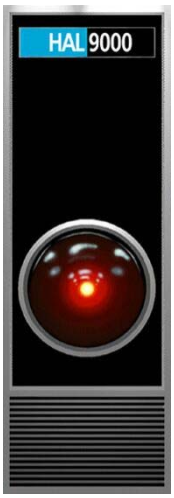

Does Cancer Cause Smoking?

Even if the correlation of two variables is due to a causal relationship, the correlation itself cannot tell us what causes what.

Sir Ronald Aylmer Fisher (1890-1962) was one of the greatest statisticians of the 20th century. He virtually invented the statistics method known as Analysis of Variance, and in his honor the key statistic of that analysis is named the "F" statistic.

Fisher recognized the correlation between smoking and cancer; greater amounts of smoking are associated with higher incidence of cancer. He believed that there was a causal relationship underlying this correlation, and he wrote the following:

Is it possible, then, that lung cancer ¾ that is to say, the precancerous condition which must exist and is known to exist for years in those who are going to show overt lung cancer ¾ is one of the causes of smoking cigarettes? I don't think it can be excluded...the pre-cancerous condition is one involving a certain amount of slight chronic inflammation...

A slight cause of irritation...is commonly accompanied by pulling out a cigarette, and getting a little compensation for life's minor ills in that way. And so anyone suffering from chronic inflammation in part of the body...is not unlikely to be associated with smoking more frequently, or smoking rather than not smoking.

Fisher was, of course, entirely correct statistically. A high correlation between smoking and cancer could be due to cancer causing smoking or to something else causing both. (Fisher also suggested that a genetic predisposition to cancer might also predispose one to smoke.)

Ironically, the proof that smoking indeed is the cause of many cancers came from experiments conducted following the principles of experiment design and analysis that Fisher himself developed.