Norwegian Institute of Public Health

Overdiagnosis of Clinicllay Irrelevant Cancer

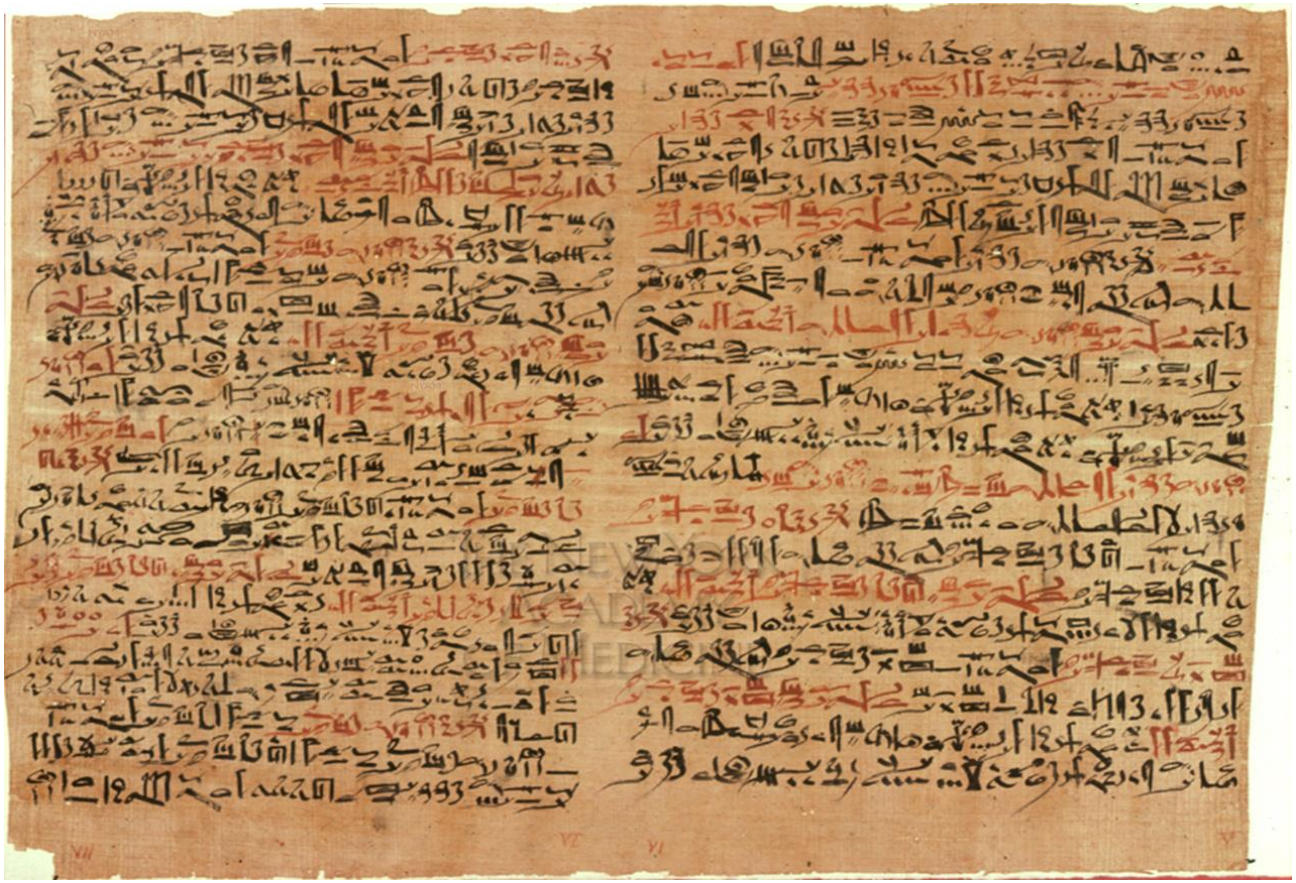
Br J Cancer 2013; 109: 2014-9.
XXX

Per-Henrik Zahl
Norwegian Institute of Public Health
Barcelona, September 20th, 2016

Summary

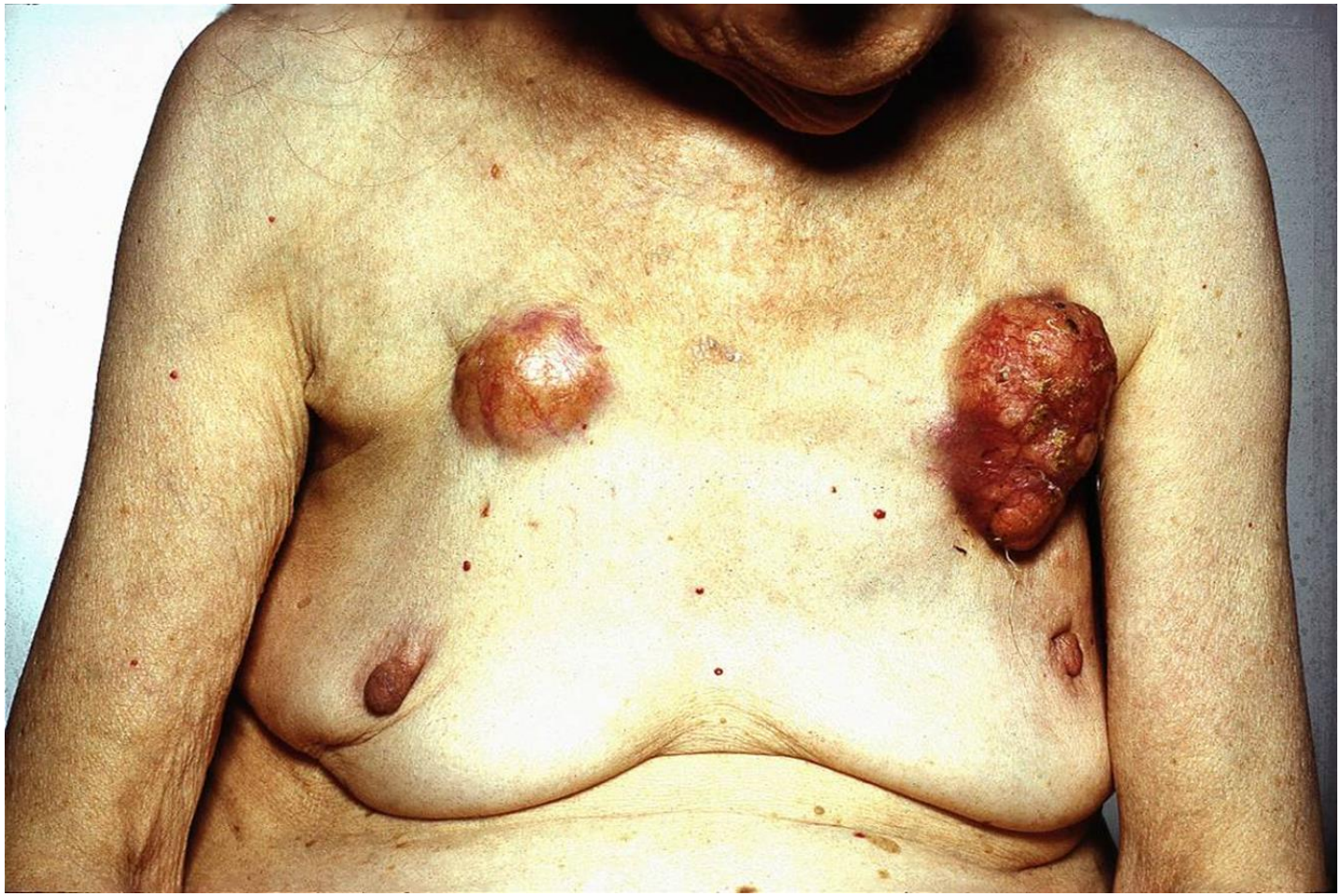
1. Short introduction to breast cancer
2. Growth rates
   - How to calculate tumor growth
   - When does dissemination starts?
3. Observed incidence when screening is not compensated by decline after screening
4. Alternative explanations to the incidence increase
5. What is lead time?
6. What is overdiagnosis?
7. Three methods to adjust overdiagnosis for lead time (early diagnosis)
8. Conclusion

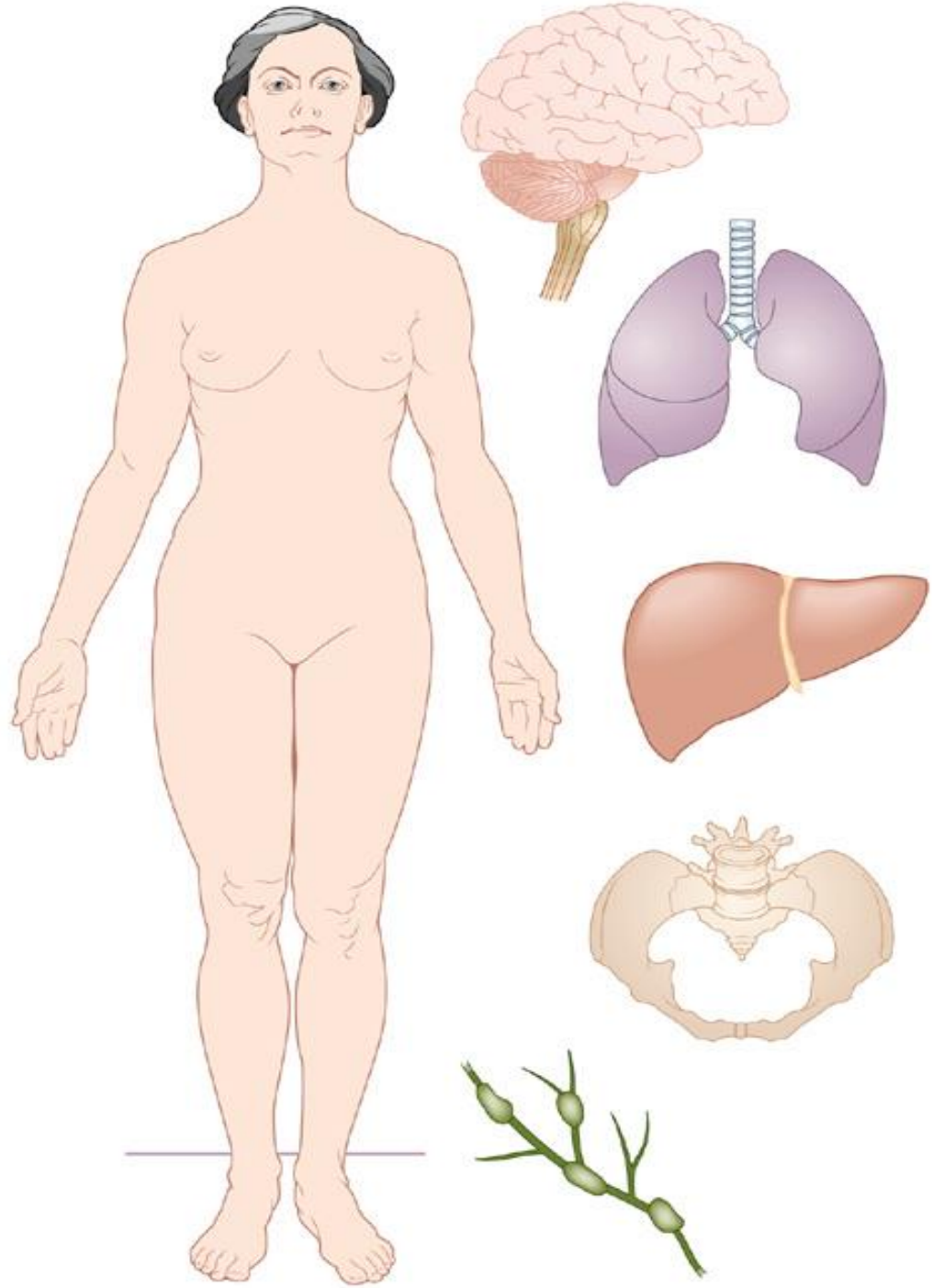
1. Short introduction to breast cancer

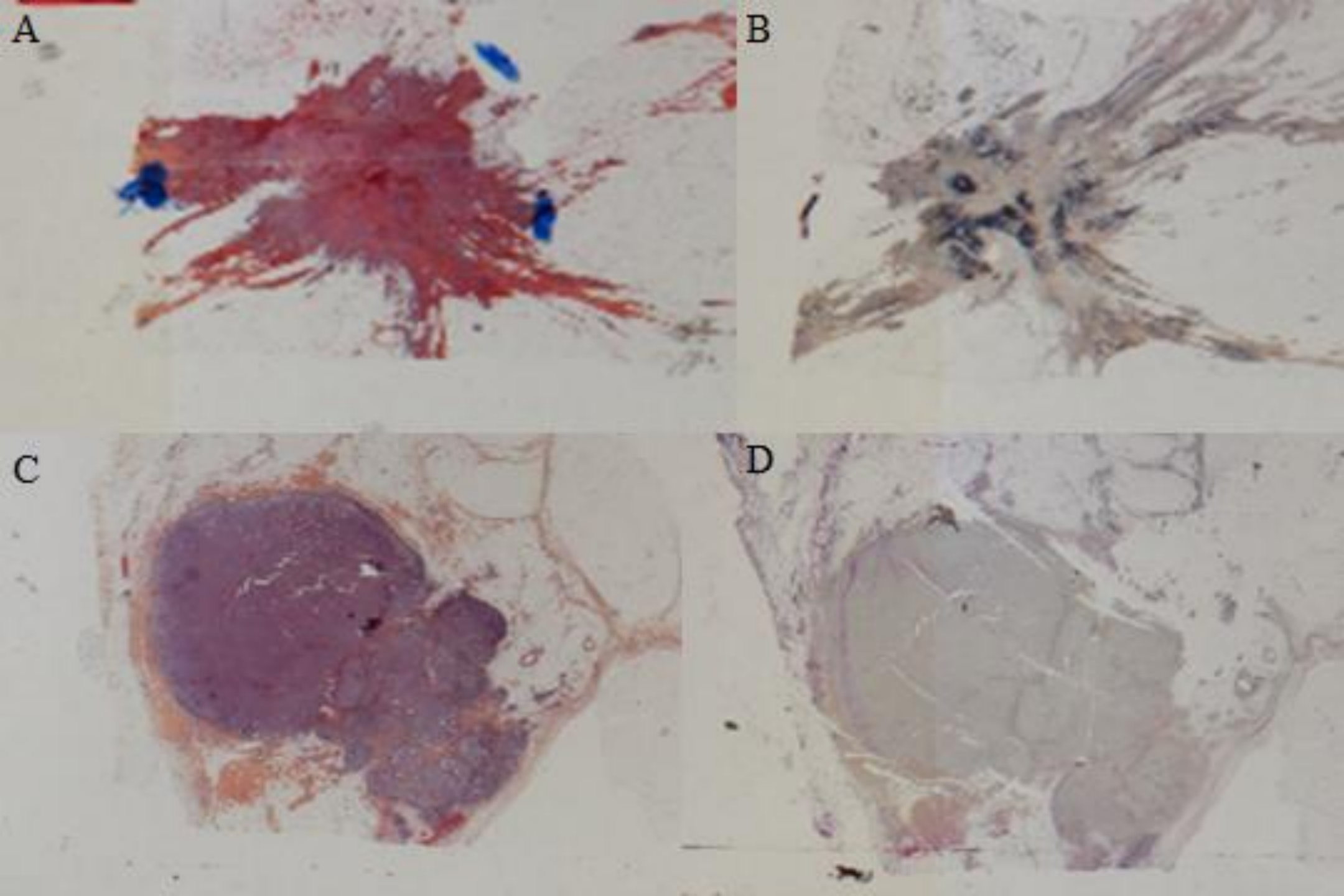


The Edwin Smyth papyrus

Primary tumor






A
B
C
D

2. The two most natural units for measuring tumor growth:

1. the ***number of cell divisions*** after the appearance of the first tumor cell which is measured by studying the absorption rate of radioactive elements
2. the tumor ***volume doubling time***, is calculated from two or more volumes observed by X-ray

Tumor volume doubling time is the same as the number of cell divisions when there is no cell death.

When does dissemination start?

Volume doubling times:

1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 (1mg), 21, 22, 23, 24, 25, 26, 27, 28, 29, 30 (1g), 31, 32, 33, 34, 35, 36, 37, 38, 39, dead.

Animal and human model studies show that dissemination starts early (10-19. volume doubling time) Folkman et al, nature 1989.

Window of opportunity (sojourn time): Mammography can maximally move time of diagnoses 2 doubling times but on average time of diagnosis is moved 1 doubling time (lead time)

Surgical treatment

1. Stage I (no metastases): Lumpectomy is equally efficient as radical mastectomy
2. Stage II (lymphatic metastases): Centennial node surgery + breast surgery
3. Stage III (huge local tumors): Drugs
4. Stage IV (distant metastases): Drugs

Primary effect is on stage I-II disease

Oncological treatment

1. Cytotoxins
2. Tamoxifen (anti-estrogene) – stops growth of metastatic disease
3. Herceptin (Trastuzumab) - monoclonal antibody blocking of HER2 receptors

Primary effect of 2-3 is on metastatic disease

Prerequisites for screening to work

1. That time of diagnosis can be moved sufficiently much backward
2. That treatment is more effective on an earlier time – note that treatment with cancer drugs primarily is effective on metastatic disease – it does not prevent dissemination of metastases

In theory: screening should not work

Results of 8 randomized mammography screening trials

- No reduction in total mortality
- 10-20% reduction in breast cancer mortality
- But no reduction in total cancer mortality
- 30% increase in breast cancer rates when screening

Gøtzsche & Jørgensen, Cochrane Review on Mammography Screening, 2011.
U.S. Preventive Services Task Force. Ann Intern Med 2009; 151;727-37.
Canadian Task Force. http://canadiantaskforce.ca/guidelines/2011-breast-cancer/

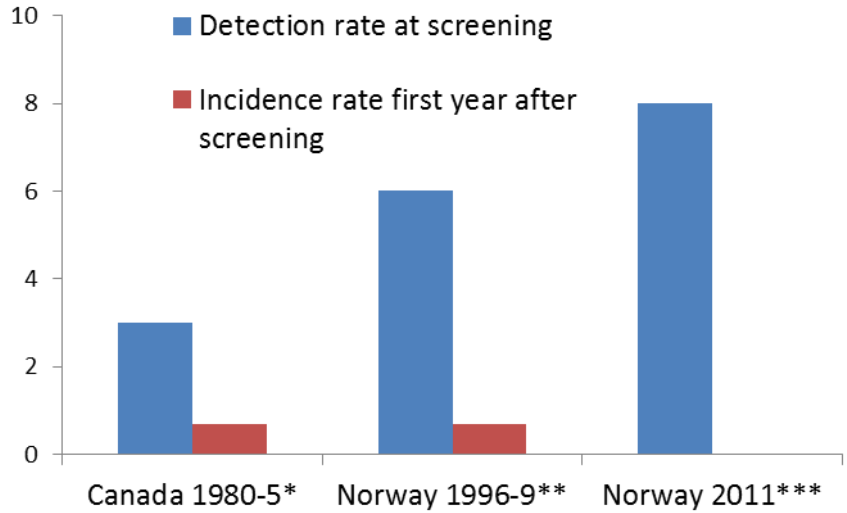
There is a disease reservoir of DCIS

Prevalence of DCIS by number of slides per breast

	Number of slides per breast	Prevalence of DCIS
Bartlow et al, 1987	9	0
Kramer et al, 1973	40	4.3
Nielsen et al, 1984	95	14.3
Nielsen et al, 1987	275	39

Detection rate at screening has increased dramatically

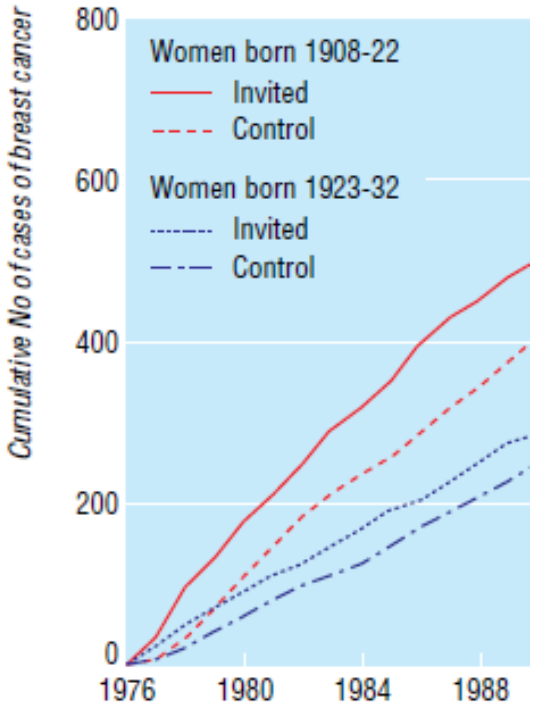
Detection rates have increased because of: double view, computer assisted reading, ultrasound, MR and 3D mammography without any decline in the rate of interval cancer.

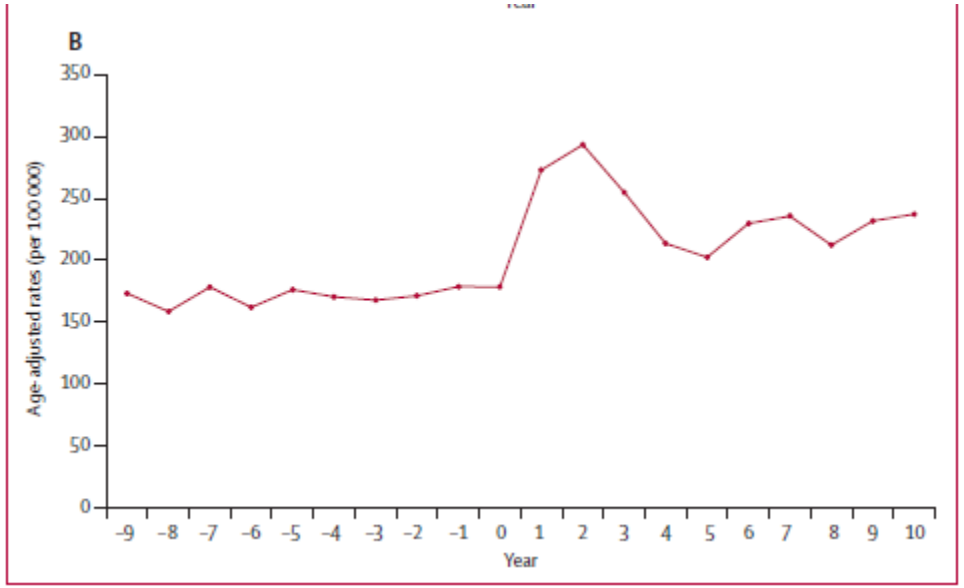


Incidence rate in the second year after a screening is 140 per 100,000 which is the incidence rate in mid 1980s.

* Miller et al (CBCSS-2), CMAJ 1992
** Evaluering av prøveprosjektet, 1999-2000
*** Skaane et al, Radiology 2013

3. Observed incidence when screening





Zachrisson et al, BMJ 2006.

Zahl, Gøtzsche & Mæhlen, Lancet Oncol 2011.

Incidence after screening

Ignored and not studied until Zahl, Strand & Mæhlen (BMJ 2004) reported that the fall after age 69 years accounted for only 1/20th of the observed incidence increase when screening women aged 50-69 years. We called the unaccounted increased for overdiagnosis.

The MISCAN model predicts

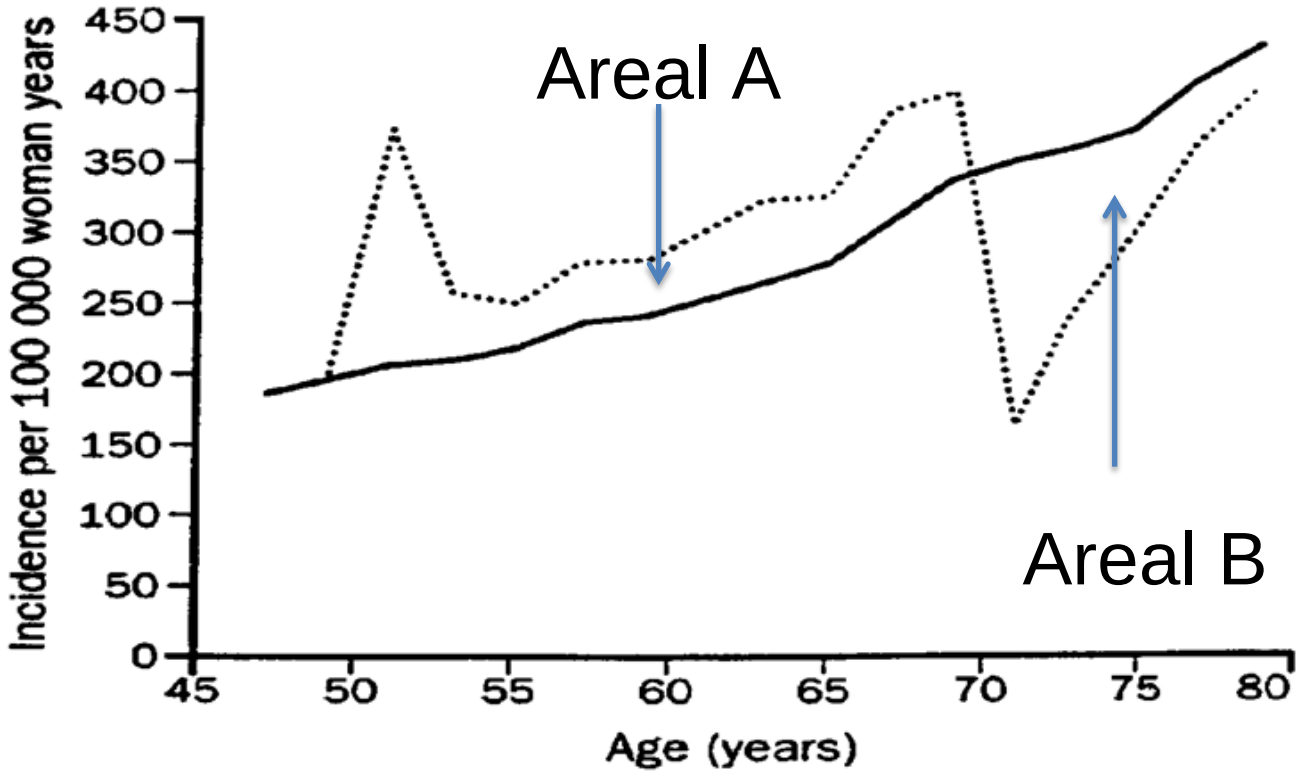
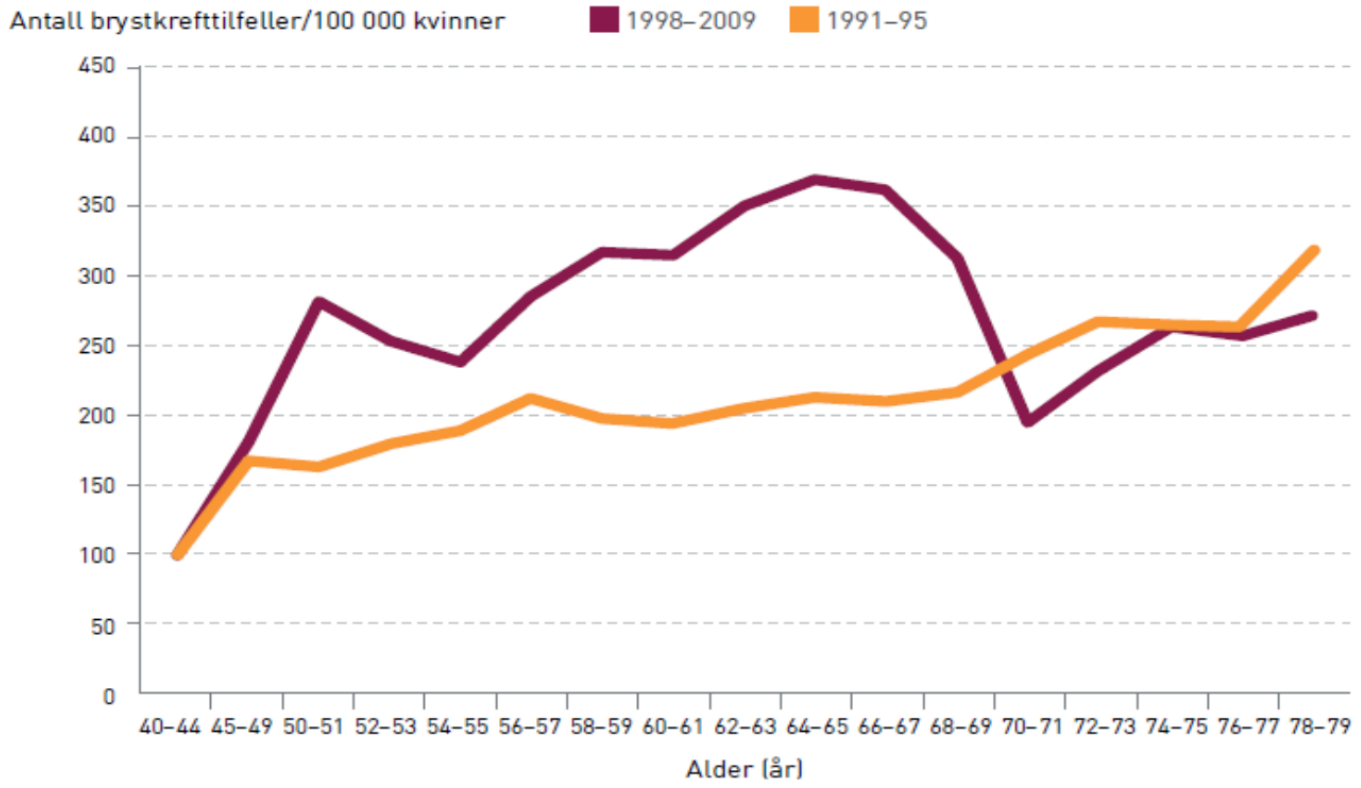


Figure: **Expected breast cancer incidence in 2-year age categories**

Solid line = not screened, dotted line = screened.

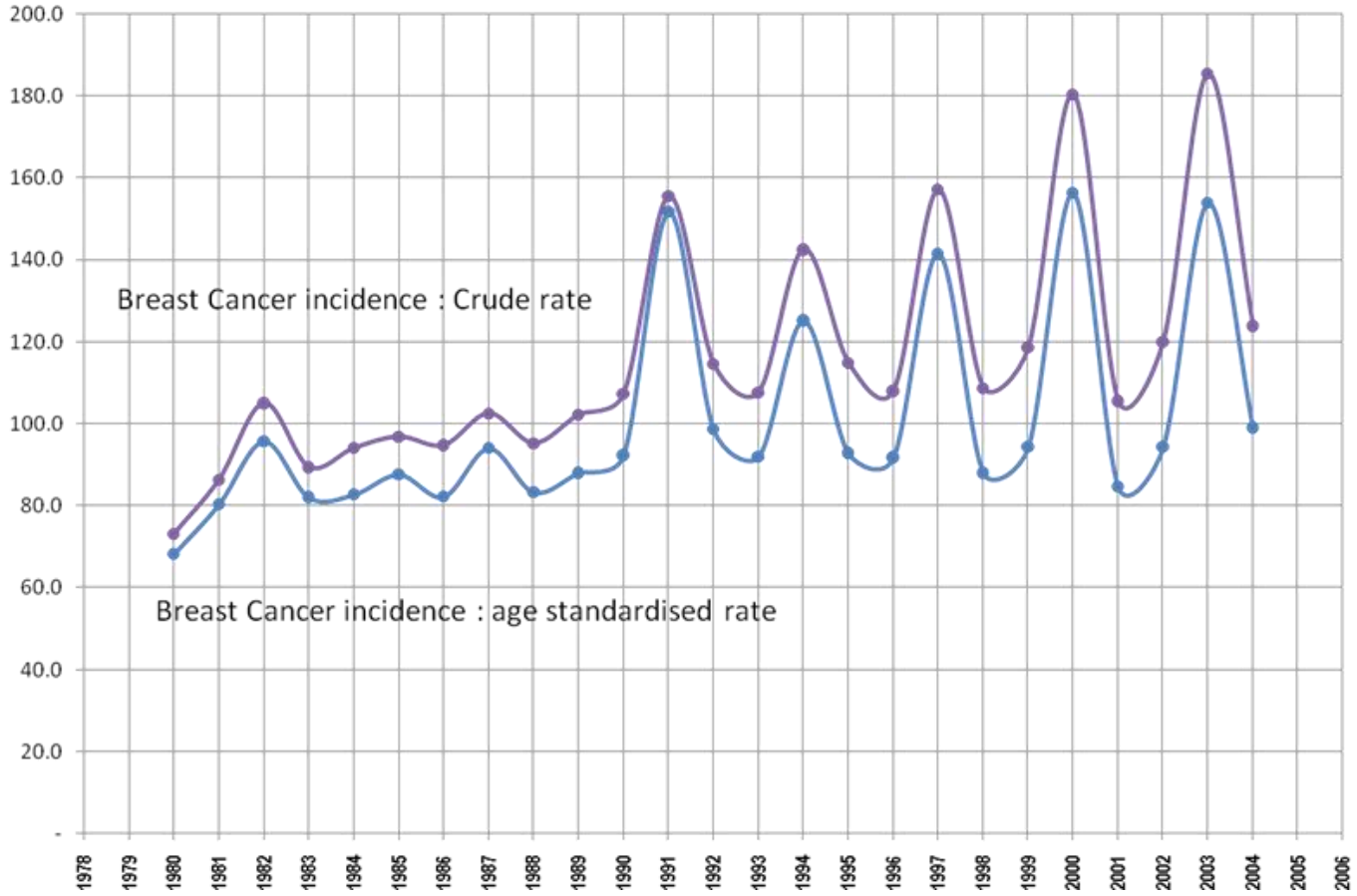
Boer et al, Lancet 1994

In reality: Areal A = 20 · Areal B



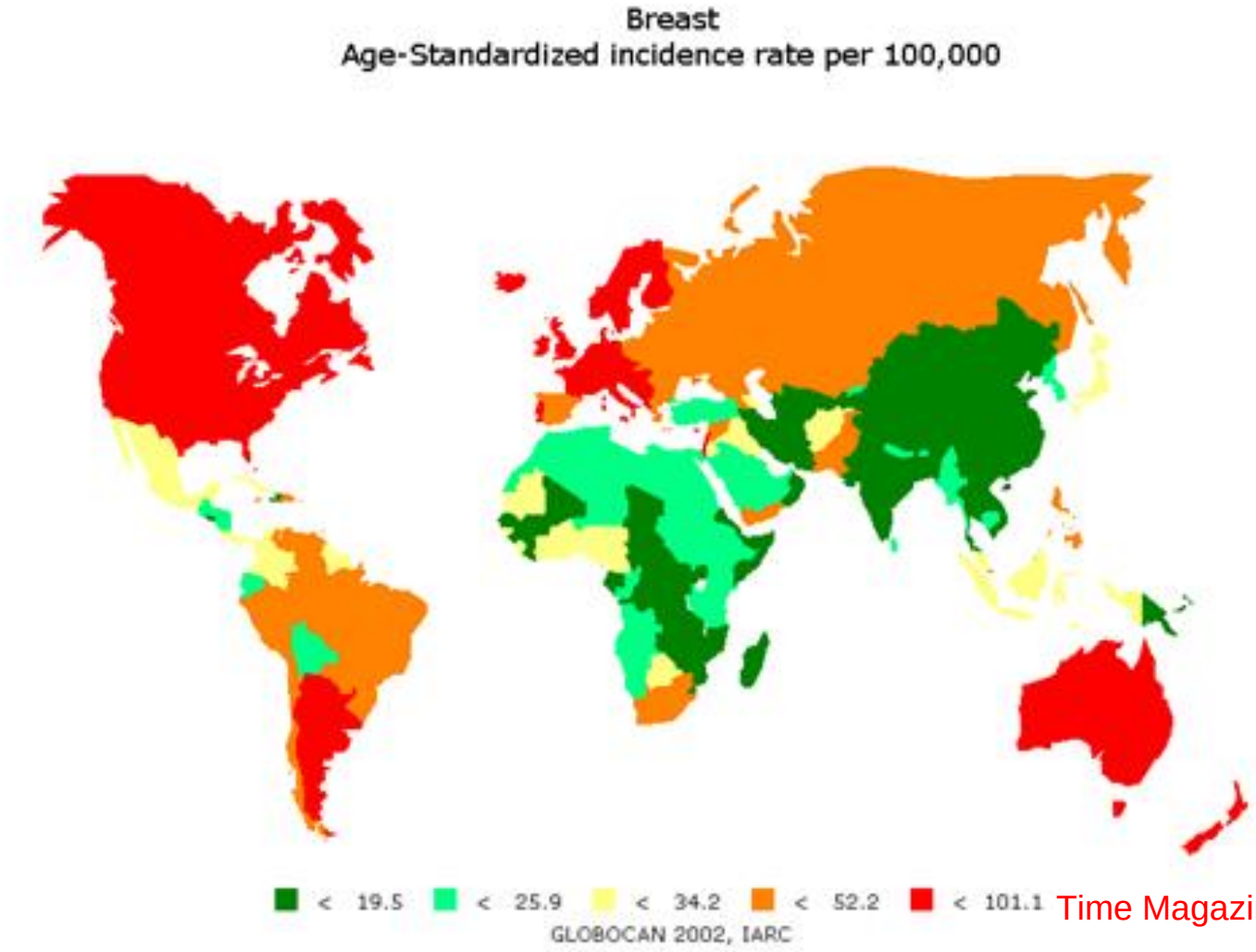
Figur 1 *Aldersspesifikk insidens av invasiv brystkreft i prøvefylkene i perioden 1998–2009 (rød linje) og i perioden 1991–95 (gul linje). Merk at de to første aldersgruppene er femårige (40–44 år og 45–49 år) og at aldersgruppene er toårige fra og med 50 år*

Invasive breast cancer incidence in Fife, Scotland



Vaidya. *BMJ* 2004; 339: b2587 (rapid responses)

This has caused a huge incidence increase in all countries with screening and only in the screened age group



Time Magazine, February 8, 2002

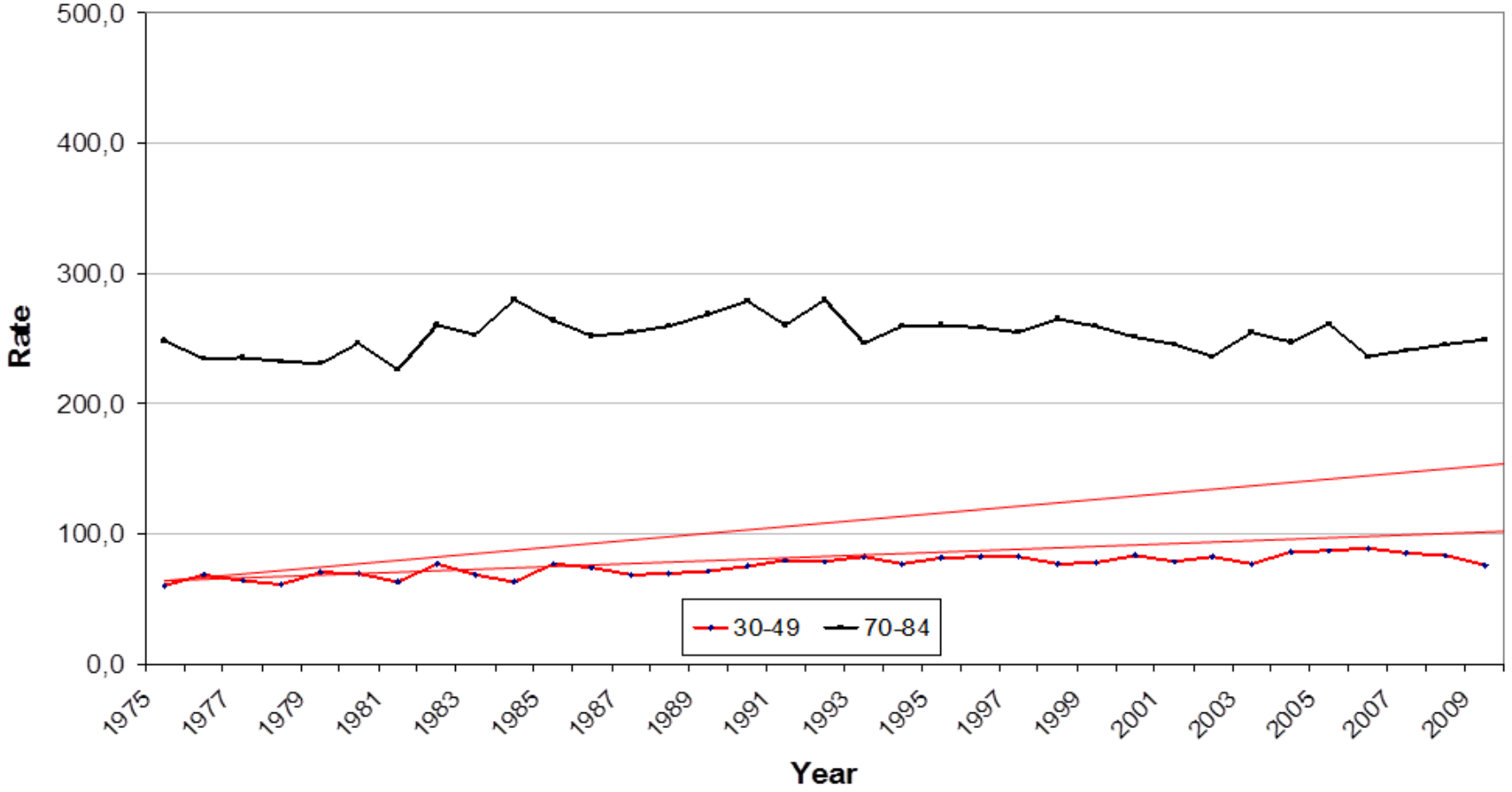
4. Alternative explanations of the incidence increase when screening

A. Underlying incidence increase

B. Increased use of hormone replacement therapy

C. Earlier diagnosis

A. Underlying incidence increase



B. Hormone replacement therapy

The randomized WHI study reported almost no effect of HT on the breast cancer incidence

Understanding recent trends in incidence of invasive breast cancer in Norway: age-period-cohort analysis based on registry data on mammography screening and hormone treatment use

BMJ 2012

OPEN ACCESS

Harald Weed
Steinar Tretli

Conclusions Changes in incidence trends of invasive breast cancer since the early 1990s may be fully attributed to mammography screening and hormone treatment, with about similar contributions of each factor.

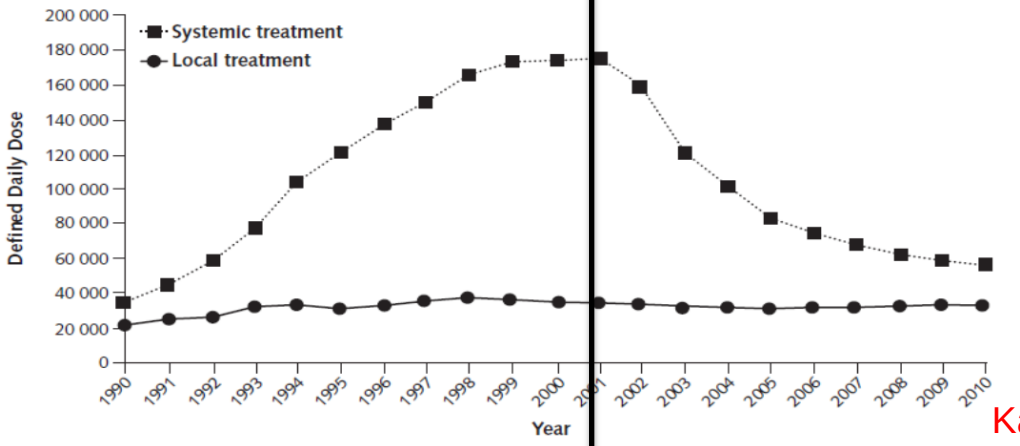
's J Vatten *professor*[3],

The model is controversial: "the authors do not discuss artefacts that can arise in ecological data and age-period-cohort analyses when non-linearities are present—problems that were noticed only after the method was introduced"

Michels (editorial) BMJ 2012

Before 2001: Breast cancers incidence increased from 180 to 290 per 100,000

Appendix Figure 1. **Defined daily doses for menopausal hormone therapy among all women in Norway from 1990 to 2010.**



Systemic treatment consisted of Anatomical Therapeutic Chemical group G03C-estrogens and G03F-combined estrogen and progesterone, and local treatment consisted of Anatomical Therapeutic Chemical group G03C. Adapted from the Norwegian Institute of Public Health.

Kalager et al, Ann Intern Med 2012.

After 2001:

Breast cancer rate > 300 per 100,000, while use of HT has dropped 80%

Zahl & Mæhlen, Tidsskr Nor Lægeforen 2012.

C. Can it be explained by earlier diagnosis?

Statisticians say

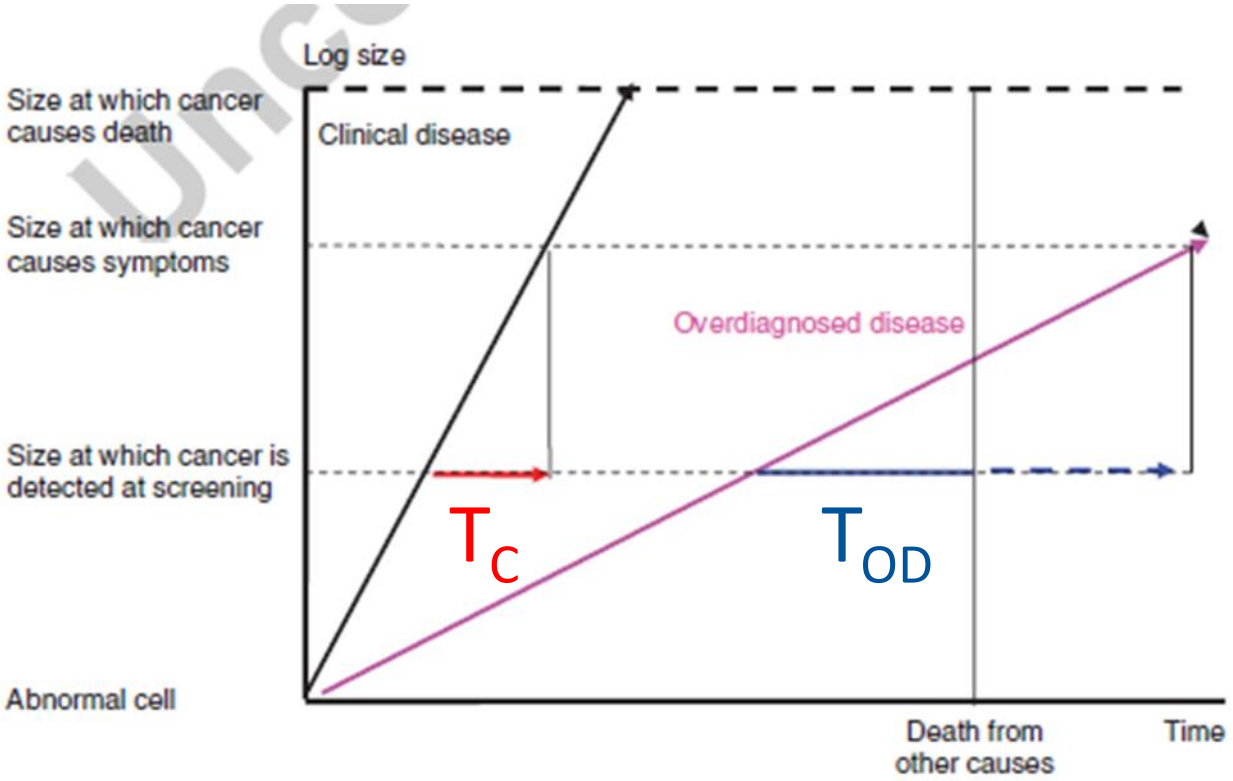
1. Lead time for breast cancer when screening with mammography is 2-7 years

Weedon-Fekjær et al, J Med Screen 2005.

2. Lead time for prostate cancer is 3-12 years when screening with PSA

Draisma, et al, JNCI 2009.

5. What is lead time?



Zahl et al, Br J Cancer 2013.

How to estimate clinical lead time: T_C

We call the annual per cent incidence reductions compared with the background incidence in a control group pr_1, pr_2, pr_3 and pr_4, respectively. It is largest in the first year after screening, that is, $pr_1 > pr_2 > pr_3 > pr_4$. The clinical lead time (in years) is then calculated approximately as the weighted average:

$$T_C = 0.5 \times pr_1 + 1.5 \times pr_2 + 2.5 \times pr_3 + 3.5 \times pr_4) / S,$$

where $S = (pr_1 + pr_2 + pr_3 + pr_4)$. Note that this estimate is not inflated by including overdiagnosed tumours and this is a novel method.

Estimated clinical lead time

Clinical lead time in the Norwegian Mammography Screening program is 1.06 year.

Zahl et al, Br J Cancer 2013.

A study of 448 women with breast cancer estimated that 90% of the doubling times were between 69 and 1622 days with a median of 260 days

Spratt, von Fournier, Spratt, Weber, Cancer 1993.

“Sensitivity” of our method

Including 5% tumours with 5 years clinical lead time increased the estimated clinical lead time by 0.14 years; assuming a 1% annual underlying incidence increase added only 0.01 year to the estimate; assuming a 50% higher incidence reduction after screening increased the estimate by 0.06 year, and combining all three extreme assumptions increased the estimate by **0.18** year.

Zahl et al, Br J Cancer 2013.

Model-based lead time: T_M

$$f(t) = \lambda \int_a^t k(t-u)x^m(u)\,du$$

Vito Volterra, 1913

This formula can be used to estimate lead time when

i) all tumors grow and

ii) there are no competing causes of deaths

Walter & Day, Biometrics 1984.

$$T_M = T_C \times (1\text{-}p) + T_O \times p$$

Zahl et al, Br J Cancer 2013.

The relation ship between clinical and model-based lead time

T_C = 1 year

Table 2. Calculated model-based lead time for a combination of clinical tumours (all with lead time of 1 year) and overdiagnosed tumours

	Overdiagnosis	Model-based lead time
Scenario 1 Lead time for overdiagnosed tumours is 10 years	10%	1.8
	30%	3.1
	50%	4.0
	70%	4.7
Scenario 2 Lead time for overdiagnosed tumours is 25 years	10%	3.2
	30%	6.5
	50%	9.0
	70%	10.9

The level of overdiagnosis varies from 10 to 70% and the lead times for overdiagnosed tumours are 10 and 25 years, respectively.

Zahl et al, Br J Cancer 2013.

6. What is overdiagnosis?

Overdiagnosis is the detection of a disease that in the absence of screening would not have been diagnosed within the patient's lifetime.

Three methods for «adjusting overdiagnosis» for lead time

1. Adjusting for earlier diagnosis of clinical cancers

Observed 530 more cancers than expected

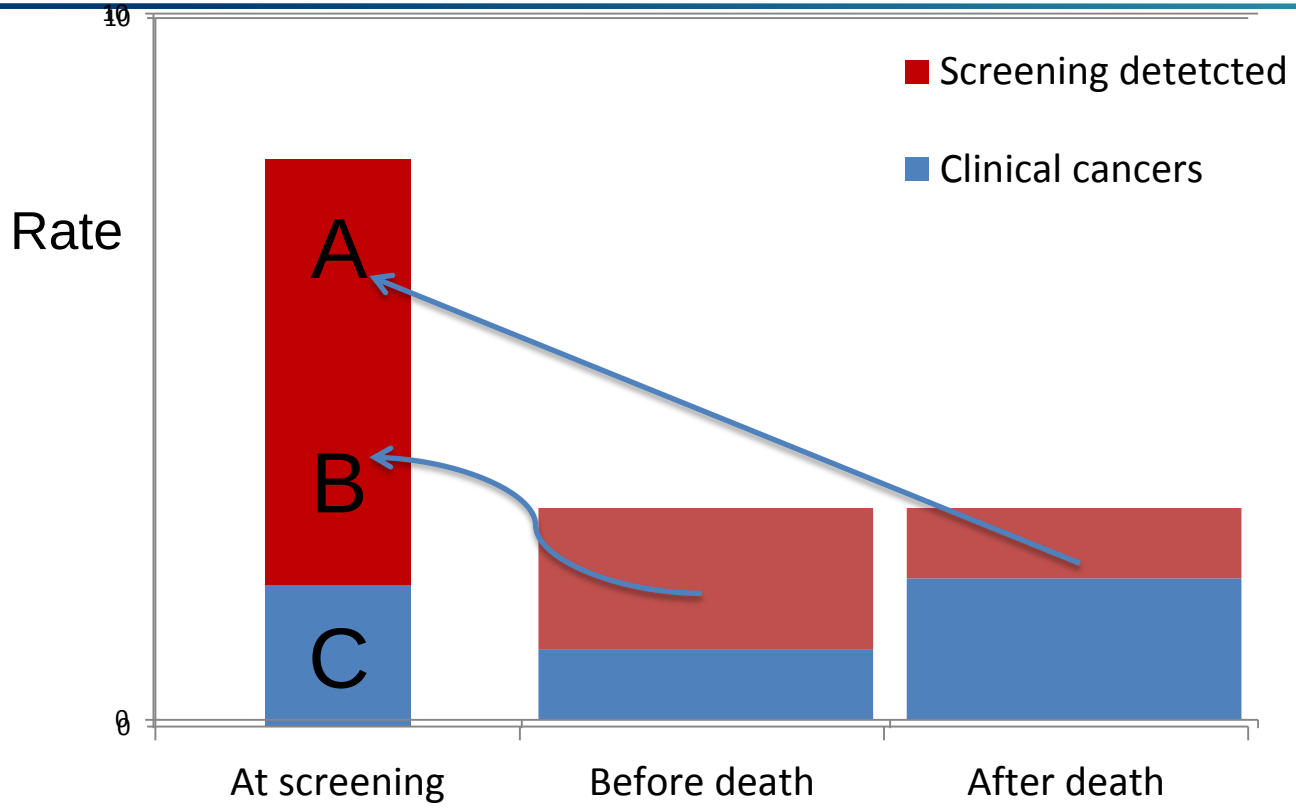
Observed 14 less after age 69

This is adjustment for earlier diagnosis

$$\frac{945+530-14}{945}=1.54$$

Zahl & Mæhlen, BMJ 2004, Tidsskr Nor Legeforen 2012

2. Calculating the proportions A and B based on a lead time model. Overdiagnosis: A/(A+B) or A/(A+B+C)



Note: Only data from the screening period are used in the calculation – they could have included data after screening has stopped

Ruth Etzioni and colleagues

Lead time when screening for prostate cancer with PSA varies from 3-12 years, and this corresponds to estimates of overdiagnosis varying from 23-42%.

Draisma et al, JNCI 2003, 2009.

If B = 0, then it is 100% overdiagnosis.

Stephen W Duffy

Overdiagnosis	Where?	Reference
At incidence screen, only 4% of DCIS	Two-County Trial and UK, Netherlands, Australia and the USA	Eur J Cancer 2003; 39:1746-54
zero (Markov models)	Two-County Trial	Radiol Clin North Am 2004; 42:793-806
around 5%	Firenze	J Med Screen 2004; 11:23-7
around 1%(*)	Two-County Trial and the Gothenburg Trial	Breast Cancer Res 2005; 7:258-65
4.8% (2 rounds plus following intervals)	Copenhagen	Breast J 2006; 12:338-42

Overdiagnosis in breast cancer screening: the importance of length of observation period and lead time

Stephen W Duffy* and Dharmishta Parmar

Breast Cancer RESEARCH

"The lead-time effect can be seen for age 50, year 1, for example, as

$410 + 0.86 \times 410 + 0.64 \times 420 + \cdots + 0.07 \times 561 = 1,812$."

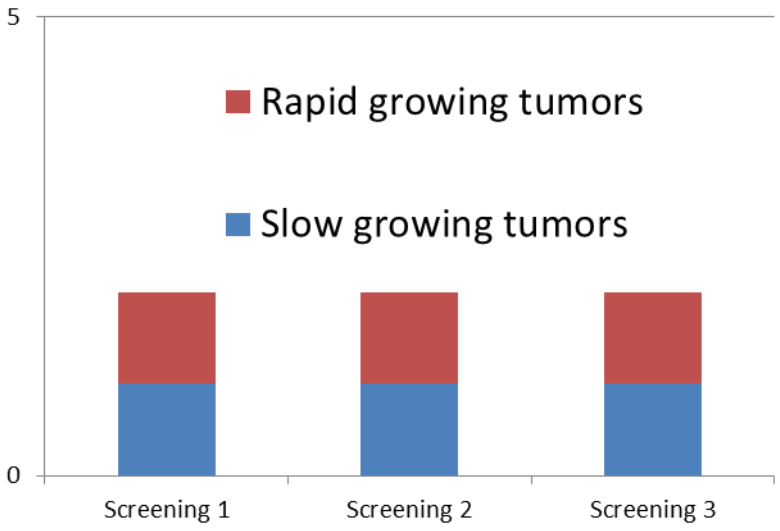
i.e. they assume that there is 86% and 64% reduction in the breast cancer rate in the first and second year after a screening. Truth is that it was about 0% in Fife, Scotland.

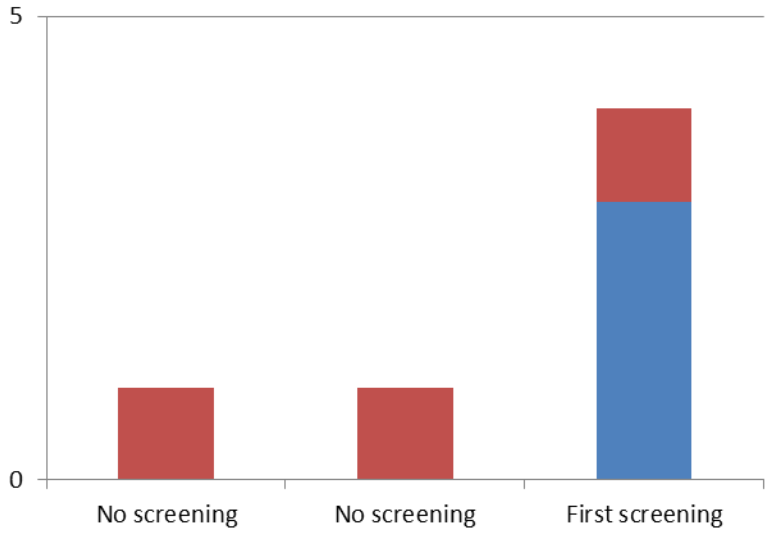
Model checking: Are there many overdiagnosed tumors with long T_{0D}?

After the screening period, you can actually check if there are many tumors with long lead time by

A) Study if there are any decline after screening has stopped like we do or Vaidya did in Fife, Scotland

B) Alternatively, you can study if slow-growing tumors accumulate over time



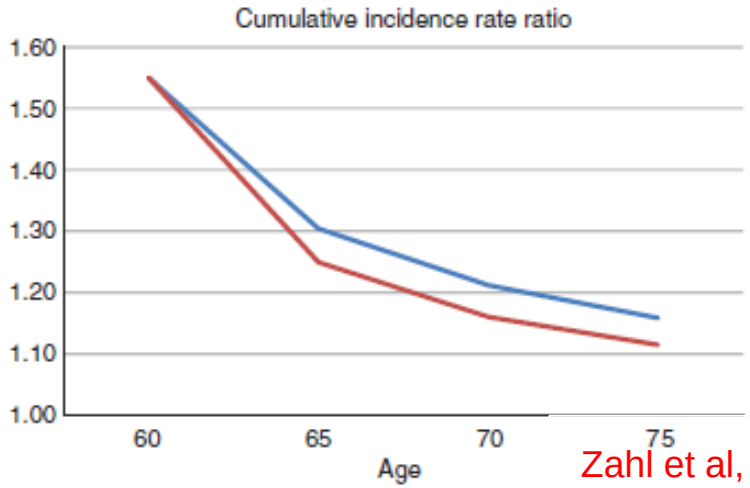


Zahl et al, Arch Int Med 2009
Zahl et al, Lancet Oncol 2011
Smith-Bindman et al, JAMA 2003

3. The «dilution method»

If you add cancers detected after screening has stopped at age 69 to both the screening and the control group, you will get a function that tend to the life time risk when all are dead

Example (Norway): 55 % incidence increase from age 50 to 59 years, whereof 5% of 55% is earlier diagnosis of clinical relevant cancers.



Zahl et al, Br J Cancer 2013.

Problems with the dilution method

- Estimates depend on i) how long you are screening and ii) how long you follow-up after screening has stopped.
- It is not the life time risk (unless you follow the cohorts until all are dead at age 100). Impractical method.
- "…the Panel thinks that the **best estimate** of overdiagnosis for a population invited to be screened is roughly 11%, defined as the excess incidence in the screening population as proportion of the long-term expected incidence."

The Marmot Report. Lancet 2012

- "11% of something that is not defined" cannot be the best estimate??

- If both groups are screened after age 59 (which is also adjusting for lead time – see the red curve), then you get almost identical estimates (curves that tend to zero), even though the proportion of overdiagnosed tumors in the population is constant over time.

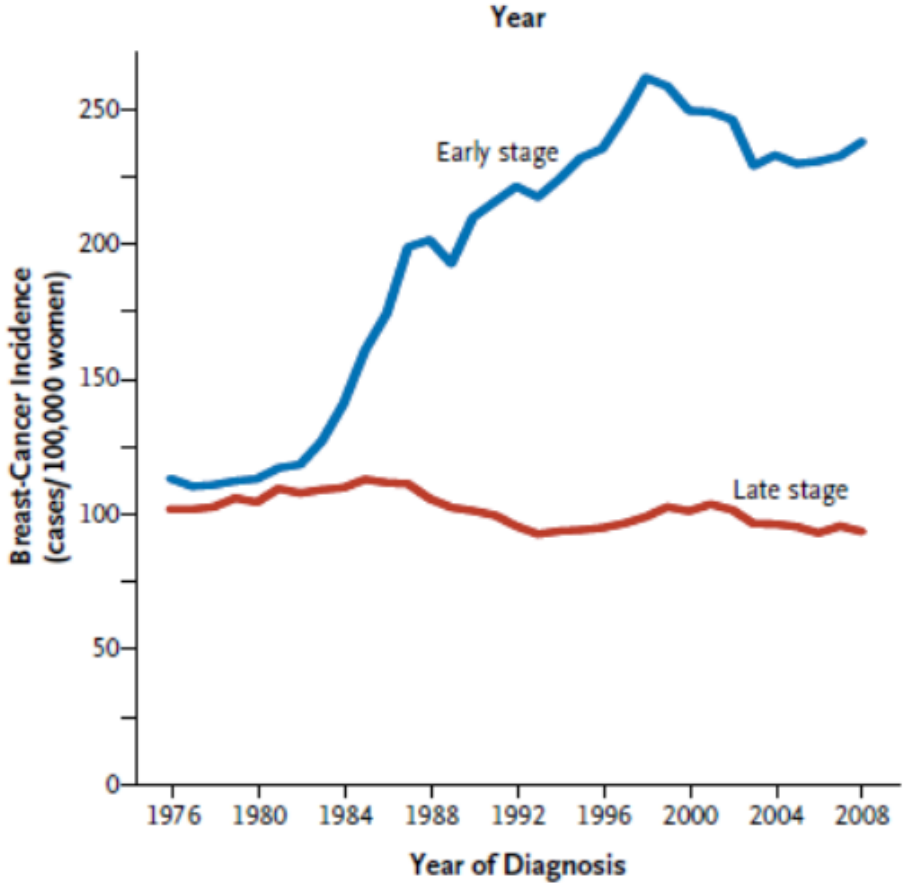
(Actually this is a test a test of cancer regression).

Conclusions

1. Model-based lead time has no medical interpretation (in contrast to lead time for clinical relevant tumors where it means earlier diagnosis)
2. Tumors with long lead times mainly exist in the head of some statisticians
3. Overdiagnosis adjusted for model based lead-time are not comparable from studies to studies
4. And cannot be compared with overdiagnosis adjusted for clinical relevant tumors = adjusting for earlier diagnosis

HIT

Not less late stage disease



Bleyer, Welch. *N Engl J Med* 2012