



Evaluating new screening tests for breast cancer

Les Irwig, Nehmat Houssami, Bruce Armstrong and Paul Glasziou

BMJ 2006;332:678-679
doi:10.1136/bmj.332.7543.678

Updated information and services can be found at:
<http://bmj.com/cgi/content/full/332/7543/678>

These include:

References

This article cites 13 articles, 8 of which can be accessed free at:
<http://bmj.com/cgi/content/full/332/7543/678#BIBL>

5 online articles that cite this article can be accessed at:
<http://bmj.com/cgi/content/full/332/7543/678#otherarticles>

Rapid responses

You can respond to this article at:
<http://bmj.com/cgi/eletter-submit/332/7543/678>

Email alerting service

Receive free email alerts when new articles cite this article - sign up in the box at the top left of the article

Topic collections

Articles on similar topics can be found in the following collections

[Clinical trials \(epidemiology\)](#) (3850 articles)
[Epidemiologic studies](#) (6899 articles)
[Breast cancer](#) (1460 articles)
[Screening \(oncology\)](#) (1394 articles)
[Radiology](#) (3047 articles)
[Clinical diagnostic tests](#) (3541 articles)
[Radiology \(diagnostics\)](#) (2051 articles)
[Screening \(epidemiology\)](#) (2614 articles)
[Screening \(public health\)](#) (2616 articles)

Notes

To order reprints follow the "Request Permissions" link in the navigation box

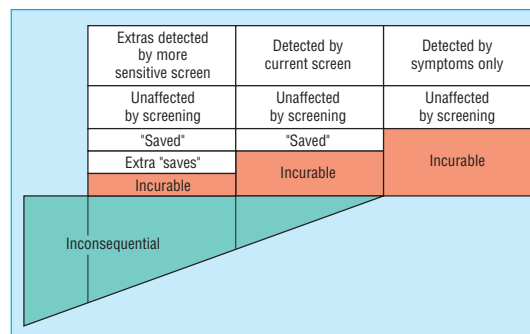
To subscribe to *BMJ* go to:
<http://resources.bmj.com/bmj/subscribers>

Evaluating new screening tests for breast cancer

May require randomised controlled trials to assess overdetection

The use of magnetic resonance imaging to screen women with high risk mutations in the genes associated with breast cancer has raised debate on what constitutes sufficient evidence for the efficacy of new screening tests.¹⁻⁴ The gold standard is evidence from randomised trials that early detection reduces mortality, as is the case for mammography and breast cancer,⁵ but how should we evaluate new tests that might detect cancer earlier?

Showing that a new test is more sensitive than others suggests that it has promise as a possible screening test,⁶ but detecting more apparent cases does not necessarily mean that using the test routinely will lead to a further reduction in breast cancer deaths. To fulfil the criteria for an effective screening test, the additional cancers detected must include ones that would both progress during the patient's lifetime and be curable by earlier treatment. The extra cancers picked up by new tests may count as cancers histopathologically but might not progress to cause symptoms in the women's lifetime: thus, a new test might lead to more overdetection rather than improved outcomes (figure). Overdetection of ductal carcinoma in situ (DCIS) is well documented,⁷ but it may also occur with cancers that seem, histologically, "invasive."⁸⁻⁹ Overdetection may cause harm through unnecessary labelling and treatment of patients as having a cancer that, without screening, might never have been diagnosed.



Balance between deaths averted and overdetection by more sensitive test

Overdetection can be identified best in a randomised controlled trial. Screening for several years should yield a higher average incidence of cancer in the screened group than in an unscreened control group during the years of screening. Once screening stops, the annual incidence of cancer in the screened group should drop below that in the unscreened group, and the eventual total number of cancers detected in the groups should equalise.¹⁰ A persisting excess of cancers in the screened group represents overdetection, as shown in the Malmö mammographic screening trial, for which the estimate of overdetection in women aged 55-69 at randomisation, followed for 15 years after the end of the trial, was 10% for all breast cancers and 7% for invasive breast cancers.¹¹

Research p 689 and
also Letters p 727

BMJ 2006;332:678-9

Methods of evaluating new and existing technology in cancer screening

Measures compared between tests	Randomised controlled trial of mortality benefit	Randomised controlled trial of interval cancer	Comparison of test accuracy (cross sectional)
Sensitivity	Yes	Yes	Yes
Stage (size and nodal status)	Yes	Yes	Yes
Interval cancer rate	Yes	Yes	No
Mortality	Yes	Not directly	Not directly
Duration of follow-up	Decades (10-20 years)	2-3 years	None
Sample size requirement	Large	Large	Small
Avoids overestimate of benefit because of overdiagnosis	Yes	Yes	No

There are three main research designs for evaluating new screening tests for cancer (table). Firstly, randomised controlled trials with long term follow-up provide the best evidence for comparing mortality from cancer among patients having different screening tests. They are unwarranted, however, once there is evidence that early detection confers benefit. Secondly, randomised controlled trials with short term follow-up can be used to compare interval cancer rates, the proportion of women whose screening yields negative results but who then present with cancer before the next scheduled screening test.¹² Reducing the rate of interval cancer rates is crucial, representing the potential benefit of early detection rather than overdiagnosis. To assess the impact of early detection further, such studies can also compare the rates of advanced cancers detected by subsequent screening rounds.

Thirdly, cross sectional studies can be used to compare the sensitivities of different tests by comparing cancer detection rates in people randomised to one or other test, or by paired studies in which people have both tests. In paired studies, sensitivities are estimated as the number of cancers each test detects divided by the total number of cancers detected by either test. Assessing relative sensitivities is valid, despite the fact that cross sectional studies do not provide data on follow-up, because missed cancers will be common to both tests.¹³ But, contrary to others' suggestions,¹ the interval cancer rate for each screening test cannot be obtained from such paired studies, even if followed over time, because cancers detected by either test will be treated: the only interval cancers are those which neither test detects.

A new screening test will seem more sensitive than an older one when it detects a larger proportion of the cancers that are collectively detected using both tests. At one extreme this apparently greater sensitivity could be due simply to overdiagnosis. At the other extreme, all the extra cancers detected might have progressed if undetected and their earlier detection should lower mortality. The increase in cancers detected by a new screening test may reflect the fact that the test detects cancers which are smaller or of a lower grade of malignancy. Without comparing interval cancer rates, the extent to which such a shift in grade reflects clinically relevant earlier detection or overdiagnosis will remain unclear.

Studies comparing new breast screening tests with mammography²⁻⁶ tend to use paired cross sectional designs. This may be adequate if the tests being compared are similar (for example, digital mammography versus film mammography¹⁴) or if the aim of the comparison is to establish equivalence. More sensitive screening tests which differ substantially from the comparator tests should be evaluated in randomised studies with short term follow-up over two or three years. Randomised controlled trials of the accuracy of alternative screening tests, such as those evaluating different tests for faecal occult blood in a programme for bowel cancer screening, are also practical and realistic.¹⁵

Critics may consider trials to detect interval cancer rates unnecessary or even unethical in people who are at substantially increased risk of developing cancer—for example, women at high risk of breast cancer because of gene mutations. But rigorous scientific evaluation is both ethical and essential to establish that a test does more good than harm, whether for the general population of women or for those with a greater risk of breast cancer.

Les Irwig *professor of epidemiology*

(lesi@health.usyd.edu.au)

Nehmat Houssami *senior lecturer*

Bruce Armstrong *professor of public health*

Screening and Test Evaluation Program (STEP), School of Public Health, University of Sydney, NSW 2006, Australia.

Paul Glasziou *professor*

Department of Primary Health Care, University of Oxford, Oxford OX3 7LF

Funding: Supported in part by programme grant 211205 from the Australian National Health and Medical Research Council and a University of Sydney Medical Foundation Program Grant.

Competing interests: None declared.

- 1 Yaffe M. What should the burden of proof be for acceptance of a new breast-cancer screening technique? *Lancet* 2004;364:1111-2
- 2 Robson ME, Offit K. Breast MRI for women with hereditary cancer risk. *JAMA* 2004;292:1368-70.
- 3 Liberman L. Breast cancer screening with MR—what are the data for patients at high risk? *N Engl J Med* 2004;351:497-500.
- 4 Elmore JG, Armstrong K, Lehman CD, Fletcher SW. Screening for breast cancer. *JAMA* 2005;293:1245-56.
- 5 International Agency for Research on Cancer. *IARC handbooks of cancer prevention, volume 7: breast cancer screening*. Lyon: IARC Press, 2002.
- 6 Irwig L, Houssami N, van Vliet C. New technologies in screening for breast cancer: a systematic review of their accuracy. *Br J Cancer* 2004;90:2118-22.
- 7 Ernster V, Barclay J, Kerlikowske K, Grady D, Henderson I. Incidence of and treatment for ductal carcinoma in situ of the breast. *JAMA* 1996;275:913-8.
- 8 Zahl PH, Strand BH, Maehlen J. Incidence of breast cancer in Norway and Sweden during introduction of nationwide screening: prospective cohort study. *BMJ* 2004;328:921-4.
- 9 Paci E, Warwick J, Falini P, Duffy SW. Overdiagnosis in screening: is the increase in breast cancer incidence rates a cause for concern? *J Med Screen* 2004;11:23-27.
- 10 Boer R, Warmerdam P, de Koning H, van Oortmarssen G. Letter: Extra incidence caused by mammographic screening. *Lancet* 1994;343:979.
- 11 Taylor R, Supramaniam R, Rickard M, Estoesta J, Moreira C. Interval breast cancers in New South Wales, Australia, and comparisons with trials and other mammographic screening programmes. *J Med Screen* 2002;9:20-5.
- 12 Zackrisson S, Andersson I, Janzon L, Manjer J, Garne JP. Rate of over-diagnosis of breast cancer 15 years after end of Malmö mammographic screening trial: follow-up study. *BMJ* 2006;332:689-92.
- 13 Chock C, Irwig L, Berry G, Glasziou P. Comparing dichotomous screening tests when individuals negative on both tests are not verified. *J Clin Epidemiol* 1997;50:1211-7.
- 14 Pisano ED, Gatsonis C, Hendrick E, Yaffe M, Baum JK, Acharyya S, et al. Diagnostic performance of digital versus film mammography for breast-cancer screening. *N Engl J Med* 2005;353:1773-83.
- 15 Australian Government Department of Health and Ageing. *Bowel cancer screening pilot program*. www.cancerscreening.gov.au/bowel/bcaust/pilot.htm (accessed 15 Nov 2005).