# 1.1: Data Analysis

Much of experimental science comes down to measuring changes. Does one medicine work better than another? Do cells with one version of a gene synthesize more of an enzyme than cells with another version? Does one kind of signal processing algorithm detect pulsars better than another? Is one catalyst more effective at speeding a chemical reaction than another?

Much of statistics, then, comes down to making judgments about these kinds of differences. We talk about “statistically significant differences” because statisticians have devised ways of telling if the difference between two measurements is really big enough to ascribe to anything but chance.

Suppose you’re testing cold medicines. Your new medicine promises to cut the duration of cold symptoms by a day. To prove this, you find twenty patients with colds and give half of them your new medicine and half a placebo. Then you track the length of their colds and find out what the average cold length was with and without the medicine.

But all colds aren’t identical. Perhaps the average cold lasts a week, but some last only a few days, and others drag on for two weeks or more, straining the household Kleenex supply. It’s possible that the group of ten patients receiving genuine medicine will be the unlucky types to get two-week colds, and so you’ll falsely conclude that the medicine makes things worse. How can you tell if you’ve proven your medicine works, rather than just proving that some patients are unlucky?