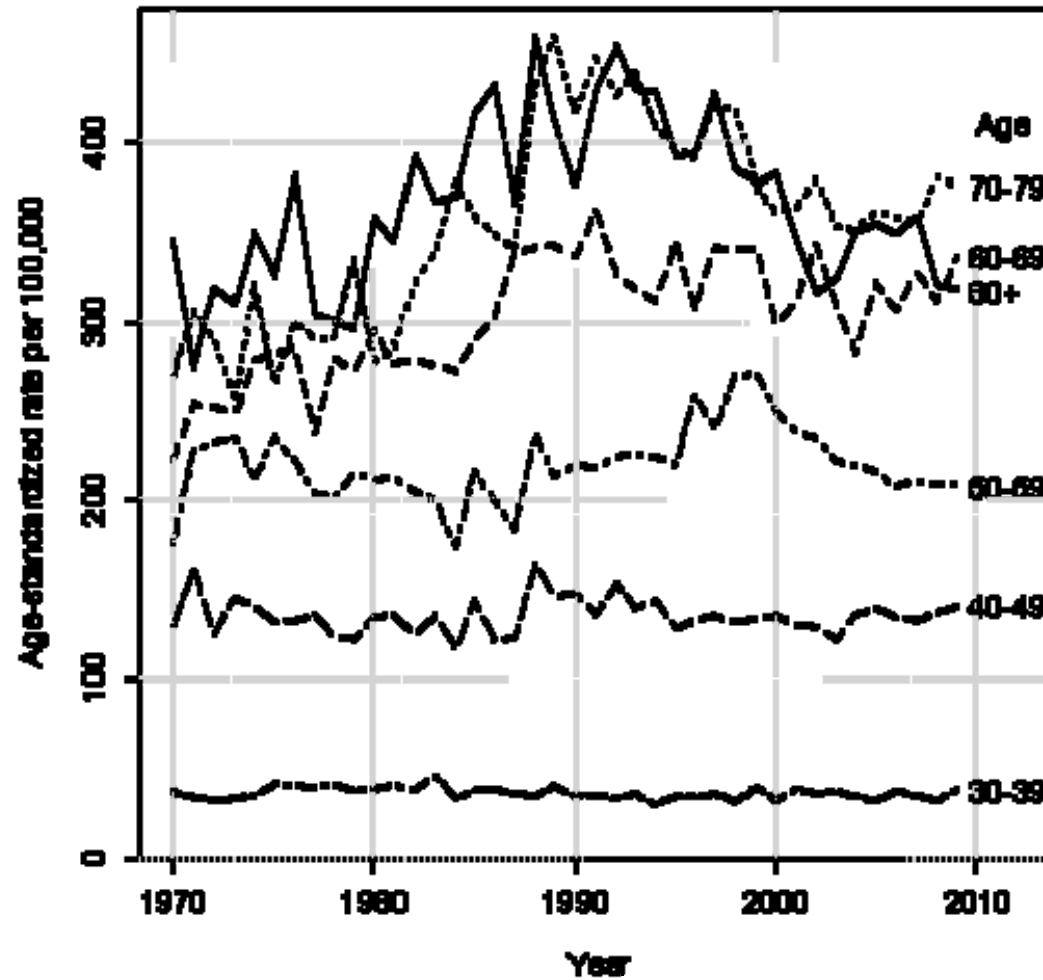
- Did not adjust for prevalence screen or include a compensatory drop
- Authors comment that they believe their estimates to be conservative (underestimates)
- Inverse age relationship is counter-intuitive and may reflect the influence of prevalence screens

Coldman and Phillips, Canada

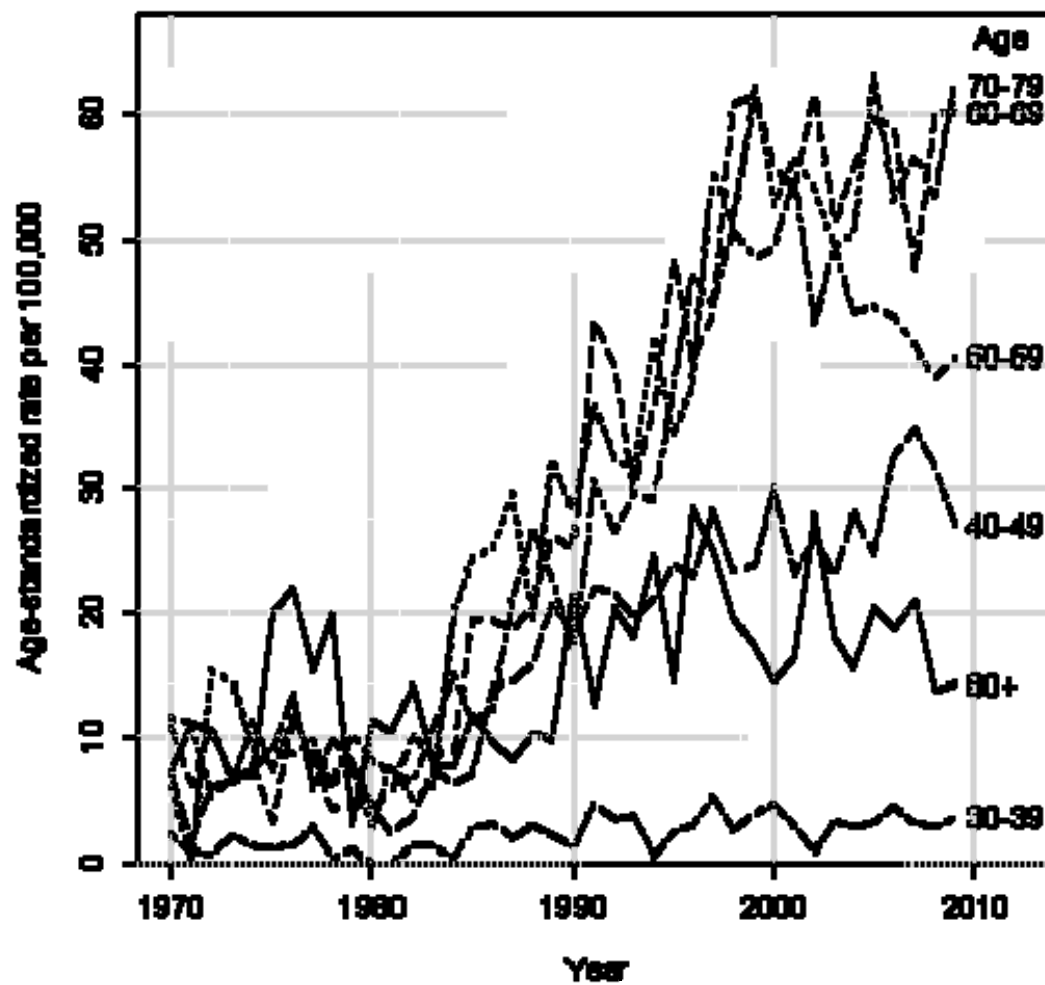
Approach Taken

- This study used data from the British Columbia Cancer Registry to examine changes in breast cancer incidence in women 50-74 between 1970 and 2009.
- Used data from population screening program to dynamically allocate the population and cancer cases by screening history
- Also used extrapolated incidence rates using two different approaches
- Used compensatory drop approach (10 year follow-up) and considered screening between the ages of 40-79
- Estimated overdiagnosis based upon screened versus un screened and also based on extrapolated population rates

Incidence of Invasive Breast Cancer by Age, 1970-2009, BC Registry



Incidence of DCIS by Age, 1970-2009, BC Registry



Coldman and Phillips, Canada

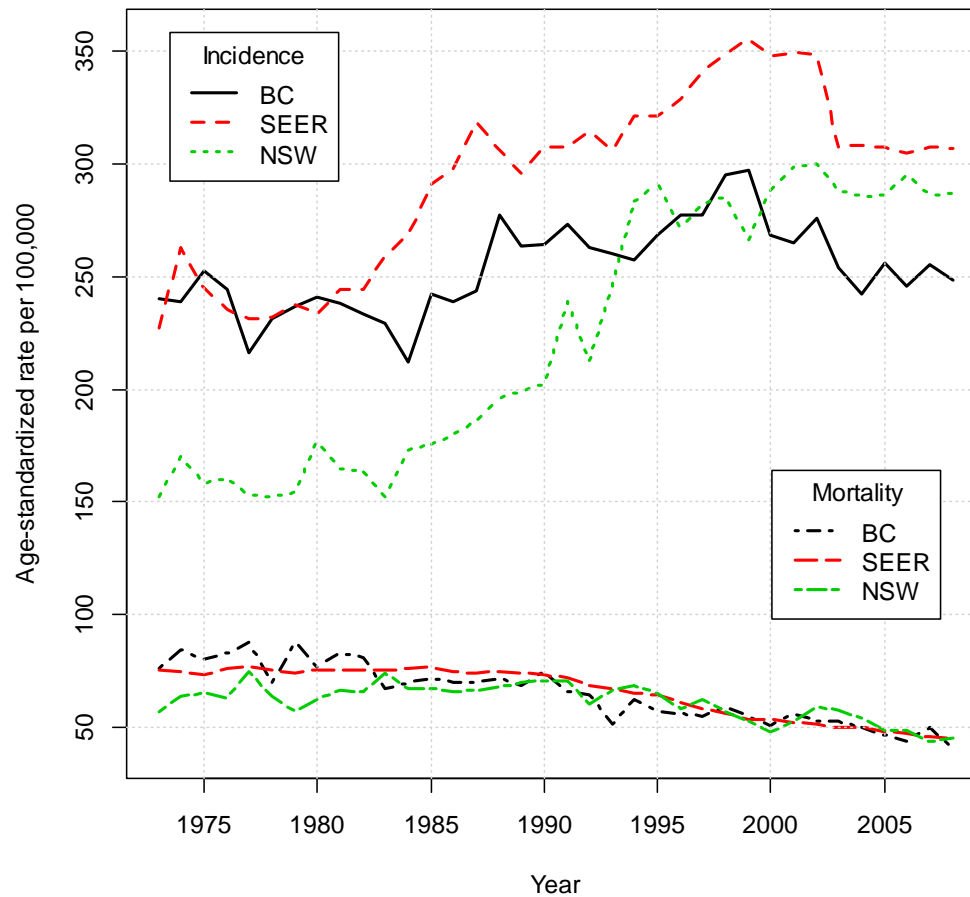
Main Finding

- Estimated that overdiagnosis of invasive cancer was between 5.4% and -0.7% depending on comparison (screened versus non or population based estimate)
- Inclusion of DCIS increased estimates of overdiagnosis to 6.7% and 17.3% respectively

Comment

- Used synthetic cohort to estimate effect of screening over large age range
- Validity of comparison of screened v non-screened uncertain
- Population based estimates of overdiagnosis had wide CI's
- Used similar time period to Bleyer and Welch analysis.

Comparison of Trends in 3 Jurisdictions: Age-standardized invasive breast cancer incidence and mortality trends age 50-69 in BC, New South Wales and US (SEER)

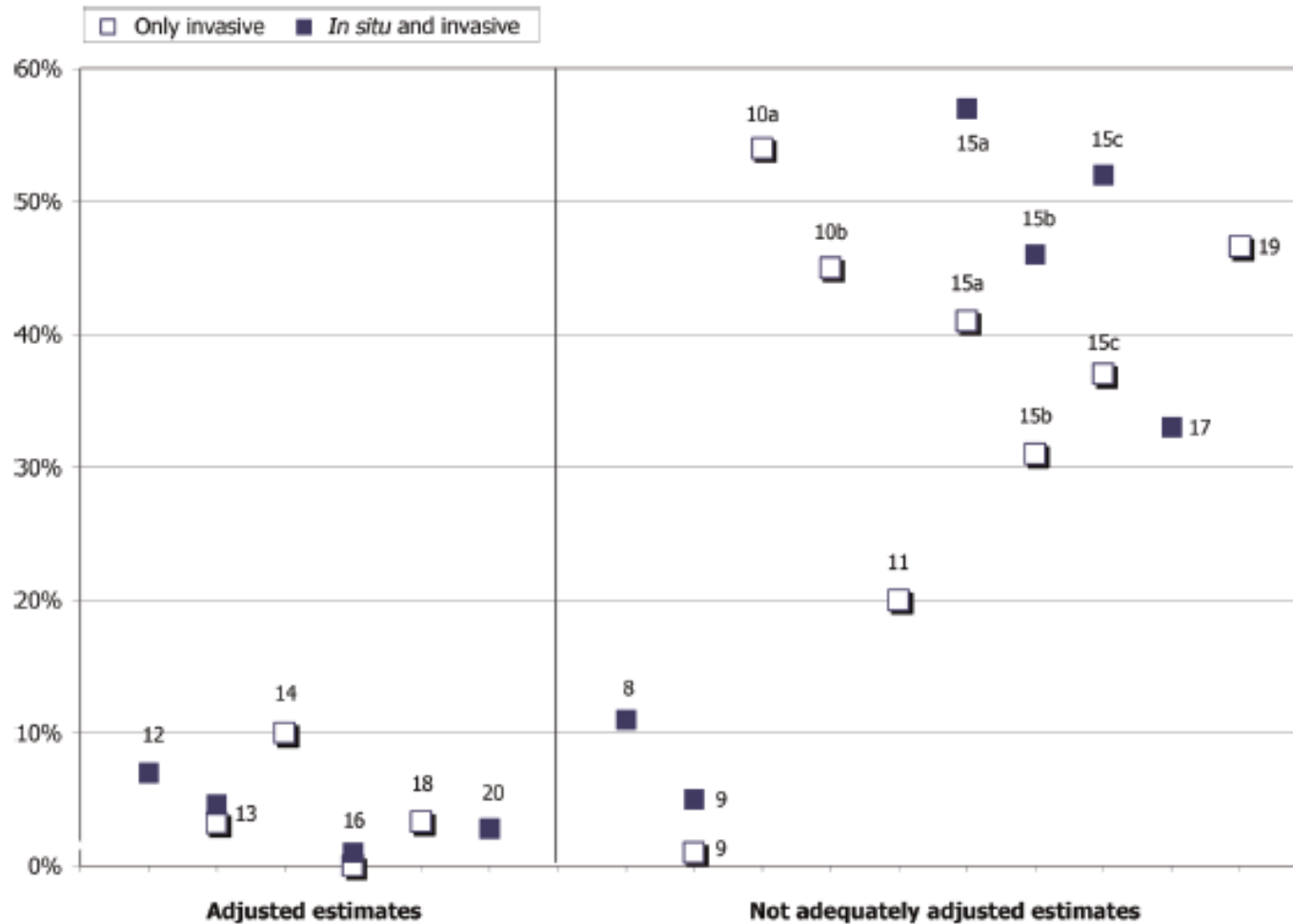


Puliti et al, Europe

Approach Taken

- This study reviewed published studies on overdiagnosis from European screening programs
- Sixteen studies satisfied eligibility criteria
- Studies were critically reviewed for the methods employed to estimate counterfactual rates and how adjustment was made for lead-time
- Studies were categorized as having adequate or not adequate adjustment for these two factors

Summary of Findings from Euroscreen Review



Puliti et al, Europe

Main Findings

- Reported rates of overdiagnosis ranged between 0 and 54%
- Among studies judged to be adequately adjusted for lead-time and counterfactual risk overdiagnosis ranged from 1.0 to 10.0%

Comment

- The authors provide a critical review of multiple studies using a common perspective
- Studies varied in the methods they employed to estimate overdiagnosis
- Inadequate adjustment lead to increased estimates of overdiagnosis

Conclusions

- Findings from randomized trials indicate that overdiagnosis is most associated with first screens
- Multiple approaches have been used to estimate overdiagnosis in population studies
- Different metrics are used to express overdiagnosis
- Population studies with adequate adjustment for lead time and counterfactual risk lead to estimates of <10% overdiagnosis which is similar to that from RCT's