The New York Times
nytimes.com

March 13, 2007

ESSAY

How Two Studies on Cancer Screening Led to Two Results

Of all the forms of cancer, lung cancer is by far the deadliest. So doctors have long hoped to come up with a screening test that would find it early, before it can grow and become untreatable.

Last fall, The New England Journal of Medicine published a study concluding that spiral CT screening (a kind of three-dimensional chest X-ray) would make most lung cancers curable. It sounded like wonderful news. For proponents of screening, it was a call to action: the Lung Cancer Alliance is starting an advertising campaign featuring sports celebrities trying to persuade you to make the "right call" and get screened.

But just last week The Journal of the American Medical Association published a study concluding that spiral CT screening is not only ineffective, but may actually be harmful, prompting unnecessary surgery that carries risks of its own.

How could these two studies — in the country's two most prestigious medical journals — arrive at diametrically opposite conclusions? An answer requires a clear understanding of the goal of cancer screening.

That goal is to save lives — or, in scientific terms, reduce mortality. Simply finding cancer early is not enough.

Finding cancer early saves lives only if two conditions are met: the cancers detected are the ones that kill people; and early treatment prevents these deaths.

It is not enough to increase survival. While that may seem to be the mirror image of mortality, it can be a terribly misleading measure of the value of screening.

In the 1970s and '80s, there was great interest in screening smokers for lung cancer using conventional chest X-rays. The question was seen as so important that it was examined using the gold standard of medical studies, a randomized trial. Half the participants were randomly selected to receive regular chest X-rays; half did not and served as the control group.

Three such randomized trials were conducted, and all three showed that screening did not reduce mortality. In fact, two reported slightly higher death rates in the group receiving chest X-rays.

The most famous of these trials, at the Mayo Clinic, showed how misleading survival can be. Although the 10-year survival rate doubled with screening, mortality was not reduced; indeed, screening may have

increased it. The Mayo trial also showed that more than a decade after screening was stopped, there were still more cancers in the screened group. This shouldn't happen: in two large randomly selected groups, there should be the same number of cancers in both. The chest X-rays must have detected some lung cancers that were never destined to cause symptoms or death — a phenomenon known as overdiagnosis.

This phenomenon challenges our conventional view of cancer as an inexorably progressive disease. Research in screening has demonstrated that what pathologists call cancer encompasses a broad spectrum of disorders: some cancers rapidly progress to death, some do so more slowly, and some don't progress at all (or may even regress).

Overdiagnosis is even more of a concern for spiral CT, because it can detect far more abnormalities than chest X-rays. In fact, a screening program in Japan found about 10 times as many lung cancers with spiral CT as had been found in the same population using chest X-rays. More remarkably, the chance of having lung cancer detected by spiral CT was almost the same in nonsmokers and smokers.

This flies in the face of everything we know about lung cancer and smoking — 50 years of research showing that smokers are 10 to 20 times as likely as nonsmokers to die from lung cancer. This is powerful evidence that spiral CT detects some lung cancers that will never affect patients.

Because all lung cancer patients get treated, overdiagnosis means some people receive treatment that can't help them (because they do not need it) and can only cause harm. Most patients given diagnoses of early lung cancer undergo surgery to remove part of a lung, a major operation from which about 5 percent die within a month.

With this background, let's look at the two recent studies on screening.

The New England Journal study reported screening about 31,000 people with spiral CTs and finding 484 with lung cancer. These patients had a 10-year survival of 80 percent — compared with 10 percent for current lung cancer patients in the United States. The JAMA study reported screening about 3,200 people and finding 144 with lung cancer. (The detection rate was higher because this study had older patients and longer follow-up.) Of 3,200 people, 38 died from lung cancer — the same mortality rate expected for people of similar age and smoking history in the absence of screening.

In short, The New England Journal reported increased survival; JAMA reported no difference in mortality.

Most of us interpret "increased survival" to mean fewer deaths. But it does not, because survival is subject to two powerful distortions.

The first is called lead-time bias. Simply advancing the time of diagnosis (as with CT screening) will always increase survival.

Imagine two patients with lung cancer. Even if both die at age 70, a patient with cancer diagnosed by

spiral CT screening at age 59 has a longer survival than one with cancer diagnosed because of symptoms (cough, weight loss and so on) at age 67. The first patient survives 11 years; the second 3 years. But both died at the same age. Survival is increased, but mortality is the same.

A second source of distortion results from overdiagnosis, when screening finds cancers that were never destined to progress and cause death. Overdiagnosis bias can also drastically inflate survival statistics, even if mortality is unchanged.

To understand why, you need to understand the definition of the two statistics. Both are fractions. Survival is calculated over a fixed period, for example 5 or 10 years.

Overdiagnosis inflates both the numerator of the survival statistic (number alive at a specified time) and the denominator (number of diagnoses). For the mortality statistic, overdiagnosis has no effect on the numerator (number of deaths) or the denominator (number studied). Perhaps the easiest way to understand this is to imagine if we told all the people in the country that they had lung cancer today: lung cancer mortality would be unchanged, but lung cancer survival would skyrocket.

The goal of lung cancer screening is to reduce mortality — to save lives. Because the New England Journal study examines only survival, it cannot tell us whether any lives are saved. Because the JAMA study examines mortality, it is the more valid study. It also corroborates the Mayo trial finding that a significantly increased survival rate can coexist with no difference in mortality.

The JAMA study also highlights the tradeoffs involved in lung cancer screening. The findings show that compared with no screening, if 1,000 people are screened over five years there would be 48 more lung cancer diagnoses, 46 more lung cancer operations (which would be expected to cause 2 deaths) and no lung cancer deaths prevented. The study data are consistent with as many as eight deaths prevented by screening, or eight extra deaths caused by it.

But neither study is definitive, because neither was a randomized trial. And both required assumptions. Given the potential benefit (so many people die from lung cancer) and the potential harms (some die from treatments), no one should have to assume anything.

Luckily, two randomized trials are under way — one a Dutch-Belgian collaboration, the other sponsored by the National Cancer Institute. Recent experience, notably with hormone replacement in postmenopausal women, has demonstrated how presuming benefits in the absence of randomized trials can cause real harm. To avoid repeating these mistakes, we should not screen for lung cancer unless the trials demonstrate a reduction in mortality.

This essay is by H. Gilbert Welch, Steven Woloshin and Lisa M. Schwartz. Dr. Welch is the author of "Should I Be Tested for Cancer? Maybe Not and Here's Why" (University of California Press). He, Dr. Schwartz and Dr. Woloshin are senior research associates at the VA Outcomes Group in White River Junction, Vt.