Toward a More Disproportionate Epidemiology

Jay S. Kaufman

I was lured into epidemiology by a friend in environmental engineering. "But don't worry," he assured me. "You don't have to take any classes or anything. It's not a real science like chemistry or physics."

Years later, a fictional department chair heard this story and was intrigued by the idea that teaching epidemiology might offer no benefit whatsoever for a large number of graduate students. He decided to save scarce funds by paying me for teaching only those students who would pass my course because they attended the class. There are 3 kinds of students, he reasoned: the type (A) who would pass the examination with or without attending the lectures, the type (B) who would pass the examination if they attended but fail if they did not, and the type (C) who were doomed to fail the examination regardless. He told me to figure out the number of type Bs in the student population, because this is the only group worth teaching.

Fortunately, I had myself studied epidemiology, so I knew to partition my incoming class at random, which assured that the expected representation of each of these latent types would be the same in the 2 groups. Then I taught one group of 30 students my usual course and assigned the other group of 30 to stay away. Everyone was compliant with their assignments and there was no communication about epidemiology among the students (needless to say). At the end of the term, I gave my examination. Only 6 of 30 students passed the examination without the class, but 18 of 30 passed with instruction. I reasoned that the number who passed in the group with instruction was A+B, whereas the number who passed in the group without instruction was simply A. The ratio of these numbers would be the causal effect of teaching, which was 18/6 = 3.0. It seemed that my teaching had tripled the pass rate, which made me happy, but did not please my fictional department chair. He wanted to pay me on the basis of the number of people who were Type B, whereas I had only identified the quantities A, A+B, and their ratio (A+B)/A.

Again fortune came to my rescue, as my research assistant K. had been assigned to the group without epidemiologic instruction. Free from the strictures of epidemiologic habit that I had learned and passed on to my other students, K. advised me to simply take the difference between the 2 numbers instead of their ratio. The difference of (A+B) in the treated group minus A in the untreated group would yield B. A total of 18 − 6 = 12 students passed the examination because of the instruction, a number completely obscured by the relative contrast that I had made without K.'s assistance.

K.'s radical analytic insight proved even handier the following year, as the worsening economy drove many highly qualified applicants back to graduate school, and so the overall failure rate on exams decreased dramatically. There were again 60 students, with 30 assigned to each group. This time, 8 of 30 students passed the examination without the class, but 24 of 30 passed with instruction. The ratio (A+B)/A showed that once again my instruction had tripled the pass rate, since 24/8 = 3.0. Without K.'s help, I might have claimed to have had the same effect on my students as in the previous year,

but instead I was pleasantly surprised to find that $24-8 = 16$, and so I could collect from my fictional department chair an additional salary for having caused 4 more students to pass than in the previous year. My cohort size and my ratio measure of effect were both identical, and yet I had affected 33% more students this year than the year before.

How could the ratio effect measure—surely our most familiar analytic approach—lead one so far astray? And if it could, why is it still our favored approach? In June of 2009, the editors of Epidemiology convened a symposium at the annual meeting of the Society for Epidemiologic Research in Anaheim, California. The purpose was to ask why epidemiologists have come to rely almost entirely on relative measures of effect (odds ratios, risk ratios and hazard ratios), even though this practice generates considerable confusion, especially over interaction, effect modification, and the potential public health benefits associated with reported effects. Many of these problems could be avoided simply by a greater attention to the baseline risks in our research and the reporting of our results.

The 3 presentations from this symposium are provided here as commentaries. Charles Poole[1] begins with a historical quest to understand the roots of our prejudice favoring relative contrasts, and the rationale for the common wisdom that relative effect measures are more suited for etiologic hypotheses. His account highlights the unintentional adverse effects of "causal criteria," such as those popularized by Bradford-Hill, and represents the kind of insightful intellectual history of our field we seldom see. Bryan Langholz[2] follows with a discussion of the case-control study design, and the pervasive yet erroneous belief that this design restricts us to the odds-ratio scale. As he shows, we can almost always choose to conduct studies that retain absolute risk information, yet for reasons that may range from habit to ignorance, we rarely do. Moreover, even when this information is available, we seldom use it. These failings cannot be attributed to the case-control design itself, only to our apparent reluctance to apply it more effectively. Finally, Miguel Hernán[3] comments on a built-in bias lurking in hazard ratios. The problem is that the hazard at a particular time is, by definition, calculated in the subset of individuals who survived through that time. Thus hazard ratios are calculated in a surviving cohort that is increasingly selected over time. One solution is a contrast of absolute survival curves – an analysis that does not refer to the relative frequency of events occurring in a surviving subset, but rather simply to cumulative proportions of the original cohort that have failed or survived at each time point.

The editors of Epidemiology hope these essays will spur reflection and discussion. While we won't insist on one effect scale or another as a blanket policy, we encourage authors to think carefully about what design and analysis best fit their study questions and subject-matter needs. We suspect that such reflection is likely to lead to much more research being situated on the absolute scale. Old habits die hard, even when they no longer serve us well (if indeed they ever did). But if we can transcend our outdated aphorisms and shed our comfortable old misunderstandings, perhaps epidemiology may be a real science after all.

REFERENCES

1. Poole C. On the original of risk relativism. *Epidemiology*. 2010;21:3–9.
2. Langholz B. Case-control studies—odds ratio: Blame the retrospective model. *Epidemiology*. 2010;21:10–12.
3. Hernán MA. The hazards of hazard ratios. *Epidemiology*. 2010;21:13–15.