

Why so much breast cancer?

How many people have I known who have had breast cancer? Without racking my brains too hard, I was able to count nearly ten, far more than for any other serious illness. There seems to be a lot of it about, something that you especially notice once someone dear to you has been affected. Cancer, of course, is not just one thing, but hundreds of different diseases, each with its own risk factors (and risks). Breast cancer itself can be any of several different cancers.

One clue concerning the prevalence of breast cancer(s) is that it varies so much from country to country, which raises the possibility that diet may be relevant. Cancer has been described as a disease of the developed world; if that is true, then the reason would not have to be that the developed world has more risk factors for cancer. And Japan has a famously low incidence of breast cancer.

So it is natural to wonder whether some differences between the Japanese and Western diets are associated with the differences in the incidence of breast cancer. Lots of possibilities there. The Japanese consumption of milk is known to be a fraction of that in countries like USA, UK, France, Australia and New Zealand, and it has been suggested that milk consumption may be a risk factor for breast cancer.

But is it? With the help of a computer and internet access, I did a little informal research. First, I found a webpage with per-capita milk consumption by country. Unfortunately there were just 26 countries listed, though they were well-assorted. Next I found a webpage with breast cancer death rates by country. Then, for the 26 countries, I entered the figures from both webpages into a spreadsheet. Finally I used one of the spreadsheet's statistical functions to calculate the correlation coefficient (known as *rho*) for the two columns of figures. This gave a number that indicates the association between milk consumption and breast cancer deaths for the 26 countries.

In case your statistics is a little rusty, let me remind you about interpreting the value of *rho*. A positive value would indicate a more-milk-more-deaths trend. A negative value would indicate a more-milk-fewer-deaths trend. Zero would indicate a lack of association (one way or the other) between milk consumption and deaths from breast cancer. The further away from zero the value of *rho* is, the stronger the association between milk consumption and breast cancer deaths, the range of possible values being from -1 to +1. Elegant, don't you think?

I can now reveal that the value of *rho* for my spreadsheet is ... (dramatic pause) ... just under +0.7. So that's a strong positive correlation between milk consumption and deaths from breast cancer for those 26 countries. (If we had individual values of *rho* for the several different cancers that go under the breast-cancer heading, then most probably not all of them would have *rho* about 0.7, and so for at least one breast-cancer disease the value of *rho* would be even more than 0.7.)

I admit that it would have been better to use the number of cases of breast cancer rather than the number of *deaths*, since the nature of health care is not the same in every country. My guess, though, is that this would have given an even bigger value of rho, as the countries with the better health care tend to be those with higher milk consumption *and* higher breast cancer death rates.

So what does this 0.7 prove? What, if anything, does it tell us about cause and effect? Broadly, there are these five alternatives.

1. The calculation used unreliable or inappropriate data, and so the 0.7 tells us nothing.
2. Chance played a big part in giving the value 0.7. If for instance I had used figures for different years or for a different set of countries, then the result could have been a lot different.
3. Milk consumption is a risk factor for breast cancer.
4. Breast cancer is a risk factor for milk consumption. Don't laugh. The possibility needs to be mentioned. Back when the correlation between smoking and lung cancer was becoming widely known, some of those nice friendly people in the tobacco industry pointed out that the explanation could be that people who were going to develop lung cancer were more likely to take up smoking.
5. There is no *direct* risk factor linking milk consumption and breast cancer, but one or more indirect associations via other factor(s). For instance, it could be that there is some technique that can be used in milk production and that happens to increase the incidence of breast cancer, and that this technique is more common in countries with high milk consumption. In such a case, the milk itself might not be implicated – it could be, say, that some practice in dairy agriculture put something into the water supply. Or the risk factor could be something commonly consumed along with dairy produce. Or the risk factor could be the *combination* of milk consumption and something else, not milk consumption on its own, nor the something-else on its own.

So, a five-horse race, and caution is needed. Occasionally a race will be won by a horse with a limp. If it does turn out that milk consumption is a risk factor for breast cancer, then there will be dilemmas. Milk has health benefits too. Being alive is a risky business.