

1.15% APY

A HIGH-YIELD SAVINGS
ACCOUNT FROM
AMERICAN EXPRESS

LEARN MORE NOW



PERSONAL SAVINGS
from American Express

Accounts offered by American Express Bank, FDIC. MEMBER FDIC

Dow Jones Reprints: This copy is for your personal, non-commercial use only. To order presentation-ready copies for distribution to your colleagues, clients or customers, use the Order Reprints tool at the bottom of any article or visit www.djreprints.com

See a sample reprint in PDF format.

Order a reprint of this article now

THE WALL STREET JOURNAL
WSJ.com

THE NUMBERS GUY | APRIL 2, 2011

Making a Stat Less Significant

By CARL BIALIK



A group of mathematicians has been trying for years to have a core statistical concept debunked. Now, the Supreme Court might have done it for them.

Last month, the court considered a case brought by investors in Matrixx Initiatives Inc. They alleged the company failed to disclose material information by neglecting to reveal it had received reports an over-the-counter medicine, Zicam Cold Remedy, caused a loss of sense of smell.

When those reports came to light, the company's stock fell. Eventually, the Food and Drug Administration warned consumers not to use certain Zicam products.

The company argued in court that the initial reports of the possible side effect didn't rise to the level of "statistical significance," and therefore didn't need to be disclosed.

Last week, the Court rejected that argument, ruling unanimously that the case against Matrixx pending in a lower court could proceed. In their opinion, the justices said companies can't only rely on statistical significance when deciding what they need to disclose to investors.

Numbers Guy Blog

A Statistical Test Gets Its Closeup

Amen, say several statisticians who have long argued that the concept of statistical significance has unjustly overtaken other barometers used to determine which experimental results are valid and warrant public distribution. "Statistical significance

doesn't tell you everything about the truth of the hypothesis you're exploring," says Steven Goodman, an epidemiologist and biostatistician at the Johns Hopkins Bloomberg School of Public Health.

A point on which most statisticians agree is that statistical significance is difficult to explain.

Here, with guidance from statisticians, is an attempt at an explanation, by way of example: Suppose an experiment looks at whether a drug induces a certain side effect in people who use it.

The results indicate that, on average, people who use the drug experience that effect more than those who don't use it, all else being equal.

While a greater percentage of people taking the drug experienced the symptoms than those in a control group, what if those increased symptoms were a result of bad luck, or a statistical fluke, rather than something caused by the drug? If they were, what, then, would be the probability that the experiment would have produced a false positive as large as the one found? If the probability of a false positive is less than a certain level—generally 5%—then the link shown between the drug and the side effect is deemed to be statistically significant.

But skeptics of statistical significance note that the side effect can pass the statistical test, without having actual clinical significance. In other words, the effect could still be rare enough to not pose a major threat.

To determine whether a medical side effect is significant in an experiment requires knowing that every incidence of that side effect is being reported. Researchers can feel confident that is happening in a controlled clinical trial of a drug, but they can't be sure when a drug is being sold to the general public, as was the case with Zicam.

Because of the pitfalls with statistical significance, when the FDA analyzes the safety of a drug or device, "we certainly don't make decisions based on a single number or a single statistical test," says William Maisel, the chief scientist for the agency's Center for Devices and Radiological Health. Just because a side effect isn't statistically significant "doesn't mean there isn't an issue there," Dr. Maisel says.

But some critics say the high court's ruling undermines one of the few standards researchers have for assessing risks.

Erik Olson, a partner with law firm Morrison & Foerster in San Francisco who filed a brief supporting Matrixx on behalf of BayBio, a California biotechnology industry group, says the high court's ruling risks leaving companies without a clear benchmark for deciding when they need to disclose adverse events. A Matrixx spokesman didn't return calls seeking comment.

While many statisticians agree that significance tests often are misused, they are split on whether results that aren't statistically significant might sometimes be worth reporting anyway. Stephen Ziliak, an economist at Roosevelt University in Chicago, argues that a bigger effect produced in a study with a big margin of error is more impressive than a smaller effect that was measured more precisely.

Prof. Ziliak applies his belief in the misuse of statistical significance to his own life, with occasional adverse effects. When his doctor advised him to take a cholesterol-reducing drug, she told him it was significantly better than its competitor. He asked her what the power of the statistical test was, meaning the likelihood that if there were a real difference between the pills, researchers would find it. "She stopped taking my requests for appointments," Prof. Ziliak says.

—Email numbersguy@wsj.com.

Copyright 2011 Dow Jones & Company, Inc. All Rights Reserved

This copy is for your personal, non-commercial use only. Distribution and use of this material are governed by our [Subscriber Agreement](#) and by copyright law. For non-personal use or to order multiple copies, please contact Dow Jones Reprints at 1-800-843-0008 or visit www.djreprints.com