Reviewer's report

Title: Prognosis of screen-detected breast cancers: results of a population based study

Version: 2 Date: 25 August 2005

Reviewer: Adri C Voogd

Reviewer's report:

General
Although the paper certainly gives a good overview of the state of the art of breast cancer screening in the Modena area it does not contribute to our knowledge regarding the effectiveness of breast cancer screening. Of course a cancer registry is a useful tool to evaluate some of the surrogate markers for the effectiveness of breast cancer screening, such as detection rate and stage of disease at diagnosis. When these parameters are favourable, it is likely that breast cancer mortality rates will start to decline. Mortality rates, analysed at a population-based level, are the second best way (after randomised controlled trials) to prove the effectiveness (see also the paper of Otto et al., Lancet 2003; 361:1411-7). However, to conclude (as the authors do) that improved short-term breast cancer survival rates are a proof or even an indication of the success of a breast cancer screening programme is not justified from a methodological point of view.

Major Compulsory Revisions (that the author must respond to before a decision on publication can be reached)

By just comparing screen-detected cases with non-screen detected cases, occurring in patients of 50-69 years of age in the period 1995-2000, there is a large risk of introducing different kinds of biases. First of all it is not clear if the non-screen detected cases represent the patients that did not participate in the screening programme as well as the ones that have not been invited yet. I think this is the case, as it is not very likely that all women of 50-59 years in the Modena area have been screened at once. For a better interpretation of the results, a distinction has to be made between these two groups in the analyses. Second, more information is needed about the actual screening round during which the tumour was diagnosed. From other studies (for example Hakama et al. Lancet 1995;345:221-4), we know that there is a large difference in the aggressiveness of prevalent breast cancer cases (i.e., diagnosed during the first screening round) and incident cases (i.e., the ones diagnosed in subsequent round). The low proportion of interval cancers suggests that many women have been invited only for the first round yet. The patients diagnosed in the first round represent a special category, with slow growing tumour and a long preclinical phase, making lead-time bias very likely. Therefore, the authors should indicate which part of the screen-detected cases was diagnosed in the first screening round and which part in the second or third round. Lead-time bias remains the most serious form of bias undermining their results, even though the authors argue that this is not a problem because the median age at diagnosis did not differ between the screen-detected and non-screen-detected cases of the patients diagnosed in the pre-screening area. However, knowing that sensitivity of mammography is increasing with the age of the screened population, one would expect the screen-detected cases to be older than the ones not detected by screening, which surprisingly was not the case. Thus, analysing median age at diagnosis is not a good method to rule out lead-time bias.

Other comments:
- The introduction is too long-winded and should justify why the authors think that analysing the short-term outcomes of any scientific merit.
- Could the authors provide any information on the detection rate of the programme and the risk of a
false-positive outcome? The low proportion suggests a high detection rate, and possibly also a high rate of false-positive mammographic tests.
- No information is given on the number of patients that were lost to follow-up. Does this mean that follow-up was 100% complete?
- The fact that the screening programme is free of charge does not rule out “class bias”, i.e. the selective attendance of the higher social classes to the programme.

What next?: Reject because scientifically unsound

Level of interest: An article of limited interest

Quality of written English: Needs some language corrections before being published

Statistical review: No

Declaration of competing interests:
I declare that I have no competing interests.