

“Get screened!”

“Find it early!”

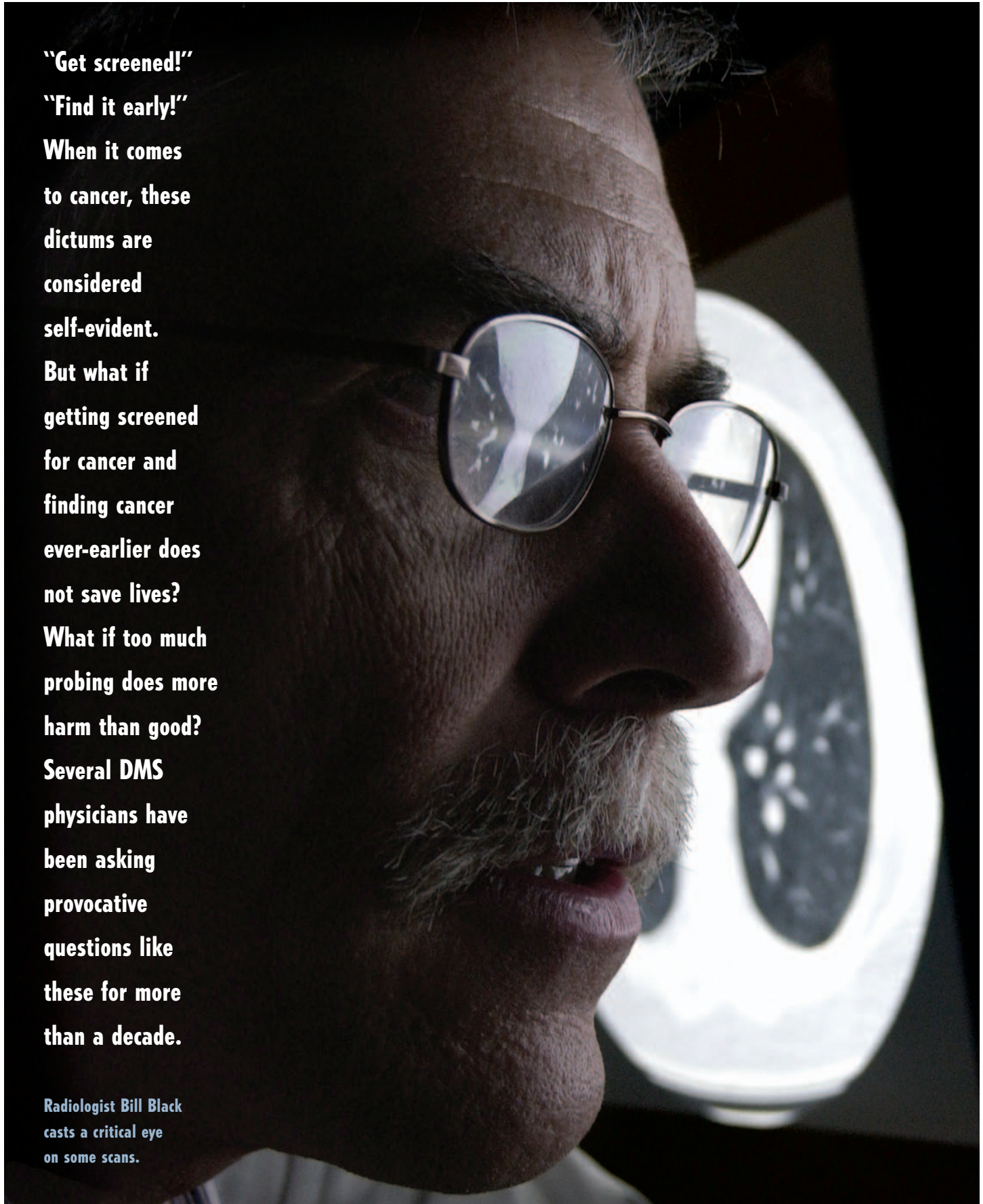
When it comes to cancer, these dictums are considered self-evident.

But what if getting screened for cancer and finding cancer ever-earlier does not save lives?

What if too much probing does more harm than good?

Several DMS physicians have been asking provocative questions like these for more than a decade.

Radiologist Bill Black casts a critical eye on some scans.



The human body is full of imperfections—most of which we simply didn't know about until recently. Now, thanks to sophisticated scanning technology, like computed tomography (CT) and magnetic resonance imaging (MRI), we're able to see ourselves at a level of detail that has never before been possible. "Because we're now able to see every millimeter of the body, we of course find a lot more abnormalities in the body than we ever knew existed," says Dartmouth radiologist William Black, M.D. "What the imaging does is it makes us think, 'Oh, there is this ton of tumors out there and other diseases, so disease must really be increasing in frequency.'" But is it?

All cancers are not created equal. Some grow rapidly and invade other tissue, others grow slowly and remain noninvasive, and some don't grow at

screening is overdiagnosis, but you don't know which people have been overdiagnosed," says Black. "And so a person who has been overdiagnosed will think they've been cured."

These ideas may seem revolutionary, even radical, but they're hardly new. Black and fellow DMS physician-researcher H. Gilbert Welch, M.D., M.P.H., have spent more than a decade exploring the potential harms of looking harder and harder for disease. As early as 1993, in a paper published in the *New England Journal of Medicine*, Black and Welch drew attention to the way that advances in diagnostic imaging can distort physicians' perceptions of disease prevalence and inflate the effectiveness of treatments.

During his 14-year career at Dartmouth, Black has become nationally known and respected as one

ARE WE HUNTING TOO HARD?

By Jennifer Durgin

all or may even recede. Many of the cancers that doctors are finding and treating today, says Black, are what's called "pseudodisease"—tumors that will never cause harm, let alone death. The trouble is that pseudodisease is nearly impossible to identify for sure in an individual who is still living, because the medical community doesn't know enough about some cancers to predict how they will behave over time. So it's safer, they reason, to label a questionable abnormality as "cancer" and to treat it, than it is to risk its growing out of control. Only after an untreated person dies from other causes can a cancer be declared pseudodisease. Only then is it clear that treatment of the cancer would have provided no benefit, only potential harm. In other words, you can't tell an "overdiagnosed," or overtreated, person from a person who has been cured. "One of the biggest downsides to cancer

of the most vocal radiologists when it comes to questioning screening tests. He has coauthored numerous papers on the topic, often with Welch, and served on several National Cancer Institute (NCI) committees, including the 1993 International Workshop on Screening for Breast Cancer, which concluded that there is no proven benefit from mammography for women in their forties.

Although four medical groups—including the American College of Physicians (ACP) and the U.S. Preventive Services Task Force—supported the NCI workshop's conclusion, many other medical and cancer advocacy organizations—including the American Medical Association and the American Cancer Society—strongly opposed the NCI position. Even Congress got involved and, after several hearings in 1994, condemned the workshop's recommendation in a report titled "Misused Science." Eventu-

Jennifer Durgin is DARTMOUTH MEDICINE's senior writer.



Radiologist Bill Black, top, and internist Gil Welch, bottom, have worked closely together for more than a decade—and have put Dartmouth on the map as a place where tough questions get asked about the risks as well as the benefits of routine cancer screening.

ally, after pressure from Congress, the public, and medical and scientific organizations, the NCI reversed its decision, as did the U.S. Preventive Services Task Force. (The ACP has no current recommendation regarding mammography screening but usually supports the Task Force's guidelines on other matters.)

Today, Black is a cochair of the NCI's multicenter National Lung Screening Trial (NLST), as well as the principal investigator for the DHMC-based arm of the trial. The NLST is comparing two ways of detecting lung cancer—spiral computed tomography (CT) and standard chest x-rays—to determine if either can reduce deaths from the disease. “You can't figure out” whether certain screening tests really work, Black maintains, “short of doing a huge experiment” like the NLST, which will randomly assign 50,000 patients to one treatment or the other.

While Welch credits Black with initially sparking his own interest in cancer screening, Welch has made significant contributions to the debate, too, in both the scientific literature and the popular press. Most recently, Welch authored a book titled *Should I Be Tested for*

Cancer? Maybe Not and Here's Why. Early in the book, he takes on the concepts of overdiagnosis and pseudodisease, using prostate cancer as an example.

“The most compelling evidence that pseudodisease is a real problem comes from our national experience with prostate cancer,” Welch writes. Prostate cancer is the second-leading cause of cancer-related death in American men, and over the last 30 years, more and more of it has been found. In 1975, about 100,000 new cases were diagnosed; in 2003, about 220,000. At first glance, one might conclude that prostate cancer is on the rise. However, if a cancer is “really increasing,” says Welch, “you'd expect death rates to rise.”

And that hasn't happened with prostate cancer. The death rate has remained more or less constant, hovering around 30,000 deaths per year in the U.S., with a slight decline in recent years. (See the graph on the facing page for a visual representation of this data.) Some argue that the decline “attests to the usefulness of testing for prostate cancer,” writes Welch, “but it could just as easily reflect better treatment.” Regardless of the small changes in the death rate, Welch believes that most of the new cases represent “nothing more than pseudodisease: disease that would never progress far enough to cause symptoms—or flat-out would never progress at all.

“But what, you might ask, is the harm in finding all this pseudodisease?” Welch writes. “Simply put: unnecessary treatment. Most of the million men whose prostate cancer is found because of superior screening have to undergo some sort of treatment, whether radical surgery or radiation. . . . [And] many experience significant complications: 17% need additional treatment because they have difficulty urinating following surgery; 28% must wear pads because they have the opposite problem—they cannot hold their urine; and more than half are bothered by a loss of sexual function.”

In addition to causing harm from unneeded treatment, overdiagnosis can distort our perceptions of how well certain cancer treatments work, says Black. Because very slow-growing and potentially harmless cancers are relatively easy to control and eliminate, finding and treating more of them make therapies seem more effective. “We see that [such cancers] behave well when we treat them, and we falsely attribute their good behavior to our treatment,” explains Black. “By definition, survival only pertains to people who are diagnosed with the disease. So when you have overdiagnosis, survival is very, very misleading. Survival is going to be very long in people who are overdiagnosed,” as well as in people whose cancers are found very early. For these reasons, Black, Welch, and others at DMS are critical of using survival rates—such as the oft-cited five-year survival rate—as a measurement of the effectiveness of screening. Looking at death rates is a better way of evaluating screening, they argue, but even that approach has problems.

“You can't just look at disease-specific mortality, because we're not sure what causes death in a lot of people,” Black explains. “There are a lot of deaths that are difficult to determine the cause of, and you can bias your results strongly in one direction or the other with disease-specific mortality.”

The best measure of a cancer screening technique, Black and Welch contend, is total deaths—known as all-cause mortality. Yet the most widely accepted measure remains disease-specific mortality. In 2002, Black and Welch wrote a paper for the *Journal of the National*

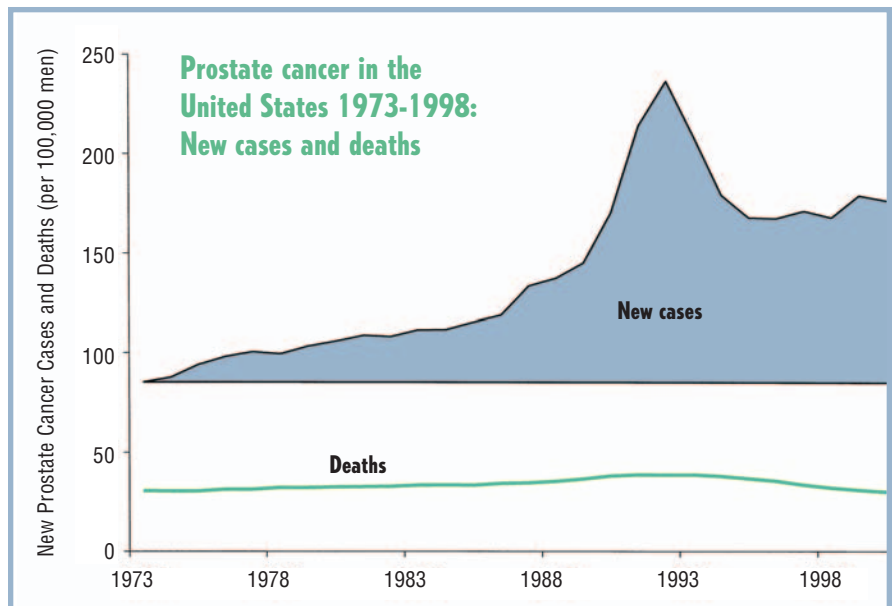
Cancer Institute (JNCI) showing just how shaky a foundation disease-specific mortality can be. They looked at the death rates in several screening trials for breast cancer, colon cancer, and lung cancer, each of which included a screened population and a control, or non-screened, population. Of the 12 trials they reviewed, seven showed major inconsistencies between disease-specific and all-cause mortality. In two there was a significant difference in the magnitude of the result, and in five the disease-specific and all-cause results actually contradicted one another—with one method showing a screening benefit and the other a screening harm.

For example, among the trials Black and Welch examined was the well-known 1989 Swedish Two-County mammography study, which reported that mammography reduces breast-cancer mortality. But when Black and Welch looked at the number of deaths from all causes in the screened group and the non-screened group, they found that there were actually slightly more deaths in the screened population. Likewise, Black and Welch examined a 1993 *New England Journal of Medicine* study that showed a 30% reduction in colon-cancer mortality in people screened with the fecal occult blood test. But the screening benefit disappeared when all-cause mortality was considered.

Such findings beg the question that if screening doesn't reduce all-cause mortality, why be screened? "Think for a minute about why you might want to be screened," writes Welch in his book. "Is it because you want to reduce your chance of a cancer death? Or is it because you want to reduce your chance of death, period?"

Just asking these questions makes Black and Welch controversial. Nationally, researchers and physicians who are asking such questions are "a small minority, maybe one percent of the entire cancer community," says Black. "And they are actually [considered] extreme in their views relative to everybody else." The skeptics of screening are "not usually the loud voices because nobody wants them to be loud," Black continues. At Dartmouth, however, Black and especially Welch have been very outspoken. Since the release of his book, Welch has been quoted in the *Wall Street Journal*, the *New York Times*, *Harper's*, and many other prominent national publications, often as the voice of opposition against ubiquitous cancer screening.

And to the dismay of some advocates of screening at DMS, Welch and Black have earned Dartmouth a reputation as "the center of antagonism for screening," as one visiting National Cancer Institute senior scientist recently put it. But Welch and Black insist that they are not anti-screening.



The graph above shows that although the number of new cases of prostate cancer has risen dramatically in the past 25 years (blue field), the number of deaths has remained nearly constant (green line)—not what one would expect if the actual disease rate were rising.

In addition to causing harm from unneeded treatment, overdiagnosis can distort our perceptions of how well cancer treatments work, says Black. Slow-growing and potentially harmless cancers are relatively easy to eliminate, so finding and treating more of them make therapies seem more effective.

"Gil and I aren't saying that people shouldn't be screened," explains Black, clarifying their perspective. "Our main point is that there are pros and cons and none of them, none of the screening tests, is a slam dunk in terms of producing a lot more good than harm."

"I want to emphasize, there is not a right" answer to the screening question, says Welch. "It's about personal preferences. It's about how you weigh the different outcomes. No one is suggesting that no one should get invited [to be screened], but they should be aware it's a close call."

There is also "a real theology here," he says, "and I understand exactly where it comes from. The idea is so appealing: Earlier is better. Prevention is better than cure. Finding small, bad breast cancers must be good." But the closest Welch can come to the "truth" about cancer screening, he says, is that "the effects of screening are probably mixed in general. Very few are helped. And very few are hurt. And most [screenings] have no effect." He repeats the last point for emphasis: "Most have no effect."

In other words, most people who get screened never develop cancer and never would have. Thus



Lisa Schwartz, top, and Steve Woloshin, bottom (center), have conducted a number of studies assessing how medical messages are communicated and how well they're understood by the public. Schwartz maintains that "scary messages" about screening are too simplistic.

for them, screening confers no benefit. Even advocates of screening like Patricia Carney, Ph.D., codirector of the Cancer Control Program at Dartmouth's Norris Cotton Cancer Center, agree with this point. "It's complicated to study the effectiveness of screening," says Carney, who is also principal investigator of the New Hampshire Mammography Network, a statewide data-collection registry and research group.

"I can't tell you whether you are going to have breast cancer or not. You may have screening all your life," she points out. "You may encounter harm from screening, like unnecessary biopsies, unnecessary workup, or unnecessary radiation exposure, and I can't tell you whether we'll screen you every one to two years for the next 30 years and whether that's going to help you or not."

Where Carney and others at the Cancer Center differ with Welch and Black is on the weight they give to various studies and in their perspectives on what's best for individual patients and the general population.

"The tension that comes up with Gil's book, and with screening in

general," says Carney, "is that as a policy approach you have to consider the entire population," which includes people at high risk and people at no risk and everyone in between. Until cancer researchers and physicians can predict with accuracy who is likely to get cancer, she argues, it makes sense to screen everyone.

And while Welch still questions whether mammography has more benefits than harms, Carney doesn't. "At this point, the evidence is strong enough for me," she says. She places her trust in the U.S. Preventive Services Task Force, which recommends mammography screening every one to two years for women over 40. "I think if the U.S. Preventive Services Task Force has reviewed these studies as carefully as I know they do, and says it is worth doing, then that's strong evidence. . . . That's a panel of people who have looked at every single paper that's out there—not one individual who has picked and chosen what he wanted to present."

Carney believes that it is "incredibly valuable" for Welch to raise the questions he does, but, she adds, "Gil likes to stir things up."

Talking with Welch, it's obvious how passionate he is about the topic of cancer screening. On one recent afternoon in his office, Welch gets visibly perturbed at the mention of a mammography poster that used to be displayed at DHMC. The poster showed the silhouette of a woman with the following words in large, bold-faced type: "The greatest risk is doing nothing." And, in smaller type: "Get the facts about breast cancer."

"So whatever you do, get worried!" exclaims Welch. "That's the message. The most important thing you can do for your health is to worry about it. It's just . . ." he says, not finishing the thought.

"On the one hand," he continues, "I sort of accept it. I've been hearing these messages for years. But it's a little irritating, isn't it? The physicians are out there trying to get people to worry about their health." Welch believes this is the wrong approach. Instead, he'd like physicians "to get people to think positively about their health" and to move away from what he calls "medical correctness" about screening—making patients feel guilty if they choose not to pursue testing. "I object to the emerging mindset that patients should be persuaded, frightened, coerced into undergoing these tests," he writes.

Welch would like the medical and public-health communities to "turn down the rhetoric"—rhetoric like the breast cancer poster and ads that convey the message that screening could save your life. Carney, an advocate of screening, is critical of such messages, too.

"I think a lot of public-health messages have played on the fear thing," she says. "I think that's unfortunate [because] it causes unnecessary anxiety." And, as Welch and his colleagues might add, unwarranted enthusiasm for screening.

Two of Welch's research collaborators, Lisa Schwartz, M.D., and Steven Woloshin, M.D., have been looking at how medical messages are communicated to the public for the past 10 years and have found that Americans are extremely pro-screening. In a 2004 study that was published in the *Journal of the American Medical Association (JAMA)*, Schwartz and Woloshin showed that approximately 87% of adults believe routine cancer screening is almost always a good idea and about 74% believe that finding cancer early saves lives. (See the graph on the facing page for a visual representation of some of the data that appeared in this paper.)

"Less than one third believe that there will be a time when they

will stop undergoing routine screening,” the authors wrote. And “a substantial portion believe that an 80-year-old who chose not to be tested was irresponsible: ranging from 41% with regard to mammography to 32% for colonoscopy.”

Schwartz, Woloshin, and Welch, along with several of their colleagues who work at the VA Medical Center in White River Junction, Vt., make up the VA Outcomes Group, a collaboration of physician-researchers who are concerned with what they perceive as excesses in American medicine. “We question the assumption that patients always stand to gain from having more health care,” the group states on its Web site (www.vaoutcomes.org). “We are concerned about advertising and other messages that exaggerate the benefit of health care and minimize the harm (or ignore it entirely). And we are troubled by the increasing enthusiasm for seeking diagnoses in the well and initiating interventions for those identified as ‘sick.’”

“We have given people persuasive, scary messages for a long time to tell them that earlier is better,” explains Schwartz, who codirects the VA Outcomes Group with Welch. “It’s not just cancer screening. It happens with heart disease, osteoporosis. With all diseases there is this push to go earlier.”

Many of the public messages about screening communicate two things, says Schwartz: “It’s really dangerous if you don’t” get screened and “This is what a good person does.” Framing the decision to be screened or not in a moral context makes it “harder for people to make an informed decision,” she maintains. “How much are we destroying whatever sense of health and well-being people have by continually scaring them?” she wonders.

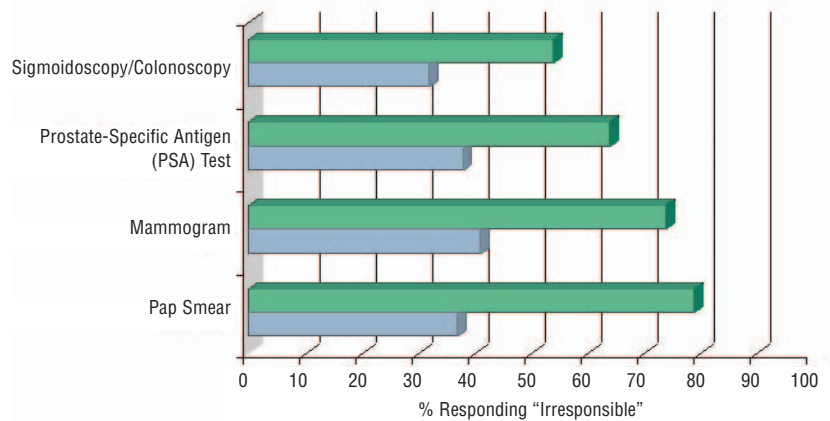
But Americans have been scared of cancer for a long time—long before cancer screening became popular and celebrities like Katie Couric began pushing various tests (see page 4 in this issue for a related story). Consider, for instance, the Neil Simon play *Brighton Beach Memoirs*, set in Brooklyn in 1937. The play centers on a working-class Jewish family, with the main character a baseball-crazed teenager named Eugene. Early in act one, Eugene addresses the audience:

Eugene: [To the audience.] If my mother knew I was writing all this down, she would stuff me like one of her chickens. I’d better explain what she meant by Aunt Blanche’s “situation.” You see, her husband, Uncle Dave, died six years ago from [he looks around] this thing—they never say the word.

What the public thinks about screening . . .

Survey participants were asked: “Do you feel that a person in average health who did not have routine [test] was irresponsible?”

Age of Person
Not Being Screened
■ 55 years old
■ 80 years old



The graph above and the box below show the discrepancy between what the public believes is the responsible stance regarding cancer screenings, and what the experts recommend. That sense of obligation is not necessarily a good thing, according to several DMS faculty.

Many public messages about screening communicate two things, says Schwartz: “It’s really dangerous if you don’t” get screened and “this is what a good person does.” Framing the decision to be screened or not in a moral context makes it “harder for people to make an informed decision,” she maintains.

. . . and what the experts recommend

Colorectal Screening: The U.S. Preventive Services Task Force (USPSTF) strongly recommends colon cancer screening for men and women age 50 and older, but found insufficient data to recommend a specific test.

PSA Test: The USPSTF concluded that evidence is insufficient to recommend for or against routine screening for prostate cancer using a prostate specific antigen (PSA) test (or a digital rectal examination).

Mammogram: The USPSTF recommends screening mammography, with or without clinical breast examination, every one to two years for women age 40 and older.

Pap Smear: The USPSTF strongly recommends routine Pap smears through age 65 for women who have been sexually active and who have a cervix.



Pathologist Wendy Wells, top, says that deciding if a tissue sample is cancerous or not sometimes isn't clear-cut. But it *is* clear from the number of paper cups on their conference table, bottom, that meetings of the outcomes group to explore such issues are very collaborative.

They always whisper it. It was [he whispers]—cancer! I think they're afraid if they said it out loud, God would say, "I HEARD THAT! YOU SAID THE DREAD DISEASE! [He points his finger down.] JUST FOR THAT, I SMITE YOU DOWN WITH IT!!"

Yet even today, despite the fact that researchers have a much better understanding of the broad spectrum of cellular abnormalities that are considered cancer, many such attitudes about the disease persist. Black explains why: "Our knowledge about the natural history of cancer is based largely on historical cases." Even though "a tiny little nodule in a CT scan" goes by the same name as the large tumors that were diagnosed in previous centuries, "people often assume that all these tiny abnormalities . . . will lead to the serious outcomes we observed in the past."

Wendy Wells, M.B.B.S, a British-trained Dartmouth pathologist who specializes in breast cancer diagnosis, agrees that a lot of confusion and anxiety surround the C-word. "As Gil would say, there's cancer and there's cancer," Wells puts it. She explains that labeling abnormal cells as cancer or not can be a close call—a fact that many pa-

tients are not aware of. In some cases, "if you understand how close you are to either getting the C-word or not getting the C-word, it's actually quite terrifying," she adds.

"We've hit the morphology wall for all those descriptive things that pathologists learn about, that define normal or hyperplastic [that is, abnormal but not definitely cancerous] or cancer," explains Wells. "We're very good at that, but we can't predict behavior. So the race now is to look for other prognostic indicators . . . not only at the protein and molecular level but also at the gene level."

Pathologists diagnose cancer based on the characteristics of cells—known as cytology—as well as on the way the cells arrange themselves—known as architecture. By analyzing the cytology and architecture of cells in a tissue sample, pathologists decide whether the cells are cancerous or not. Next, if they are cancerous, the pathologists decide whether they are very abnormal (high-grade) or slightly abnormal (low-grade) and whether they have invaded the surrounding tissue (invasive) or not (noninvasive).

Beyond these rather general categories, however, are more complicated classification systems. And as Wells discovered through research that she and Patty Carney did with the New Hampshire Mammography Network, pathologists often don't agree on how to classify a given tissue sample. In other words, one pathologist might diagnose the cells from a breast biopsy as atypical ductal hyperplasia, which is *not* considered cancer, while another pathologist might label the very same sample as low-grade ductal carcinoma *in situ* (DCIS), which *is* considered to be cancer.

"If you get the diagnosis of atypia, nothing else is done—no further excision, no radiation. If you just nudge the criteria and make it a low-grade DCIS, suddenly: Excision! Radiation! Everything full hog! And yet, the criteria are so nebulous," says Wells. She adds, however, that only about 2% of the breast biopsies she evaluates fall into this difficult-to-classify category. And when she can't reach a conclusion on a sample, Wells will consult with her colleagues and sometimes with experts elsewhere. If a consensus on a diagnosis still can't be made, the patient is notified. This is "a huge onus on the patient because you're literally saying, 'We really can't make up our mind,'" she says. Until pathologists know more about how certain cancer cells behave over time, these uncertainties will likely remain.

Wells says she is neither an advocate for nor an opponent of screening—though, at age 45, she does get regular mammograms—because not enough is known about how certain cancers behave. "I don't think we have enough data," says Wells. "Why Patty's work is so interesting is that she's doing exactly that. She's trying to get all the data" through the mammography network. But, she adds, "I see where Gil is coming from. I do think we are harming some patients."

All sides seem to agree on that point: some patients are being harmed by screening. In his book, Welch offers some advice to those who are concerned about being overdiagnosed. "Probably the best way to minimize the harmful effects of screening is to be willing to take some time" with the small, questionable abnormalities, he writes. "Even when 'cancer' is agreed on, it may make sense to wait and be sure the cancer is really growing." Watchful waiting, as the wait-and-see strategy is called, can be hard for patients and doctors alike because in some ways, just like screening, watchful waiting is a gamble.

One DHMC patient who opted for watchful waiting believes he's

won the “bet.” Robert Aliber, Ph.D., an emeritus professor of international economics and finance at the University of Chicago who now lives in Hanover, was diagnosed with an early-stage, moderately aggressive prostate cancer in 2001, when he was in his early seventies. Since Aliber had been getting a prostate-specific antigen (PSA) test every year since 1996, his doctors had a good baseline from which to judge his unusual scores. They recommended that he have a biopsy because his PSA scores were variable and increasing. When one of the 10 snips from the biopsy came back positive for cancer, Aliber was faced with four treatment options—surgery to remove the prostate, external radiation, internal radiation, or watchful waiting—or some combination of the four.

“Economists deal with data, and they often deal with data under uncertainty,” says Aliber. To him, choosing a treatment course was similar to making an investment. “I was trying to figure out what the aftereffects were of different treatments—in a sense, that’s the cost.” The benefit, of course, would be a potentially longer life.

“My choice is straightforward,” Aliber wrote in a February 2002 memo to his doctors. “I can ‘buy’ a treatment in the next few months, or I can adopt a policy of watchful waiting and ‘buy’ a treatment at some future date if the PSA or other data indicate that the cancer has become more aggressive.

“The choice between whether to ‘buy’ a prostate-cancer treatment at this time or perhaps at a later time can be modeled like an investment decision,” Aliber continued. “The traditional investment decision is ‘which investment has the higher present value (or, alternatively, the higher rate of return)?’” In other words, was it more valuable to him to maintain his quality of life in his seventies in exchange for potentially shortening his life? He decided it was.

Four years after the diagnosis, Aliber’s PSA levels have stabilized, a sign that the cancer probably isn’t growing. “It would have been costly without any benefit if I’d had the treatment three years ago,” he says. So Aliber appears to be beating his cancer by doing little more than watching and waiting.

In the end, the decision to be screened, or not screened, for a particular cancer—as well as what course of treatment, or nontreatment, to pursue if cancer is detected—should be made in the context of each individual’s own preferences. Welch and Black and others of like mind hope that most physicians would agree with that premise.

Where much of the controversy centers, however, is on the population and policy-making levels. Should the medical community be enthusiastical-



Hanover resident Robert Aliber, an economist, chose “watchful waiting” when he had a positive biopsy after his prostate-specific antigen test results began to rise. He approached the decision as if he were making an investment—weighing the costs against the gains.

Four years after the diagnosis, Aliber’s PSA levels have stabilized, a sign that the cancer probably isn’t growing. “It would have been costly without any benefit if I’d had the treatment three years ago,” he says. So Aliber appears to be beating his cancer by doing little more than watching and waiting.

ly endorsing cancer screening? Is saving a few lives worth harming many? There’s a bigger question at work here, too. What should be the medical community’s philosophy about taking action in the absence of conclusive data?

For physicians, like Aliber’s DHMC urologist, John Seigne, M.B., the tendency is to want to do *something*. “I’m a urologic oncologist,” Seigne explains. “Sitting in front of me every day are people who have prostate cancer, people who are dying of prostate cancer, people who I can’t cure. . . . So I get a completely different perspective from somebody who is” looking at the big picture. Somebody like Black. Or Welch.

Yet Seigne, as does Carney, values the questions that Black, Welch, and others have been raising. “What do you do when you don’t know the answer to a question? Do you do nothing?” Seigne reflects. “One of the tenets of the Hippocratic oath is [first] do no harm—*primum non nocere*. So do you do nothing? Or do you try and synthesize the best available information that you have and do something with it? . . . I don’t know enough about philosophy to answer that.” Does anyone? ■