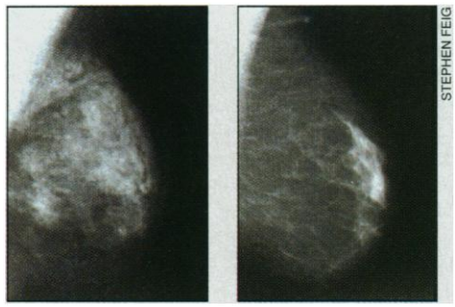
The Breast-Screening Brawl

The controversy over whether regular mammograms should be recommended to women in their forties has been stoked by uncertain evidence, opposing world views, and plenty of invective

The question is deceptively simple: Should regular mammography screening be recommended for women in their forties? But going into a crucial meeting at the National Institutes of Health (NIH) last month, which was supposed to come up with a consensus, the researchers and clinicians involved seemed to agree only that there was no consensus. In the words of Harvard Medical School radiologist Daniel Kopans: "Everybody who knows anything about it has taken one position or another." This was a bad sign. The planning for the meeting was described by one member of the planning committee as "a brawl," and the meeting itself then lived down to expectations. National Cancer Institute (NCI) director Rick Klausner said he was "shocked to see firsthand the vitriol and animosity" surrounding the issue. And Barry Kramer, editor of the *Journal of the National Cancer Institute (JNCI)*, described the meeting simply as "raucous" and "like no scientific meeting I'd ever seen before."

After a hectic two-and-a-half days, it ended with public accusations of fraud and bureaucratic chicanery (see box). One radiologist opined publicly, and with excellent press coverage, that the panel's conclusion that every woman in her forties would have to decide for herself on mammography constituted a "death sentence" for those women. He would mourn for them, he said. Now, the combatants are regrouping for another round, which will come next week, when the National Cancer Advisory Board (NCAB) will meet to consider the next step toward resolving the controversy, one of the most emotional in medicine.

The irony is that the contentious ground lies so near an area of agreement. There's little argument that mammography beginning at the age of 50 saves lives. But biology complicates the question for younger women. Mortality from breast cancer is 30 per 100,000 in women in their forties, compared to 126 per 100,000 for women 65 and over. Yet, cancer in younger women tends to be more virulent, growing faster, and killing faster. These factors mean that any test to find it has to be that much more sensitive to be useful, but cancer is also more difficult to detect in younger women. Their breasts tend to have more glandular tissue and less fat than those of older women, and the glandular tissue is often of the same density as tumors, hiding them on x-ray.



STEPHEN FEIG

Cloudy tissue. Cancers can be harder to detect in a dense breast *(left)*, more typical of younger women, than in a fatty "lucent" breast.

As a result, while everyone agrees that screening should be recommended for women in their fifties and not for women in their thirties, the decade in between has become the battleground. Over the past 30 years, evidence of a benefit from screening in this age group has gradually climbed toward statistical significance. But the studies, many of them decades old, remain riddled with ambiguities. And it isn't just the sparse and contradictory evidence that is driving the controversy. To understand its emotional pitch, say those involved, requires knowing the two sides and their differing world views.

The experts who believe screening should be officially recommended to women in their forties are mainly radiologists, with a few surgeons and a very few epidemiologists. To them, says radiologist Laszlo Tabar of the Central Hospital in Falun, Sweden, the question is whether doctors can save lives by giving mammograms to younger women. And the answer—seemingly supported by the latest meta-analyses of the clinical trials—is yes. "We're talking about mortality," he says. Radiologists live every day with the tragedy of breast cancer, and they see firsthand how that tragedy can be averted by screening. As University of California, San Francisco (UCSF), radiologist Ed Sickles puts it, "People who are on [the other] side are primarily people who are too far removed from actually seeing what's happening, what we're doing at the clinical level."

Those who are skeptical of the benefit of screening younger women are mainly epidemiologists and public health physicians versed in the science of evidence-based medicine. For them, the relevant figures are not just how many lives could be saved, but how many women have to be screened to save one life, and at what cost in false positives and unnecessary surgery. Russell Harris, for instance, director of a University of North Carolina Medical School program on health promotion and disease prevention, has pointed out that if mammography can reduce breast cancer mortality by 20% in women in their forties, then screening 100,000 women for a decade will reduce deaths in that group from 30 to 24. A radiologist who personally screens 1000 women every year through their forties, says Harris, will "have prolonged no one's life 9 years after beginning screening."

To this group—which includes, curiously enough, virtually all of the women active in the controversy, with the exception of breast cancer

RANDOMIZED TRIALS AND RESULTS, AS OF 1996

Study	Dates	Age at Entry	No. of Women Study/Control	Screening Frequency (months)	Years of Follow-Up (median)	Breast Cancer Mortality (study/ctrl)	95% Confidence Interval
HIP (Health Ins. Plan of NY)	1963–69	40–49	14,849/14,849	12 (4 rounds)	18	0.77	0.53–1.11
Malmö	1976–86	45–49	3658/3769	18–24 (5 rounds)	15.5	0.66	0.32–1.27
Two County	1977–85	40–49	19,844/15,604	24 (4–5 rounds)	14–15	0.91	0.59–1.39
Edinburgh	1979–88	45–49	5913/5810	12	11	0.73	0.43–1.25
Canadian (NBSS-1)	1980–87	40–49	25,214/25,216	12 (5 rounds)	10.5 (mean)	1.10	0.78–1.54
Stockholm	1981–85	40–49	14,375/7103	28 (2 rounds)	11.9	1.05	0.53–2.09
Gothenburg	1982–88	40–49	10,821/13,101	18 (4 rounds)	9.9	0.62	0.36–1.08

SOURCES: P. GLASZIOU; FALUN MEETING REPORT

How One Radiologist Turns Up the Heat

In the debate over mammography screening for women in their forties, plenty of factors conspire to raise the emotional pitch. There is the horror of breast cancer, the ambiguous data on the benefits of screening, and the conflict between the outlooks of the radiologists who favor regular screening and the epidemiologists and public health specialists who urge caution (see main text). But one individual has contributed more than his share of acrimony, according to members of both camps: Harvard Medical School radiologist Daniel Kopans.

"I am just trying to get the truth out in a clear and balanced fashion."
—Daniel Kopans

Kopans's latest and most public assault on his opponents came last month just after a National Cancer Institute panel refused to recommend regular screening for women in their forties. The panel's conclusion was "fraudulent," Kopans told the press. He later explained that the panel had ignored new data and come to conclusions that were not only "grossly misleading" but "meant to mislead." The accusation is only one of many he has made over the years.

For the past decade, Kopans, who is director of breast imaging at Massachusetts General Hospital, has been an outspoken proponent of screening in the scientific community and one of the most widely cited in the popular press. Researchers in the field say Kopans's scientific arguments have prompted them to look at the data more rigorously. But they have also portrayed his methods as "intellectual terrorism" or "scientific McCarthyism." National Cancer Institute (NCI) director Rick Klausner, for instance, says Kopans employs "a pattern of inflammatory, accusatory approaches that are antithetical to the requirements of scientific discourse."

Kopans's charges have included the following:

- that the NCI has stacked allegedly objective review panels with "opponents of screening";
- that administrators at the NCI misled congressional investigators looking into its 1994 decision to eliminate screening guidelines;
- that the *Journal of the National Cancer Institute* (*JNCI*), the *Journal of the American Medical Association* (*JAMA*), and the *Annals of Internal Medicine* are biased and will only publish articles that do not conclude that screening is beneficial, while rejecting any articles that do. Editors of the journals deny the accusations. Nonetheless, says Barry Kramer, editor of *JNCI*, "it has a chilling effect, when letters written about your integrity as an editor are copied to the director of the institute and the NIH, and copied to the press and the White House and to politicians and congressmen"; and
- that two University of California, San Francisco (UCSF), epidemiologists "manipulate[d] data in a fashion that borders on scientific fraud" in papers in *JAMA*. Kopans sent a letter with this accusation to George Lundberg, editor of *JAMA*, and to Haile Debas, dean of the UCSF medical school. The letter "made accusations that I would not expect from a reasonable clinician or scientist," says Debas. Lundberg says an investigation showed that there was "no substance" to the charges.

In defense of his accusations, Kopans says he is "not a terrorist. ... I am just trying to get the truth out in a clear and balanced fashion." He acknowledges that some of his accusations may sound "grossly paranoid." But he stands by his claims and says his aggressive approach is necessary to combat the politicization of science by screening opponents. Administrators at the NCI and researchers who don't endorse screening, he says, are motivated by a need for "petty vindication" and manipulate data to suit preconceived opinions. "People have staked out positions, and now they can't back down," he says. "I'm amazed they have been allowed to get away with it."

Even some of those who favor screening, however, suggest that Kopans may be going too far. Laszlo Tabar, for instance, a radiologist who directs the Swedish Two County Trial of mammography, says Kopans is overly belligerent. Ed Sickles, a UCSF radiologist who has collaborated with Kopans, was a co-author of one of the *JAMA* papers Kopans attacked. "It saddened me that he wrote [the letters]," says Sickles. "[My co-authors] haven't done anything scientifically wrong; they've just looked at the data with a different perspective and come to different conclusions, as have others. It's just frustration on [Kopans's] part." –G.T.

survivors or advocates—the chance of causing harm "is greater than the chance of having the disease or dying from it," says UCSF clinical epidemiologist Karla Kerlikowske. "We're not here to promote a sense of vulnerability and illness," says internist Suzanne Fletcher of Harvard Medical School. "Those of us in prevention are here to promote health."

The dispute is not about money, even though screening advocates occasionally suggest that those skeptical of screening are motivated by a desire to save precious health care dollars. Nor is it about the dangers of radiation—both sides agree that there's little evidence supporting the proposition that mammograms themselves increase the risks of breast cancer. Rather, the animosity is driven by beliefs that are "almost like a religion for some people," says Barbara Rimer, a Duke University behavioral scientist and chair of the NCAB. "They have made up their minds," agrees Sickles, speaking of the epidemiologists and public health experts who believe that regular mammography screening should not be recommended for younger women. "And to me, the answer is clear also."

Nonstatistical fluctuations

The level of vitriol in the field has been building up for a decade as the studies—and the government's stance on screening—have flip-flopped repeatedly. In 1989, the NCI signed onto an agreement among 12 professional organizations recommending screening in women aged 40 to 49 every 1 or 2 years. The data supporting the decision came from six randomized control trials. The first began in New York state in 1963 and was followed by four studies in Sweden and one in Scotland. None had been designed specifically to test the benefits of screening in younger women. Rather, they looked at the effect of screening on all women 40 and over or, in two cases, 45 and up.

The studies concluded that there is a "very convincing benefit for women screened at age 40 and up," says Sickles. To determine a benefit specifically for women in their forties, however, required breaking out that subgroup from the data and trying to analyze it separately. None of the studies had enrolled enough women in their forties to do so with reasonable confidence. When researchers tried this subgroup analysis, they found hints of a benefit for 40-something women, although it was smaller than in older women.

One study was already under way, however, that was advertised as powerful and specific enough to provide an answer: The Canadian National Breast-Screening Study, which set out in the mid-1980s to test annual mammography specifically in women in their forties. The trial, which enrolled 50,000 women in their forties and 39,000 in their fifties, ended in 1988. By 1992, it was already an open secret that the study had come up with

a dramatic and soon-to-be controversial finding: Not only did it show no added benefit from screening in either age group, but the 40-something women who were screened regularly had experienced *more* breast cancer deaths than the control group had. Indeed, the results had become public in part because study organizers had been openly talking to other researchers about how to explain this bizarre result. "We were really rather worried about what was going on here," says Anthony Miller, a University of Toronto physician turned epidemiologist.

Still, the preliminary results spurred the NCI to convene an international workshop in February 1993 to reexamine the evidence supporting screening. The workshop, chaired by Suzanne Fletcher, concluded that a meta-analysis of all the screening trials showed a 39% mortality reduction from screening in older women and no evidence for a benefit for women in their forties. It also noted, however—in a caveat that would help shape the emerging controversy—that the original trials were showing hints of a benefit much later, 10 to 12 years after screening had ended.

What happened to these conclusions as they worked their way up through the NCI bureaucracy was largely responsible for leavening the controversy to the present level of animosity. According to *JNCI*'s Kramer, the NCI's Board of Scientific Counselors of the Division of Cancer Prevention and Control was the first to examine the evidence. It agreed with the workshop's conclusion that screening younger women appeared to provide little benefit and suggested that the NCI drop its recommendation for screening. Next up in the NCI hierarchy was the NCAB, which is composed of scientists, consumers, physicians, and advocates, and which voted 14 to 1 to continue recommending screening, contrary to the conclusion of the workshop.

This decision then went up to Sam Broder, the NCI director at the time, who overturned it. Broder decreed, says Kramer, that NCI "would get out of the guideline business," leaving it to the U.S. Clinical Preventive Services Task Force.

ANNUAL MAMMOGRAPHY SCREENING FOR WOMEN AGED 40–49

DO RECOMMEND

American Cancer Society

American College of Obstetricians and Gynecologists

American College of Radiology

American Medical Association

DO NOT RECOMMEND

American College of Physicians

U.S. Preventive Services Task Force

National Cancer Institute

American Academy of Family Practice

"Even screening every 6 months might not make a difference. That's something that's very hard for even clinicians to understand."
—Suzanne Fletcher

Faultfinding

The twists and turns of events sparked two counterattacks from screening advocates. The first was a 1994 congressional investigation into Broder's decision, led by Brooklyn Representative Edolphus Towns (D–NY). The second was an attack, led by radiologists, on the validity of the Canadian trial. The gist of it, Central Hospital's Tabar explains, was that if the trial had more deaths in the screening group than in the controls, something was wrong with the trial. "You start screening and you expect to provide a benefit, and suddenly people die at a higher rate. Now, hold it, we're not going out and killing women. This demands an explanation," he says.

For starters, the radiologists argued that the Canadian mammograms were unsatisfactory, which might explain the lack of benefit from screening, albeit not the higher mortality. The key criticism, however, focused on the fact that in the very first year of the trial, 19 women in the screening group were diagnosed with the most virulent grade of cancer, while the number of equivalent cancers in the control group was only five. So large a disparity so early in the trial suggested, said the critics, that the method used to assign women to the study and control groups was faulty and had skewed the results.

Because the trial enrolled tens of thousands of women at centers throughout Canada, it relied on a simple method of randomization. When women entered a clinic, explains Robert Phillips, executive director of the Canadian National Cancer Institute, they received a clinical breast examination from a nurse or a physician, who then randomly assigned them to either the control group or the screening group. That's where the excess of cancers in the screened group must have originated, says Stephen Feig, a radiologist at Thomas Jefferson Medical College in Philadelphia who was a consultant for the Canadian trial but quit, he says, because of doubts about mammogram quality. "The women came in with palpable masses, and a nurse in the center, with all good intentions, thought, 'You have a breast mass; let's put you into the arm that's getting mammography,' " says Feig.

Harvard's Kopans, who first publicly voiced this speculation, said he heard from people involved in the study who said this was indeed what had happened. The Canadian NCI, which funded the study, decided it had to investigate and called on epidemiologists John Bailar of the University of Chicago and Brian MacMahon of the Harvard School of Public Health. Their investigation was finished last fall and appeared in the *Canadian Medical Association Journal* the week before the NIH consensus meeting. With the help of Canadian forensics experts, the two epidemiologists searched for evidence that the randomization process had been manipulated. They found, says Bailar, "that there was no plausible evidence that this had occurred." This finding, not surprisingly, failed to appease the critics. "The [Canadian researchers] screwed up so dearly, so beautifully," says Tabar. "Everything they touched was wrong."

The human price

As Broder's decision to drop the screening guidelines played out, two new trends emerged. Follow-up data from the original trials continued to climb toward statistical significance, but at the same time, the terms of the controversy were shifting. For the first time, public health physicians began to look at what Fletcher calls "the human price" of screening millions of healthy women. Studies estimated that from 5% to 11% of mammograms in 40-something women would come up as false positives, requiring further mammograms or biopsies to come to the correct diagnosis and taking a psychological toll.

Researchers also began looking at the complications of what was becoming an epidemic of abnormalities or small cancers known as ductal carcinomas in situ, or DCIS. Beginning in the 1980s, says UCSF epidemiologist Virginia Ernster, there was a 200% to 500% increase in DCIS incidence in the United States, almost assuredly caused by increased use of mammography, which could easily pick up DCIS's characteristic speckled pattern of "microcalcifications" in the ducts.

The complication is that no one knows whether or how much DCIS goes on to become invasive breast cancer. Although studies suggest that anywhere from 30% to 75% of DCISs will not become invasive, all have to be treated as though they will, as soon as they are detected. According to Ernster, over 40% of women with DCIS are having mastectomies, at a rate of some 10,000 a year. "There is the rub," says UCSF's Sickles.

"Once detected, we have to treat it."

What's at issue is whether and how this problem should be taken into account in the screening controversy. Epidemiologists and public health experts argue that it has to be. Screening advocates tend to consider it an issue that doesn't belong in the debate. Tabar, for instance, says he is "flabbergasted" that epidemiologists raise it. "It has nothing to do with mortality, and we're talking about mortality." Adds Kopans: "The goal line keeps moving. ... Most people admit there's a benefit now, so now they're [asking] can we afford it; think of all the harm we're doing."

To Kopans and his colleagues, any doubts about whether screening benefits women in their forties were dispelled last May, at a meeting in Falun, Sweden. One Swedish trial—in Malmö—was reporting a 41% decrease in mortality in younger women who were screened, and a meta-analysis of all the trials together, not including the Canadian one, showed a 23% mortality reduction that verged on statistical significance. Add the Canadian study, and screening still seemed to reduce mortality by 15%. Tabar called it a "landmark meeting." The bottom line, says Sickles, "was now we could tell the world there is a benefit [from screening]. It's probably less than in older women, but there is a benefit."

"People who [oppose screening for these women] are primarily people who are too far removed from actually seeing what's happening."

—Ed Sickles

The new evidence prompted NCI's Klausner to request that the NIH host the consensus conference. But it raised questions of its own. How much of the benefit actually applies to women in their forties, and why does it take so long to appear? In women over 50, the reduction in mortality from screening appears after 5 years, while younger women showed little hint of a benefit until more than 10 years out. David Atkins, an internist and consultant with the U.S. Preventive Services Task Force, points out one possibility: Because some of the women who entered the trials were in their late forties, many of them were still being screened as part of the trial through their early fifties, when mammography is known to be beneficial. "Clearly some of the [late] benefit," he says, "could have come by starting screening at age 50. The question is how much."

The radiologists counter with an analysis of data from the Swedish Two County Trial by researchers including Tabar, who led the trial, and Steve Duffy and Nick Day of the British Medical Research Council. These researchers took the characteristics of the tumors at the time of diagnosis—their size, for instance, and the number of lymph nodes to which they had spread—and used a mathematical model to predict how long the cancer would have remained in what's known as the "preclinical detectable phase," when it could only be detected by mammography, not physical examination, and is still curable. The model suggested that the interval was much shorter in younger women than in older ones.

"In the body of a young woman, the cancer worsens very fast," Tabar explains, "making the preclinical detectable phase so short that if you screen with long intervals, you miss these cancers." Thus, the logic went, the mammograms in the trials—done at intervals as long as 3 years—were missing many of the fast-growing tumors in the 40-something women and catching mostly the slowest growing ones, which is why any benefit was delayed so long. This picture also implied that more frequent screening of these women would yield better results. "It's not surprising that the majority of trials, which screened every 2 years, had comparatively poorer performance in women in their forties than in women older," says Duffy. To achieve a significant reduction of mortality, he says, "you have to screen every year."

The skeptics, however, consider the argument little more than a hypothesis, and they say that detecting cancers at an earlier stage is not the same as saving lives. In younger women, says Fletcher, cancer "seems to go from the breast to the lymph nodes to the whole body very quickly, in which case even screening every 6 months might not make a difference. That's something that's very hard for even clinicians to understand, let alone patients." The classic example, she says, is lung cancer, where randomized trials showed that chest x-rays did not reduce mortality—even if they were done every 4 months.

No consensus

That was where the debate stood going into the consensus conference. Klausner believed that the conference—operating under a set of formal rules designed to ensure that the speakers represent the full range of views and that the panel is unbiased—would be the best way to discuss the benefits and risks of screening in "as neutral a setting as possible." This time, however, the conference system was up against a severe challenge. Kopans, arguing that the NCI was planning to stack the deck against screening, says he had to fight his way onto the planning committee. And many of the researchers involved in the controversy wondered how a group, no matter how erudite, could read 1500 pages of papers in preparation, listen to 32 speakers in less than 2 days, and then come to an informed, intelligent decision overnight. As Kerlikowske put it, "The U.S. Preventive Services Task Force spends several years making a decision. To get a bunch of hotheads in a room for 2 days to do the same thing doesn't make sense to me at all."

The outcome surprised nearly everyone. Even the skeptics seemed to expect that the consensus panel would accept that the trials showed some benefit from screening. But it didn't. "During the first 7 years following mammography, breast cancer mortality is no lower in women [aged 40 to 49] who were assigned to screening than in controls," the panel members wrote in their draft report. "Some studies find lower mortality from breast cancer in screened women after 10 years, but others do not." After weighing the risks from radiation and taking into account false positives and increased diagnosis of DCIS, the panel concluded that "each woman should decide for herself whether to undergo mammography."

The process of reaching that conclusion, says Jeanne Petrek, a breast cancer surgeon at Memorial Sloan-Kettering Hospital, "was like a bunch of blind people examining the elephant and telling each other what they thought." Petrek herself resigned from the panel because she felt it was overstating the risks of mammography and understating the survival benefits. But the one thing everyone agreed upon, says Petrek, was that "we can't recommend [that] every women in their forties have mammography."

"[The panel] reinforced the reason for the conference in the first place," says *JNCI*'s Kramer. "The conference was held because of the uncertainty. And having gone through all the information, they concluded the uncertainty was justified."

While that conclusion disappointed many of the screening advocates, Duke's Rimer, who will preside over the controversy's next official airing at the NCAB, suggests that the discord may have finally peaked. "The fact that it became so inflammatory, that people became so angry, may have made it in some ironic way easier for people to get down to try to search for common ground," she says. "It seemed to sharpen the difference so much that many of the people who represent responsible organizations now feel an obligation to women to try to seek that common ground."

–Gary Taubes