STATISTICS AND EVIDENCE-BASED MEDICINE

STATISTICAL SIGNIFICANCE

The medical community is tasked with making decisions about treatment choices and recommendations based on evidence-based research. Statisticians play a key role in helping determine the relationship between exposures and outcomes. The significance of these relationships have a huge impact on our daily life.

EVIDENCE-BASED DECISIONS

In the scientific community, we use studies to learn about unknown phenomena that will ultimately inform our decisions. These studies come in many different forms. For example, you could select a representative population and randomly assign them to a treatment in a very controlled environment and see what happens (these are called randomized controlled trials). Alternatively, you could take data that already exists and explore what has already happened, accounting for all of the important factors you would normally control for in a randomized controlled trial (these are called observational studies).

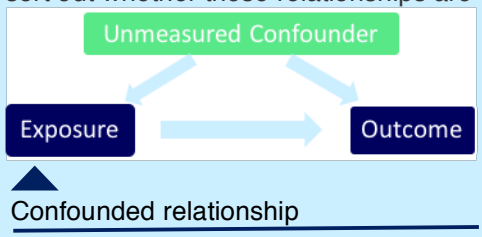
There are many reasons why a researcher may choose the latter design over the former, such as limited resources or ethical reasons. For example, there never has been, and there never will be a randomized controlled trial to assess whether smoking causes lung cancer, however the empirical evidence points to a strong association.

IDENTIFYING RELATIONSHIPS

Since these observational studies are not performed in controlled environment, we need to be sure that we are accounting for all important factors that may distort the relationship between the exposure and the outcome. In the case of smoking and lung cancer, the great statistician Ronald Fisher, argued that there may be an unmeasured factor, perhaps a genetic predisposition to smoke, that is distorting the relationship between smoking and lung cancer.[1] Another great statistician, Jerome Corn field countered that even if such a gene existed and caused lung cancer with near certainty, it would need to be much more prevalent among smokers to explain the observed relationship between smoking and lung cancer. He quantified this by showing the 9 times greater risk of lung cancer seen in smokers would require this mystery gene to be 9 times as great among smokers than non-smokers.[2] This was deemed extremely unlikely, and the discussion was put to rest.

CAUSAL RELATIONSHIPS IN MEDICINE

In the medical field, we are often interested in the relationship between various exposures and outcomes. Some examples of this are the relationship between smoking and lung cancer, UV radiation exposure and skin cancer, and exercise and heart disease. Statisticians can help sort out whether these relationships are *causal* (the exposure is causing the outcome), *confounded* (the relationship is distorted by the presence of another factor), or an *illusion* (there is not really a relationship between the exposure and the outcome, we are just seeing one by chance). It is often difficult to distinguish between *causal* and *confounded* relationships. Statisticians can help by quantifying how important an outside factor would have to be in order to distort the observed relationship between an exposure and an outcome. For example, this is how the association between cigarette smoking and lung cancer was deemed to causal and very unlikely to be due confounding from an unknown factor.



Confounded relationship

Jerome Cornfield, circa 1950

STATISTICS & MEDICINE

Using advanced techniques, statisticians can help to sort out whether observed relationships are causal, confounded, or merely illusions, helping to inform decisions made by the medical community.

Lucy D'Agostino McGowan &
Robert Alan Greevy Jr.

1. Fisher, R. A. (1958), *Nature*, 182, 108.
2. Cornfield, J., et al (1959), Smoking and Lung Cancer: Recent Evidence and a Discussion of Some Questions, *Journal of the National Cancer Institute,* 22:173-203.

Copyright: *Photograph of Jerome Cornfield appears courtesy of IMS. Copyright IMS.*