Why we need more breast cancer screening trials

Joshua J Fenton,[1] Diana L Miglioretti[2]

[1]Department of Family and Community Medicine and the Center for Healthcare Policy and Research, University of California, Davis, Sacramento, California, USA
[2]Group Health Research Institute, Group Health Cooperative, Seattle, Washington, USA

Correspondence to:
Dr Joshua J Fenton
Department of Family and Community Medicine and the Center for Healthcare Policy and Research, University of California, Davis, 4860 Y Street, Ste 2300, Sacramento, CA95817, USA;
joshua.fenton@ucdmc.ucdavis.edu

Despite a high false-positive rate, screening mammography fails to detect one in five breast cancers and even fewer in women with dense breasts. New technologies have been developed to address these limitations, including digital mammography, breast tomosynthesis, MRI and computer-aided detection (CAD). Conceivably, a new technology—either alone or alongside mammography—could yield net benefits to women, ushering in a new era of breast cancer screening. But what sort of data will be needed to infer that a new screening method is better than mammography alone?

New breast cancer screening modalities would ideally be evaluated in head-to-head randomised trials comparing breast cancer mortality in patients screened with the new modality versus patients screened with conventional mammography. In light of variable interpretation across radiologists and the relatively small incremental benefits of new screening modalities, head-to-head trials would likely require huge sample sizes of patients and radiologists to achieve sufficient statistical power.[1] In addition, for mortality outcomes, many years of follow-up are required, so evaluated technologies may be obsolete by the time trial findings become available.

Despite these challenges, we still believe that head-to-head clinical trials are necessary to inform clinical and policy decisions regarding breast cancer screening. However, future trials of breast cancer screening will necessarily rely on near-term surrogate outcomes, ideally outcomes that strongly correlate with decreased breast cancer mortality, such as the incidence rate of *interval cancers* or incidence rate of *late-stage cancers*. (Interval cancers are cancers diagnosed between screening rounds and putatively reflect both missed cancers and highly aggressive cancers unlikely to be screen-detected.) However, both outcomes are rare, and trials designed with these endpoints would be costly due to the very large sample sizes.

Nevertheless, from the societal perspective, the cost of such trials may still be small compared with the cumulative costs of premature technology adoption. Although no trial has evaluated its impact on interval or late-stage cancer incidence, CAD is now used on most screening mammograms in the USA (increasing the cost of each mammogram by at least 10%). It is difficult to estimate the cost of an adequately powered trial testing CAD's impact on these outcomes, but the Digital Mammographic Imaging Screening Trial (DMIST) cost ~$26 million to compare sensitivity and specificity of digital versus film-screen mammography in over 49 000 women who each received both examinations.[2] If the cost of a head-to-head trial of CAD use versus non-use were fivefold greater than DMIST (~$125 million), this cost would still be one-fourth the approximate total annual cost of CAD use within the USA (~$500 million).[3]

What is the role of further screening trials like DMIST that assess more proximate surrogate outcomes, such as sensitivity and specificity? For increased sensitivity to lead to reduced breast cancer mortality, cancers must be detected significantly earlier (when treatments are more likely to improve survival) than with an alternative method with lower sensitivity. With improved breast cancer treatments, this is a challenging goal to meet. In addition, more sensitive examinations usually reduce specificity and may increase overdiagnosis. Thus, by themselves, trials examining screening accuracy cannot directly address whether the benefits of new technologies are likely to outweigh potential harms.

But data from trials assessing sensitivity and specificity can be used in natural history models of breast cancer that can evaluate the long-term impacts of screening under a variety of real-world scenarios, ranging from varying screening performance to differences in the starting ages or the intervals of screening.[4][5] Microsimulation models can explicitly weigh the mortality benefits of new technologies (often mediated by reduced incidence of late-stage disease) and potential harms (eg, reduced quality of life following overdiagnosis and non-beneficial treatment). Model inputs can also be modified based on community-based observational studies as they emerge.

We recognise challenges to implementing large screening trials, including limited funding and opportunity costs. Although formidable, challenges are probably not insurmountable with sufficient push from funding and regulatory bodies. By orchestrating the roll-out of new screening regimens in different regions, leaders in Norway have planned a series of randomised trials to address crucial questions about colorectal cancer screening.[6] Although screening is not delivered by a national programme in the USA, Medicare could condition coverage of new breast cancer screening technologies based on trial participation, or the collection of high-quality registry data for observational research.[7]

There remains a vital role for breast cancer screening trials that examine not only sensitivity and specificity but near-term surrogates for breast cancer mortality. The challenge of overcoming the logistical and political barriers to trial implementation will make it tempting to do nothing. National leaders and policymakers will need to articulate and sustain the argument that the societal benefits of large screening trials are too great to allow new screening technologies to disseminate without rigorous evaluation.

Competing interests None.

References
1. Jiang Y, Miglioretti DL, Metz CE, et al. Breast cancer detection rate: designing imaging trials to demonstrate improvements. *Radiology* 2007;243:360–7.

2. **Pisano ED,** Hendrick RE, Yaffe MJ, *et al.* Diagnostic accuracy of digital versus film mammography: exploratory analysis of selected population subgroups in DMIST. *Radiology* 2008;**246**:376–83.

3. **Fenton JJ,** Taplin SH, Carney PA, *et al.* Influence of computer-aided detection on performance of screening mammography. *N Engl J Med* 2007;**356**:1399–409.

4. **Berry DA,** Cronin KA, Plevritis SK, *et al.* Effect of screening and adjuvant therapy on mortality from breast cancer. *N Engl J Med* 2005;**353**:1784–92.

5. **van Ravesteyn NT,** Miglioretti DL, Stout NK, *et al.* Tipping the balance of benefits and harms to favor screening mammography starting at age 40 years: a comparative modeling study of risk. *Ann Intern Med* 2012;**156**:609–17.

6. **Bretthauer M,** Hoff G. Comparative effectiveness research in cancer screening programmes. *BMJ* 2012;**344**:e2864.

7. **Tunis SR,** Pearson SD. Coverage options for promising technologies: Medicare's 'coverage with evidence development'. *Health Aff (Millwood)* 2006;**25**:1218–30.