Overdiagnosis: 45 years and still in the making

From shadowy idea to acknowledged reality

Professor Alexandra Barratt
School of Public Health
Is it time to burn our bras again?
Part 1: Beginnings, 1968
PRINCIPLES AND PRACTICE OF SCREENING FOR DISEASE

J. M. G. WILSON
Principal Medical Officer, Ministry of Health, London, England

G. JUNGNER
Chief, Clinical Chemistry Department, Sahlgren's Hospital, Gothenburg, Sweden

WORLD HEALTH ORGANIZATION
GENEVA
1968
A BAS les ordonnances et le 1er plan
A BAS la réforme des Arts plastiques et l'ambition
A BAS DE GAULLE
response). The figures may therefore be to some extent biased. However, it seems unlikely that they are far wrong, since they are supported by data from the National Health Survey carried out by the National Center for Health Statistics of the United States Department of Health, Education, and Welfare. According to these findings, 15.5% of persons aged 18 to 79 had a blood-sugar level of 160 mg/100 ml, or over, one hour after taking 50 g of oral glucose (see Chapter 4, page 85). Whether some 13%-15% of the entire adult population of a country of this type is in need of treatment for diabetes, and, if so, which kind of treatment is required, urgently need discovering. Since the treatment of border-line diabetics (see page 83) is not at present known to affect prognosis it seems reasonable in case-finding to advise treatment only for declared, or established, diabetics until the outcome of surveys, at present in progress, is known. (The question of what should be regarded as declared, or established, diabetes is discussed on page 85.)
Part 2: Early warnings
Some Pitfalls in the Evaluation of Screening Programs

Munson Feinleib, MD, DrPH, Bethesda, Md, and Marvin Zelen, PhD, Amherst, NY

The interpretation of the results of any screening program is always hazardous. This is due in part to all of the "real life" problems of bias, chance variation, and small numbers which are inherent in most, if not all, clinical trials and epidemiologic studies. But it is also due to certain intrinsic theoretical difficulties in ascertaining a complex process from very limited data.

For example, it is well known that in a one-shot screening program increased survival time after diagnosis does not necessarily mean that early diagnosis was helpful. It might simply mean that the time of diagnosis has been advanced without necessarily meaning that the time of death has been delayed. In order to evaluate this aspect of the problem, two methods are available. First, we can try to estimate the "lead time," i.e., the time between diagnosis with the screening program and the usual time of diagnosis under current medical practice. The lead time is the time that diagnosis has been advanced. If this is known, it can be subtracted from the observed survival time, and the residual survival time in the cases diagnosed by screening can then be compared with usual survival times. Only if there is improvement in the residual survival time can any benefit be claimed.

A second approach is to ascertain the age-specific mortality from the disease in the entire screened population during the years following the screening program. If these age-specific rates are lower than those for an unscreened control population, then a benefit can be claimed.

To the best of our knowledge, only in the Memphis-Shelby County Cervical Cancer Study and in the ongoing Health Insurance Plan of Greater New York (HIP) Mammography Study have attempts been made to estimate the lead time gained by the screening program. Hutchinson and Shapiro have estimated that the lead time for breast cancer screening is between ten and 20 months. Only if the survival of the screened patients is increased at least by this amount will they have some justification in claiming that early diagnosis has improved survival. The HIP study is also the only one to our knowledge that has a randomized, unscreened control group with which to compare the subsequent age-specific mortality rates. It should be noted that the age-specific "discovery" rates will initially be higher in the screened population, as is shown in Table 1, which is based on data provided by S. Shapiro, BS, in October 1968. The reason for the higher discovery rates in the screened population is that the program is diagnosing the prevalence cases which have been in existence for varying durations, some of which may never go on to clinically recognizable disease, as well as diagnosing the recently incident cases.

Because the cases discovered at screening are prevalence cases which have been in existence for some length of time prior to the screening, they tend to have longer preclinical phases than the average case. This is shown in the Figure where the lines indicate pre-clinical disease and usual diagnosis.
Furthermore, long duration of preclinical disease may be correlated with long duration of the clinical phase. That is, the slow-growing tumor with favorable prognosis at time of usual diagnosis may also have an indolent preclinical phase. Thus, a screening program will not only tend to predate the onset of illness, but the cases discovered by one-shot screening may tend to be biased towards more benign disease. Even if the residual life-span after accounting for early diagnosis is increased, this may simply be due to a preponderance of benign disease among the cases discovered at screening.

Nevertheless, it is important to be able to estimate the lead time for those diseases in which early diagnosis may be beneficial.

Table 2.—Estimates of Mean Duration and Lead Time for Preclinical Disease

<table>
<thead>
<tr>
<th></th>
<th>Example Fig 1</th>
<th>Breast Cancer</th>
<th>Cervical Cancer</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P$</td>
<td>5</td>
<td>281.6</td>
<td>705.3</td>
</tr>
<tr>
<td>$I$</td>
<td>2</td>
<td>152.2</td>
<td>77.5</td>
</tr>
<tr>
<td>$d = \frac{P}{I}$</td>
<td>2.5</td>
<td>1.85</td>
<td>9.10</td>
</tr>
<tr>
<td>$A_I = x + 2.5$</td>
<td></td>
<td>54.47</td>
<td>56.77</td>
</tr>
<tr>
<td>$A_P = x + 1.7$</td>
<td></td>
<td>55.01</td>
<td>51.50</td>
</tr>
<tr>
<td>$\Delta A = A_I - A_P$</td>
<td>0.8</td>
<td>-0.54</td>
<td>5.27</td>
</tr>
<tr>
<td>Lead time $\bar{\tau} - \Delta A$</td>
<td>1.7 yr</td>
<td>2.39 yr</td>
<td>3.83 yr</td>
</tr>
</tbody>
</table>
Figure 2.—Age-specific incidence rates for ductal carcinoma in situ of the breast from 1973 to 1992 for US women, all races combined (adapted from National Cancer Institute SEER tapes).
Fig 1.—Annual breast cancer incidence (x) and mortality (o) for Connecticut, age-adjusted for 1950 standard population (unpublished data, Connecticut Tumor Registry 1973, 1974, 1975).
Part 3: Storms
Should I Continue Having Mammograms to Screen for Breast Cancer?

A decision aid for women aged 70 and older at their next screening mammogram.

AUSTRALIAN SCREENING MAMMOGRAPHY DECISION AID TRIAL
Cancers which may never affect your health

You may have noticed that more breast cancers are found in women who are screened than in women who are not screened. Some breast cancers found by screening never cause any problems because women die of something else first. These breast cancers may be slow growing cancers, for example DCIS (ductal carcinoma in situ).

So, if women with these slow growing cancers had not had screening, they might never have known they had cancer and would not have had treatment. The diagram below tries to illustrate this.

**Maria is 70 years old. She needs to decide whether to continue or stop having mammograms**

**If Maria stops, this may happen:**

- Age 70
- She actually has a breast cancer, but it causes no symptoms and because she stops having screening mammograms, it is never found
- Age 75: Dies of heart attack

**If Maria continues, this may happen:**

- Age 70
- Her breast cancer is found by screening. She has surgery (lumpectomy + radiotherapy) and tamoxifen. She has no symptoms of breast cancer and her breast cancer does not recur
- Age 75: Dies of heart attack
Should I Start Having Mammograms to Screen for Breast Cancer?

Some 40 year old women start thinking about whether they should attend mammography screening now or wait until they are 50. If you are in this situation, you might find this website helpful.

Researchers from the University of Sydney have compiled the best available evidence regarding mammography screening and created what we call a decision aid. A decision aid is intended to provide you with unbiased information so that you can make a decision after considering the evidence.

We invite you to spend approximately 30 minutes reading the decision aid and completing some questions. All of your responses remain confidential, and at no time do we ask your name or an email address.

If you are interested, click 'next' to find out more.
Cancer Screening - Benefits and Harms, Part One

Monday 22 August 2005 8:30AM

Most people would say that screening to find cancer early is a good idea. Well, maybe not so good. It could be that finding small cancers earlier and earlier is doing more harm than good. According to Professor Gilbert Welch from Dartmouth Medical School, New Hampshire, USA, cancer screening is a two-edged sword with important harms as well as benefits.

This is part one of a special three-part series which looks at how screening can unearth cancers you'd rather not know about, as well as other drawbacks of cancer screening.

This series is presented by Associate Professor Alex Barratt of Sydney University.
“The main balance I assume you are trying to achieve is ….. encouraging full participation of women especially aged 50-69 in mammographic breast screening and in cervical screening-- both are now well proven major success stories. I guess we all have responsibility to discuss the evidence and talk about that and encourage the population to be screened where benefit is clear.”

“…… Thus we have a real obligation to increase participation in these screening programs”

“……Obviously it is for the public good based on this sort of evidence to give a clear message of good from breast screening."
GE Gets FDA Approval for 3-D Mammogram Machine
Win Adds Substantial Competition in Market Dominated by Hologic

A 3-D view, left, compared with a 2-D view from GE's SenoClue mammogram machine showing microcalcifications in the breast, which can indicate cancer.

GE General Electric Co. (GE $GE) won approval from the Food and Drug Administration to offer its 3-D breast-imaging technology in the U.S., the company said, adding substantial new competition to the fast-growing market dominated by Hologic Inc. (HOLX $HOLX)

Three-dimensional imaging—also known as tomosynthesis—combines X-rays taken from multiple angles to produce a more detailed picture than regular mammograms. The technology costs more than regular mammography, but finds more cancers with fewer false alarms, studies have found, and is rapidly replacing the traditional mammography in the $10 billion-a-year market for breast screenings.

Insurers and Medicare don't provide extra reimbursement for 3-D mammograms. Some hospitals charge patients extra for it—generally $50 to $75 per exam; other hospitals absorb the extra cost of the machines.
Part 4: The problem with PINK
Our Feel-Good War on Breast Cancer

I used to believe that a mammogram saved my life. I even wrote that in the pages of this magazine. It was 1996, and I had just turned 35 when my doctor sent me for an initial screening — a relatively common practice at the time — that would serve as a base line when I began annual mammograms at 40. I had no family history of breast cancer, no particular risk factors for the disease.
The benefits and harms of breast cancer screening: an independent review

Independent UK Panel on Breast Cancer Screening

Summary

Whether breast cancer screening does more harm than good has been debated extensively. The main questions are how large the benefit of screening is in terms of reduced breast cancer mortality and how substantial the harm is in terms of overdiagnosis, which is defined as cancers detected at screening that would not have otherwise become clinically apparent in the woman’s lifetime. An independent Panel was convened to reach conclusions about the benefits and harms of breast screening on the basis of a review of published work and oral and written evidence presented by experts in the subject. To provide estimates of the level of benefits and harms, the Panel relied mainly on findings from randomised trials of breast cancer screening that compared women invited to screening with controls not invited, but also reviewed evidence from observational studies. The Panel focused on the UK setting, where women aged 50–70 years are invited to screening every 3 years. In this Review, we provide a summary of the full report on the Panel’s findings and conclusions. In a meta-analysis of 11 randomised trials, the relative risk of breast cancer mortality for women invited to screening compared with controls was 0.80 (95% CI 0.73–0.89), which is a relative risk reduction of 20%. The Panel considered the internal biases in the trials and whether those trials, which were done a long time ago, were still relevant; they concluded that 20% was still a reasonable estimate of the relative risk reduction. The more reliable and recent observational studies generally produced larger estimates of benefit, but these studies might be biased. The best estimates of overdiagnosis are from three trials in which women in the control group were not invited to be screened at the end of the active trial period. In a meta-analysis, estimates of the excess incidence were 11% (95% CI 9–12%) when expressed as a proportion of cancers diagnosed in the invited group in the long term, and 19% (15–23%) when expressed as a proportion of the cancers diagnosed during the active screening period. Results from observational studies support the occurrence of overdiagnosis, but estimates of its magnitude are unreliable. The Panel concludes that screening reduces breast cancer mortality but that some overdiagnosis occurs. Since the estimates provided are from studies with many limitations and whose relevance to present-day screening programmes can be questioned, they have substantial uncertainty and should be regarded only as an approximate guide. If these figures are used directly, for every 10 000 UK women aged 50 years invited to screening for the next 20 years, 43 deaths from breast cancer would be prevented and 129 cases of breast cancer, invasive and non-invasive, would be overdiagnosed; that is one breast cancer death prevented for about every three overdiagnosed cases identified and treated. Of the roughly 307 000 women aged 50–52 years who are invited to begin screening every year, just over 1% would have an overdiagnosed cancer in the next 20 years. Evidence from a focus group organised by
Perspective

Abolishing Mammography Screening Programs? A View from the Swiss Medical Board

Nikola Biller-Andorno, M.D., Ph.D., and Peter Juni, M.D.

In January 2013, the Swiss Medical Board, an independent health technology assessment initiative under the auspices of the Conference of Health Ministers of the Swiss Cantons, the Swiss Medical Association, and the Swiss Academy of Medical Sciences, was mandated to prepare a review of mammography screening. The two of us, a medical ethicist and a clinical epidemiologist, were members of the expert panel that appraised the evidence and its implications. The other members were a clinical pharmacologist, an oncologic surgeon, a nurse scientist, a lawyer, and a health economist. As we embarked on the project, we were aware of the controversies that have surrounded mammography screening for the past 10 to 15 years. When we reviewed the available evidence and contemplated its implications in detail, however, we became increasingly concerned.

First, we noticed that the ongoing debate was based on a series of reanalyses of the same, predominantly outdated trials. The first trial started more than 50 years ago in New York City and the last trial in 1991 in the United Kingdom.1 None of these trials were initiated in the era of modern breast-cancer treatment, which has dramatically improved the survival rates of breast-cancer patients. Therefore, it is crucial to reconsider the evidence available at the time the trials were performed. We also noted that all the trials were performed in populations that are not representational of the current patient population. The current focus on screening younger women to detect breast-cancer, and in particular screenees under age 50, has also been neglected.

In this Perspective, we present our concerns about the evidence for mammography screening and our recommendations for future research.
Breast cancer screening: It’s your choice

New information to help women aged about 50 to make a decision
A Phase III trial of surgery versus active monitoring for LOw RISk Ductal Carcinoma in situ (LORIS)

Chief Investigator: Miss Adele Francis

Quality of Life Principal Investigators: Professor Lesley Fallowfield, Dr Val Jenkins

Trial Co-ordinator (quality of life): Lucy Matthews

Funded by: National Institute for Health Research Health Technology Assessment Programme (NIHR HTA)

The introduction of mammographic breast screening in the UK has resulted in a dramatic increase in the diagnoses of Ductal Carcinoma in situ (DCIS). This is because DCIS is usually found on mammograms. A diagnosis of DCIS means there are abnormal cells in the milk ducts of the breast that have not invaded the surrounding breast tissue.

DCIS is visualised as specks of white (calcium) seen on mammograms. Pathologically it can be divided into 3 grades – high, intermediate or low. High grade DCIS is more likely to turn into breast cancer, and so it is treated as though it is breast cancer. Low and low/intermediate grade DCIS is different and doctors are uncertain if it would ever become invasive breast cancer.

Only patients with low risk DCIS (i.e. low or intermediate grade DCIS) are eligible to join the LORIS trial.

In the LORIS trial patients will be randomised to either a no surgical intervention (active monitoring) or standard treatment which is surgery. DCIS is a difficult disease to describe and there is conflicting information available to patients on charity websites. A patient friendly DVD will be provided for potential LORIS trial participants, to help provide a concise explanation of DCIS and the trial.

The quality of life aspect is integral to this study, due to the fact that a diagnosis of DCIS can provoke significant psychological distress. Patients will be asked to complete questionnaires at different time points that will look at factors that may influence:
Part 5: Conclusions
1. Persistence and courage
2. Research
3. Engagement and communication
4. Advocacy
5. Collaboration
A time to burn our bras again?
My thanks to Ray Moynihan, Les Irwig and Jonathan Bogais for critical review.