Consequences of False-Positive Screening Mammograms

Anna N. A. Tosteson, ScD; Dennis G. Fryback, PhD; Cristina S. Hammond, MPH; Lucy G. Hanna, MS; Margaret R. Grove, MS; Mary Brown, MPH; Qianfei Wang, MS; Karen Lindfors, MD, MPH; Etta D. Pisano, MD

**IMPORTANCE** False-positive mammograms, a common occurrence in breast cancer screening programs, represent a potential screening harm that is currently being evaluated by the US Preventive Services Task Force.

**OBJECTIVE** To measure the effect of false-positive mammograms on quality of life by measuring personal anxiety, health utility, and attitudes toward future screening.

**DESIGN, SETTING, AND PARTICIPANTS** The Digital Mammographic Imaging Screening Trial (DMIST) quality-of-life substudy telephone survey was performed shortly after screening and 1 year later at 22 DMIST sites and included randomly selected DMIST participants with positive and negative mammograms.

**EXPOSURE** Mammogram requiring follow-up testing or referral without a cancer diagnosis.

**MAIN OUTCOMES AND MEASURES** The 6-question short form of the Spielberger State-Trait Anxiety Inventory state scale (STAI-6) and the EuroQol EQ-5D instrument with US scoring. Attitudes toward future screening as measured by women’s self-report of future intention to undergo mammographic screening and willingness to travel and stay overnight to undergo a hypothetical new type of mammography that would identify as many cancers with half the false-positive results.

**RESULTS** Among 1450 eligible women invited to participate, 1226 (84.6%) were enrolled, with follow-up interviews obtained in 1028 (83.8%). Anxiety was significantly higher for women with false-positive mammograms (STAI-6, 35.2 vs 32.7), but health utility scores did not differ and there were no significant differences between groups at 1 year. Future screening intentions differed by group (25.7% vs 14.2% more likely in false-positive vs negative groups); willingness to travel and stay overnight did not (9.9% vs 10.5% in false-positive vs negative groups). Future screening intention was significantly increased among women with false-positive mammograms (odds ratio, 2.12; 95% CI, 1.54-2.93), younger age (2.78; 1.5-5.0), and poorer health (1.63; 1.09-2.43). Women’s anticipated high-level anxiety regarding future false-positive mammograms was associated with willingness to travel overnight (odds ratio, 1.94; 95% CI, 1.28-2.95).

**CONCLUSIONS AND RELEVANCE** False-positive mammograms were associated with increased short-term anxiety but not long-term anxiety, and there was no measurable health utility decrement. False-positive mammograms increased women’s intention to undergo future breast cancer screening and did not increase their stated willingness to travel to avoid a false-positive result. Our finding of time-limited harm after false-positive screening mammograms is relevant for clinicians who counsel women on mammographic screening and for screening guideline development groups.
A substantial proportion of women who undergo routine screening mammography during a 10-year period will have a false-positive mammogram, requiring additional workup to rule out breast cancer.\(^4\) The rate of false-positive mammograms leading to unnecessary biopsies with benign results, compared with the number of cancers detected, contributed to the 2009 changes in the US Preventive Services Task Force breast cancer screening guidelines.\(^4\) Instead of recommending routine screening among women 40 to 49 years old, these guidelines recommend that women in their 40s discuss the pros and cons of mammographic screening with their health care providers before deciding whether to initiate screening. This recommendation acknowledges that individual women’s preferences may differ regarding the balance of screening benefits vs harms, which include false-positive mammograms. As the US Preventive Services Task Force reevaluates the evidence for breast cancer screening, the harms of screening are among the questions to be addressed in their systematic evidence review.\(^4\)

Although there is a growing literature on how women view false-positive screening mammograms,\(^3,5,23\) few studies have attempted to assess the effect of false-positive results on generic measures that allow comparison to a broad range of health outcomes. To include such a harm in a societal cost-effectiveness analysis, the effect of false-positive screening mammograms on generic health utility (on a scale where 0 represents death and 1 represents perfect health) is needed to compute cost-effectiveness results, which can be compared with other health care interventions.\(^24\) These facts—combined with results of early digital mammographic screening studies\(^25,29\) which suggested that digital mammography may yield fewer false-positive results than screen-film mammography—led to inclusion of a quality-of-life (QoL) substudy in the American College of Radiology Imaging Network Digital Mammographic Imaging Screening Trial (DMIST).\(^30\) The QoL substudy was designed to characterize the personal anxiety, disability, and personal time costs associated with workup of positive screening mammograms. Personal time costs of mammographic screening outcomes in DMIST were reported and used in the DMIST cost-effectiveness analysis.\(^31\) In this article, we report DMIST QoL substudy results that characterize the effect of false-positive screening mammograms on personal anxiety, health utility, and attitudes toward future screening.

**Methods**

**DMIST Description**

The DMIST was funded by the National Cancer Institute and conducted by the American College of Radiology Imaging Network, as described in detail elsewhere.\(^30\) In brief, its primary aim was to compare the diagnostic accuracy of digital mammography with that of screen-film mammography.\(^30,35\) Secondary aims included an economic evaluation of digital mammography\(^31\) and an assessment of the effect of false-positive screening mammograms on QoL.

To be eligible for DMIST, women had to present for screening, agree to undergo follow-up mammography, and provide written consent to study participation. Women were excluded if they had a focal dominant lump or bloody or clear nipple discharge, a history of breast cancer treated with lumpectomy, or breast implants, or if they believed they might be pregnant. For eligible women, both digital and screen-film mammograms were obtained, which were read independently by different radiologists. The DMIST study protocol was approved by the institutional review boards at all study sites.

**DMIST QoL Substudy Participants**

The substudy was conducted by telephone survey and included 2 groups of women: (1) a random sample of DMIST participants with a positive screening mammogram, defined as any mammogram for which additional workup or consultation was recommended; and (2) a sample of women with a negative screening mammogram. To ensure a comparable number of participants with positive and negative mammograms from each site, for each woman selected with a positive screening mammogram, a woman from the same institution who had a negative screening mammogram and was of similar age (within 5 years) was also selected for participation. Women in whom breast cancer was diagnosed at any time during DMIST were ineligible and were excluded.

**Measures**

**Anxiety**

The Spielberger State-Trait Anxiety Inventory (STAI) is a widely used measure of general anxiety that includes both a state scale and a trait scale.\(^32\) In the current study, we are interested in state anxiety—that is, anxiety of the moment as experienced by the person. To measure this, we used a validated 6-question short form of the STAI state scale (STAI-6), which yields a score between 20 (least anxious) and 80 (most anxious).\(^35\)

**Health Utility**

To characterize general health-related QoL, we used the validated EuroQol EQ-5D instrument, which consists of 5 questions, a question each about mobility, self-care, usual activities, pain/discomfort, and anxiety/depression.\(^36\) Each question has 3 possible responses categorizing the degree of problem with the particular aspect of health, ranging from 1 (no problem) to 3 (extreme problem). The 5 questions and 3 response categories are used to define 245 distinct health states. Scoring using US preference weights assigns health state values ranging from –0.11 (worst health for those with extreme problems in all 5 areas) to 1.0 (best health for those with no problem in any area).\(^37\) A current health rating scale asked “On a scale from 0 to 100 where 0 represents death or the worst health you can imagine and 100 represents perfect health or the best health you can imagine, how would you rate your health during the past 4 weeks?”

**Breast-Related Resource Use**

At follow-up, women were asked “since your mammogram approximately 12 months ago [date provided], have you received medical care for any breast-related concerns?” Women...
who answered affirmatively were read a list of tests, procedures, and health provider visits that some women have after a mammogram and were asked which they had undergone.

Future Screening Attitudes
Women were asked to think about “how your breast-related care in the past year may change your future use of screening mammography” and to choose the statement that best represented their feelings at the time of the follow-up survey. Response choices included the following: “I am less likely than a year ago to undergo screening mammography in the future,” “My use of screening mammography will not change,” and “I am more likely than a year ago to undergo screening mammography in the future.”

Two questions asked women their opinion about new types of mammograms that may be developed in the future. A “willingness to pay to avoid” approach was used to assess how women valued false-positive screening mammograms. This was implemented by asking about the duration of travel women would be willing to undertake to gain access to a hypothetical new type of mammography that would produce fewer false-positive mammograms while identifying just as many cancers. This survey item was developed for the current study and was modeled on the waiting-time trade-off developed to value transient health states encountered in radiologic cost-effectiveness analyses, with travel serving as a metaphor for lost time.

Women were also asked to imagine that they could choose between 2 new types of mammography that are just as accurate as those today: one that resulted in fewer false-positive mammograms and one that caused less breast compression (see Appendix in the Supplement for details). Development of the travel and new mammography type survey items was informed by focus groups of women who had experienced false-positive screening mammograms.

Telephone Interview Schedule
Telephone interviews were conducted shortly after the baseline mammogram and approximately 1 year later. At both times, the interview included the STAI-6, the EQ-5D, and the current health rating scale.

The baseline interview was intended to occur after notification of the need for further workup and before workup completion—during what we define as the active workup window. This was not always possible. Based on follow-up data on the date of subsequent breast-related medical care, we categorized baseline interviews as occurring during or after the active workup window. Women interviewed after the active workup window were considered to have had “resolved” false-positive mammograms at the time of the baseline interview. We hypothesized that, compared with women with negative screening mammograms, anxiety would be highest and QoL lowest for women with positive screening mammograms who were interviewed in the active workup window (ie, before the false-positive mammogram was resolved). We further hypothesized that elevated anxiety and reduced health utility associated with a false-positive mammogram would be transient, such that at follow-up there would be no differences between those with positive and negative screening mammograms.

At the second interview, additional information was collected on the use of breast-related resources and attitudes toward future screening mammography.

Statistical Analysis
Characteristics of participants were summarized using means and proportions, with comparisons between women with false-positive vs negative mammograms made using Student’s t tests and \( \chi^2 \) tests. To assess the effect of a false-positive mammogram on anxiety and health utility, the changes between baseline and follow-up scores were compared between women in the false-positive and negative mammogram groups. In logistic regression analyses, we explored factors associated with women’s self-report of their increased likelihood to undergo future breast cancer screening and with the odds of being willing to travel overnight to avoid a false-positive mammogram. Variables considered included age, breast cancer risk, baseline mammogram positivity, need for biopsy, anticipated anxiety regarding future false-positive mammograms, race/ethnicity, and institution.

Results
Among 1450 eligible women invited to participate, 1226 (84.6%) from 22 institutions were enrolled in the QoL substudy, and follow-up interviews were obtained for 1028 (83.8%). Women with false-positive mammograms tended to be younger but did not differ on any other characteristics (Table 1).

At baseline, there were significant differences in anxiety between women with positive and negative mammograms but no statistically significant differences in the EQ-5D or the current health rating scale (Table 2). At follow-up, a significant decline in anxiety was noted (mean STAI-6 difference, −1.53; 95% CI, −2.70 to −0.35) among women with positive mammograms (Table 2). Women with negative mammograms were found to have a modest but significant decline in rating scale at follow-up (mean rating scale difference, −1.22; 95% CI, −2.34 to −0.10).

Use of breast-related care reported 1 year after the initial screening examination differed by positivity status for all categories except clinical breast examination (Table 3). Biopsy procedures were used in 14.6% of women with a false-positive mammogram compared with 1.1% of those with a negative mammogram. Among women with a false-positive mammogram, anxiety was reported as moderate or higher by 50.6% and as extreme by 4.6%.

Women’s plans to undergo mammography within the next 2 years did not differ by screening outcome, but significantly more women who had a false-positive mammogram characterized themselves as “more likely” to undergo future breast cancer screening (25.7%) compared with those who had a negative mammogram (14.2%) (Table 4). However, no differences were noted in women’s attitudes toward the anticipated anxiety they would feel if they were to have a positive mammogram in the future. Many women in each group (61.2% in the negative and 58.9% in the false-positive group) thought they would experience anxiety that was moderate or worse if they
were to have a false-positive mammogram in the future, and a substantial proportion anticipated feeling a high level of anxiety (with about 27% characterizing anticipated anxiety as “a lot” or “extreme” in each group).

Women's experience of a false-positive mammogram did not influence their willingness to travel to avoid a false-positive mammogram in the future, with the vast majority of women in both groups being willing to travel up to 4 hours to...
Willingness to travel and double a woman’s stated intention to use future screening variable analyses, a false-positive mammogram more than that they would undergo future screening (odds ratio, 2.12; 95% CI, 1.54-2.93). Willingness to travel was associated with women’s anticipated feelings of anxiety surrounding a future false-positive mammogram. Instead, willingness to travel was associated with women’s anticipated feelings of anxiety and that the effects were transient and not measurable with the EQ-5D questionnaire or current health rating scale. A small proportion of women were willing to travel and stay overnight to undergo screening with a hypothetical modality that would have fewer false-positive results. Instead, willingness to travel and stay overnight to avoid a false-positive mammogram did not vary by positivity but was associated with women’s reports of anticipated anxiety if they were to have a false-positive mammogram in the future (Table 5).

Discussion

The DMIST QoL substudy provides evidence that women with a false-positive mammogram initially experienced a significant increase in anxiety and that these effects were transient and not measurable with the EQ-5D questionnaire or current health rating scale. A small proportion of women were willing to travel and stay overnight to undergo screening with a hypothetical modality that would have fewer false-positive findings, but this willingness was not influenced by experience of a false-positive mammogram. Instead, willingness to travel was associated with women’s anticipated feelings of anxiety surrounding a future false-positive mammogram. Women who anticipated feeling more moderate anxiety were nearly twice as likely to report a willingness to travel and stay overnight as those who anticipated feeling less anxiety. Our results regarding the transient anxiety associated with
false-positive mammograms may provide useful information for health care providers who counsel individual women regarding the decision whether to initiate breast cancer screening and for policy makers assessing the clinical effectiveness of mammographic screening.

Although there is concern that the health and psychological burden of false-positive mammograms may not be justified when weighed relative to the few additional breast cancers that routine screening would identify among younger women, we found that a false-positive mammogram had only transient effect on anxiety. Our finding differs from recent reports of longer-term effect of false-positive mammograms on specific psychological outcomes.21-23 However, the ongoing harms reported in the literature are related primarily to outcomes specific to breast cancer rather than general psychological measures, such as the general anxiety measure used in our study. It is an open question whether one should expect harms that are measurable only when framed in terms of a specific disease to affect a general anxiety or health utility measure. An important feature of the 2 generic measures used in our study, STAI-6 and EQ-5D, is the ability to compare potential harms associated with breast cancer screening to those associated with a broad range of other health care practices.

In addition to the transient effect on anxiety, we also provide evidence that women are motivated to avoid false-positive mammmograms. A large proportion of women were willing to travel up to 4 hours to undergo a hypothetical new type of mammography that would be just as good at finding cancer but would result in fewer false-positive mammmograms. Further supporting women's interest in avoiding false-positive mammograms, when given the choice between 2 new types of mammography—one with fewer false positives and one with reduced breast compression—most women in our study chose the type with fewer false positives.

Our study also characterized the anxiety and discomfort imposed by the health care visits and procedures used to resolve positive screening mammograms. Although anxiety and discomfort were significantly higher among women with false-positive mammograms than among those with negative mammograms, these findings seemed to have no effect on women's plans to undergo mammography in the next 2 years—something that 93.5% of women intended to do. In contrast, women's intention to use breast cancer screening in the future (ie, self-characterization as being more likely to undergo future breast cancer screening) was increased by 2-fold among women who had a false-positive mammogram. The reasons for this are unclear, but a meta-analysis of observed screening practices after a positive mammogram has similarly concluded that women in the United States who have a false-positive mammogram are significantly more likely to participate in future mammographic screening.31

The DMIST QoL substudy results were used to estimate previously reported economic time costs associated with follow-up of screening mammograms and were incorporated into the cost-effectiveness analysis of digital mammography.31 In this article, we have provided further information on the use of follow-up tests. Not surprisingly, some women with negative screening mammograms reported use of breast-related medical services within the year after screening (eg, additional imaging in 4.5% and breast biopsy in 1.1%), but use of such services were associated with lower levels of anxiety and less discomfort than was reported among women with false-positive mammograms.

The DMIST cost-effectiveness analysis did not adjust for a QoL/health utility loss when estimating quality-adjusted life years for women with false-positive screening mammograms, a decision supported by the fact that our results showed no measurable impact with EQ-5D.28 Nonetheless, measures such as the EQ-5D, which are appropriate for use in societal cost-effectiveness studies, are of limited value in clinical settings where individual women must consider how they value potential screening outcomes when weighing the benefits and harms of breast cancer screening. In this context, there were no differences reported between women in the false-positive and negative groups concerning how much anxiety they would anticipate feeling about a positive screening mammogram in the future. However, 26.9% of women reported that they would feel more than a moderate amount of anxiety/concern (ie, a lot of anxiety/concern or extreme anxiety/concern) if such a screening outcome were to occur. Thus, it seems prudent for women and their clinicians to consider the potential outcomes of screening when deciding on the best course of action for each woman.

Table 5. Correlates of Greater Intention to Use Future Screening and Willingness to Travel and Stay Overnight to Avoid a False-Positive Mammogram

<table>
<thead>
<tr>
<th>Correlates</th>
<th>Odds Ratio (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Univariate Analysis</td>
</tr>
<tr>
<td>Intention to use future breast cancer screening</td>
<td>2.07 (1.51-2.84)</td>
</tr>
<tr>
<td>False-positive screening mammogram</td>
<td></td>
</tr>
<tr>
<td>Age &lt;65 y</td>
<td>2.44 (1.39-4.35)</td>
</tr>
<tr>
<td>EQ-5D baseline score (good vs perfect)</td>
<td>1.00 (0.69-1.46)</td>
</tr>
<tr>
<td>EQ-5D baseline score (worst vs perfect)</td>
<td>1.52 (1.03-2.24)</td>
</tr>
<tr>
<td>Willingness to travel and stay overnight to</td>
<td></td>
</tr>
<tr>
<td>avoid a false-positive mammogram</td>
<td></td>
</tr>
<tr>
<td>False-positive screening mammogram</td>
<td>0.93 (0.62-1.40)</td>
</tr>
<tr>
<td>Anticipated high level of anxiety concerning</td>
<td>1.94 (1.28-2.94)</td>
</tr>
<tr>
<td>future false-positive mammogram*</td>
<td></td>
</tr>
</tbody>
</table>

* Values >1 indicate greater intention to use screening or willingness to travel.

a EQ-5D score of <0.1 and >0.8271.

b EQ-5D score of 1.0.

c EQ-5D score of <0.8271.

Women reporting that their level of anxiety/concern would be "a lot" or "extreme" if they were to have a false-positive mammogram in the future.
Studies to improve breast screening practices by finding approaches that reduce such anxiety are also warranted. Several limitations to our study deserve comment. First, we intended to interview women during the active workup window, but this was not always achieved. As a result, many women were interviewed after workup was completed. Nonetheless, we were able to document heightened anxiety associated with false-positive mammograms. We did not document any effect on overall health utility using the EQ-5D preference-based measure that is appropriate for use in estimating quality-adjusted life years. This may be due to limitations in the EQ-5D descriptive system, which combines the rating of anxiety and depression together at only 3 levels: no anxiety/depression, moderate anxiety/depression, and extreme anxiety/depression. Alternatively, it may reflect the general insensitivity of the EQ-5D to small changes in the health of healthy persons, or it may suggest a relatively low anxiety effect for false-positive mammograms. In addition, at baseline most participants rated their health at the ceiling of the EQ-5D descriptive system in both the false-positive (272 of 494 patients [55.1%]) and negative (293 of 531 [55.2%]) groups.

Second, the attitudes of DMIST volunteers may not be representative of the general population of women eligible for breast cancer screening. Enthusiasm for screening may be higher among DMIST participants than in the general population, a conjecture supported by the fact that 93.5% of women reported their intention to undergo breast cancer screening with mammography again within 2 years. However, attitudes toward cancer screening in the United States have generally been documented as highly favorable.39

Conclusions

We conclude that false-positive mammograms are associated with a measurable, small, and transient effect on personal anxiety and that further research should address opportunities for reducing this anxiety. Although the effect of false-positive mammograms on health utility for estimating quality-adjusted life years is not measurable using the current EQ-5D instrument, it is clear that women, regardless of whether or not they had a false-positive mammogram, are willing to invest the necessary travel time to avoid such results in the future. The fact that women’s anticipated anxiety about future false-positive mammograms was a correlate of willingness to travel and stay overnight to avoid such a result, but the actual experience of a false-positive mammogram was not, further highlights opportunities to educate women about screening outcomes. Although health utility scores were not measurably affected, the experience of a false-positive mammogram did increase women’s intentions to undergo breast cancer screening in the future. Ongoing studies of breast cancer screening processes of care through the National Cancer Institute’s PROSPR (Population-based Research Optimizing Screening through Personalized Regimens) initiative may further elucidate the frequency and sequelae of false-positive screening mammograms. Meanwhile, our report on women’s experience of false-positive mammograms may provide useful information for those counseling women on the decision to undergo mammographic screening and for screening guideline development groups.

REFERENCES
9. Brett J, Austoker, J. Ong G. Do women who undergo further investigation for breast screening suffer adverse psychological consequences? a multi-centre follow-up study comparing different...
Are the Harms of False-Positive Screening Test Results Minimal or Meaningful?

Kurt Kroenke, MD

That most screening test results will be normal or negative is commonplace, but the reality that abnormal results are frequently false-positive is not always well appreciated, nor is it fully conveyed to patients. How does a patient feel after a false-positive test result? Tosteson and colleagues concluded from their longitudinal study that “false-positive mammograms are associated with a measurable, small, and transient effect on personal anxiety.” However, a closer look at all the outcomes assessed in this well-done study reveal some adverse consequences that, although not serious, may nonetheless be meaningful.

The authors chose as their primary outcomes a 6-item version of the Spielberger State-Trait Anxiety Inventory (STAI-6) and the 5-item EuroQol instrument (EQ-5D). The STAI-6 is a general rather than disease-specific measure of anxiety, and research has shown that questions directed at specific cancer fears or test-related worries are more sensitive in detecting patient distress after mammography with false-positive results.2


Related article page 954

Invited Commentary

Downloaded From: by a Non-Human Traffic (NHT) User on 11/17/2018