BMJ RESEARCH

Overdiagnosis in publicly organised mammography screening programmes: systematic review of incidence trends

Karsten Juhl J0rgensen. researcher Peter C G0tzsche. director

ABSTRACT

Objective To estimate the extent of over diagnosis (the detection of cancers that will not cause death or symptoms) in publicly organized screening programmes. Design Systematic review of published trends in incidence of breast cancer before and after the introduction of mammography screening.

Data sources PubMed (April 2007), reference lists, and authors.

Review methods One author extracted data on incidence of breast cancer Oncludingcarcinoma in situ), population size, screening uptake, time periods, and age groups, which were checked independently by the other author. Linear regression was used to estimate trends in incidence before and after the introduction of screening and in older, previously screened women. Meta-analysis was used to estimate the extent of over diagnosis.

Results incidence data covering at least seven years before screening and seven years after screening had been fully implemented, and including both screened and non-screened age groups, were available from the United Kingdom; Manitoba, Canada; New South Wales, Australia; Sweden; and parts of Norway. The implementation phase with its prevalence peak was excluded and adjustment made for changing background incidence and compensatory drops in incidence among older, previously screened women. Over diagnosis was estimated at 52% (95% confidence interval 46% to 58%). Data from three countries showed a drop in incidence as the women exceeded the age limit for screening but the reduction was small and the estimate of over diagnosis was compensated in this review.

Conclusions: The increase in incidence of breast cancer was closely related to the introduction of screening and little of this increase was compensated for by a drop in incidence of breast cancer in previously screened women. One in three breast cancers detected in a population offered organized screening is over diagnosed.

INTRODUCTION

Screening for cancer may lead to earlier detection of lethal cancers but also detects harmless ones that will not cause death or symptoms. The detection of such cancers, which would not have been identified clinically in someone's remaining lifetime, is called over diagnosis and can only be harmful to those who experience it as it is not possible to distinguish between lethal and harmless cancers, all detected cancers are treated. Over diagnosis and overtreatment are there f or inevitable. [2]

It is well known that many cases of carcinoma in situ in the breasts do not develop into potentially lethal invasive disease. [1] In contrast, many find it difficult to accept that screening for breast cancer also leads to over diagnosis of invasive cancer. Hannless invasive cancer is common, however, even for lung cancer, with 30% over diagnosis after long term follow-up of patients screened by radiography. Autopsy studies have shown that invasive prostate cancer occurs in about 600/o of men in their 60s, whereas the lifetime risk of dying from such cancer is only

about 3%. [2] Autopsy studies have also found inconsequential breast cancer lesions. Thirty seven percent of women aged 40 – 54 who died from causes other than breast cancer had lesions of invasive or non-invasive cancer at autopsy, and half were visible on radiography. [3**]

Over diagnosis can be measured precisely in a randomized trial with lifelong follow-up if people are assigned to a screening or control group for as long as screening would be offered in practice, which in most countries is 20 years. Over diagnosis would be the difference in number of cancers detected during the lifetime of the two groups, provided the control group or age groups not targeted are not screened. In the absence of over diagnosis the initial increase in cancers in the screened age groups would be fully compensated for by a similar decrease in cancers among older age groups no longer offered screening, as these cancers would already have been detected.

The extent of over diagnosis and over treatment as a result of mammography screening was first quantified in reviews of randomized trials. 56 The total number of mastectomies and lumpectomies increased by 31% and mastemomies by 20%. As these trials did not have lifelong follow-up the extent of over diagnosis could have been overestimated. Underestimation is also possible, however, as the randomized design was maintained for only 4 – 9 years [6] and as opportunistic screening occurred in the control groups. [7]

Screening programmes differ from randomized trials. Radiologists outside arigorous trial setting may be less well trained that those in the trial, and technical developments resulting in higher resolution images may also affect outcomes. The basic prentiae of an unchanged lifetime risk of breast cancer in the absence of over diagnosis is, however, the same.

To estimate the extent of over diagnosis in organized screening programmes we compared trends in breast cancer incidence before and after screening, taking account of changes in the background incidence and any compensatory drop in incidence of breast cancer among older, previously screened women. We combined our results in a meta-analysis.

Methods

We included article in any language with data on breast cancer incidence for both screened and older, non-screened age groups for at least seven years before screening and seven years after screening had been fully implemented, regardless of the time it took to implement screening. We reasoned that a long period after implementation was necessary to obtain an estimate of the trend in breast cancer incidence that was unaffected by initial peak in prevalence when screening is introduced. Acquiring incidence data for age groups older than those screened allowed us to evaluate any compensatory declines in incidence among previously screened women.

When a country was described in several papers we selected the one with the most recent and best reported data as our core article, and we supplemented with other papers when relevant. When possible we also added data from the internet and supplied by authors. We did not search for articles published before 1990, as insufficient time would have elapsed after the initiation of screening.

Literature searches

Our searches in PubMed were developed iteratively and we tried several search strings. The final search, which identified all included articles, was ((Mammography" [MeSH]) AND ((Breast Neoplasms/epidemiology"[MeSH]) OR("Breast Neoplasms" [MeSH] AND incidence* [ti]))) OR (Breast cancer AND screening AND trend*[ti]) OR (Breast cancer AND screening AND overdiagnosis*[ti]).

One author scanned titles and abstracts and retrieved the full text of potentially relevant articles for evaluation of eligibility, scanned the reference lists, and contacted authors. We compared the final search with an archive of all articles on breast cancer screening published in 2004, which we have used for another study, [8] and found that we had not missed any potentially relevant papers. None of the four authors we contacted told us of additional studies but three provided unpublished data or referred us to internet resources. [9] We did not find additional studies in the reference lists.

Go to www.MindfulWellnessMedicalThermography.com/research to read more...