Science for Public Consumption

Presenters:
Lisa Schwartz
VA Outcomes Group
White River Junction, Vermont, and
Center for the Evaluative Clinical Sciences
Dartmouth Medical School
Hanover, New Hampshire

Steven Woloshin
VA Outcomes Group
White River Junction, Vermont, and
Center for the Evaluative Clinical Sciences
Dartmouth Medical School
Hanover, New Hampshire

Reporter:
Devora Mitrany
AdvancePCS
Scottsdale, Arizona

Lisa Schwartz and Steven Woloshin, a lively team of physician-scientists, presented case studies to show how statistics—although technically accurate—can be used to exaggerate research outcomes. The result, for the public, is often unnecessary anxiety and overexposure to medical care. Their prescription: "Take the mystery out of the numbers with straightforward data presentation."

Race and Sex Discrimination—Really?

"Blacks and women with chest pain are 40% less likely than whites or men to be referred for cardiac catheterization." The first question is "40% of *what*?" The authors presented odds ratios, which exaggerate relative risk when base rates are over 20%—and in this case, they were 91% for whites and 85% for blacks. Risk ratios would have been less shocking but more accurate.

The real message was that blacks or women were referred 7% less often than men or whites. Hardly sensational enough to seize the spotlight on "Nightline" or in the *New York Times*, as actually happened.

In reality, only black women had a lower referral rate—information that was obscured by how the methods were described in the original article, "Analysis of Race-Sex Interactions". Woloshin and Schwartz suggested a "black-box" warning for analyses described in this way: "999 of 1000 readers may be intimidated"!

WHI Findings—Risky Business

The hormone-therapy findings of the Women's Health Initiative have spawned intense mass-media attention. The problem with many of these mainstream media reports, Schwartz and Woloshin contend, is a lack of base rates. A 1.41 hazard ratio translates into a 41% increased relative risk. What we don't know is 41% higher than *what*.

Relative-risk reduction appears more impressive than the corresponding absolute-risk reduction, especially when the risk is small. The *New York Times* tried to circumvent the problem this way: "The data indicate that if 10,000 women take the drugs for a year, 8 more will develop invasive breast cancer. . . ." But again, the question is "eight more than *what*?"

Expressing results as absolute event rates is better: Of 1000 postmenopausal women observed over a 5-year period, 15 of those taking placebo will develop breast cancer, compared with 20 of those taking hormone therapy. Readers can then clearly judge the magnitude of the risk.

Overinterpreting Early Detection

Earlier detection of cancer doesn't necessarily mean longer survival, say Woloshin and Schwartz. In a study comparing computed-tomographic (CT) scans and chest x-ray pictures for lung-cancer detection, CT scans found 4 times more cancers, and at an earlier stage, than x-ray pictures did. But advancing the time of diagnosis is not proof that screening works.

For example, if 10 people are diagnosed with lung cancer at age 65 and all are dead by age 69, their 5-year survival rate is zero. If those same 10 people were diagnosed at age 60, their 5-year survival rate would be 100%!

Schwartz and Woloshin caution that early detection is not the same as a better prognosis. Proof of reduced mortality, or even proof that screening did more good than harm, can come only from a randomized trial of screening.

New Alzheimer's Drug: How Significant?

A clinical trial to study a new drug to treat Alzheimer disease reported statistically significant results. What wasn't so clear was their *clinical* significance.

The article lacked a description of the clinical tests used to score patients' functionality, so the results—an increase of 0.3 point on a 7-point scale—are meaningless. Unless findings can be related to patients' ability to function, the reader cannot gauge the magnitude of improvement.

Treatment Cost-effectiveness: Putting the Cart before the Horse

Sometimes researchers create cost-effectiveness models before treatment effectiveness has been established. Woloshin and Schwartz suggest that journals consider not publishing such studies. If the findings are published, a caveat may lessen their prominence: "The effectiveness of this procedure [or test] has not been demonstrated. Therefore, the question of cost-effectiveness may be premature."