On Tuesday, the D.C. Circuit decided *American Petroleum Institute (API) v. EPA*, an interesting case dealing with nitrogen oxide (NO2) levels. The standard is supposed to include a margin of safety.Under the Clean Air Act, EPA sets National Ambient Air Quality Standards (NAAQS) for airborne substances that endanger human health or welfare. EPA set such a standard for NO2 in 1971 and finally got around to revising the standard in 2010.

The innovation in the new NO2 standard is that it’s a one-hour standard covering peak exposures, and all air monitors in an area must hit the standard. The previous standard was an annual average, so local, temporary peaks could be quite a bit higher. The evidence showed that the earlier average standard did not protect people against respiratory problems from spikes in nitric oxides, particularly if they were near freeways.

Two industry groups sued to overturn the new standard, but it was unanimously upheld by a panel containing two very conservative judges and one more liberal one. The court was distinctly unimpressed by the industry claims. In response to a claim that EPA violated its own rules because it relied on a study that wasn’t peer-reviewed, the court wrote, “Perhaps the API should have had its brief peer-reviewed.” The court faulted the industry brief for deleting crucial language when quoting an EPA document, among other errors.

Notably, the court flagged a common error in using statistics. Industry relied on a study that found no statistically significant relationship between concentrations of NO2 and health effects. According to industry, the study showed that there was no health effect. The court pointed out, however, the study did not prove that there was no health effect; it merely failed to detect one. Although people commonly confuse lack of evidence of an effect with proof that there *is* no effect, there is a fundamental difference.

It’s easier to understand the difference in a more everyday context. If you have an alarm system, there’s a tradeoff in deciding how sensitive the system should be. If you have a really sensitive system, it may often generate false alarms but is guaranteed to detect an intruder. If your system is less sensitive, you’ll have fewer false alarms but an intruder may go undetected. The fact that the alarm hasn’t gone off is some evidence that there’s no burglar, but if you’re really anxious to avoid false alarms, your system may well be missing actual intruders.

The same is true of statistical tests. A statistical test may fail to detect a

relationship either because it doesn’t exist or because the test isn’t sensitive enough. Statisticians talk about Type I and Type II errors, but they’re really just talking about the same tradeoffs as with burglar alarms between false alarms and missed intruders.

The court didn’t rely on this point alone, but also explained that EPA had plausible criticisms of the study’s methodology. Thus, it was reasonable to EPA to rely on the so-called non-peer reviewed study that industry complained about. I say “so-called non-peer reviewed” because EPA just updated a published study that *had* been peer-reviewed, and its update was reviewed by the Science Advisory Board, a form of peer-review.

As far as I can tell, the industry just wanted the court to second-guess the agency’s scientific judgments. But it’s not the court’s job to play amateur scientist. EPA clearly gave plausible explanations for its expert judgments, and that’s all the law requires.