

P E R S P E C T I V E

Cooper's Analysis Is Incorrect

The authors of the work critiqued by Richard Cooper demonstrate that his findings bolster their conclusions, instead of contradicting them.

by **Katherine Baicker and Amitabh Chandra**

ABSTRACT: In his papers, Richard Cooper finds positive associations between health care quality and both specialist and generalist physicians, but he misinterprets his results. Instead of undermining the findings of our study, which found higher quality in areas with more generalists relative to specialists, his results bolster ours: they suggest that the effect of generalists on quality is ten times larger than that of specialists. Furthermore, his rejection of multiple regression in favor of exclusive reliance on isolated correlations precludes him from gauging the relative contributions of specialists, generalists, and other factors. Unfortunately, these deficiencies mean that we can learn little from Cooper's analyses. [*Health Affairs* 28, no. 1 (2009): w116–w118 (published online 4 December 2008; 10.1377/hlthaff.28.1.w116)]

THE PAPER “States with More Physicians Have Better-Quality Health Care” by Richard Cooper makes two claims.¹ First, he finds positive correlations between the quality of care and the number of both specialist and generalist physicians. He concludes from this that specialists and generalists contribute equally strongly to quality. Second, he believes that his results discredit our study, which found that quality is higher in areas where generalists constitute a greater share of the physician workforce.

Here we demonstrate that both of Cooper's conclusions are incorrect. First, he ignores statistical inference, reporting statistically insignificant correlations as if they were more definitive. Second, he confuses correlations (which tell whether or not two variables are related) with the magnitude of the effects

(which tell you how big the relationship is), and thus fails to recognize that his exhibits suggest that generalists have an effect on quality that is ten times larger than that of specialists. Third, as we illustrate with an example, his rejection of multiple regression in favor of considering each variable only in isolation results in misleading inferences. Fourth, his research fails to meet scientific standards in other important dimensions.

Scoring Cooper's Analysis

■ **Ignores statistical inference.** Cooper's own data show no statistically significant relationship between specialists and quality, but they do show a statistically significant relationship between generalists and quality. The tools of statistical inference exist to help us distinguish between correlations that rise to

Katherine Baicker (kbaicker@hsph.harvard.edu) is professor of health economics at the Harvard School of Public Health and a research associate at the National Bureau of Economic Research (NBER) in Cambridge, Massachusetts. Amitabh Chandra is an assistant professor of public policy at Harvard's Kennedy School of Government and a faculty research fellow at the NBER.

the level of a detectable relationship and those that in essence show no relationship. Cooper, however, ignores the basic standards of inference when he writes, “Quality is better in states with more physicians, both specialists and family physicians” (p. w89). This is not an isolated or off-handed statement, but rather his title and main conclusion. A more careful statement would be, “Quality is better in states with more family physicians, but no significant association was found for specialists.” This statement is entirely consistent with the results found in our paper.

■ **Confuses correlations with magnitudes (missing the key implications of exhibits).** Cooper’s analysis rests on the idea that because the correlation between quality and the number of generalists is similar to the correlation between quality and the number of specialists, both types of physicians contribute equally to quality. This is simply wrong. Correlations tell us whether two variables are associated, but not how big a change in one variable we can expect to see associated with a change in the other (that is what regression coefficients tell us). The numbers of specialists and generalists per capita may have identical correlations with quality, but they have very different size effects on quality. Cooper’s own exhibits suggest that this is the case—and that generalists have a dramatically bigger effect on quality than specialists do. Although nowhere does he report the slopes of the lines that he shows in his Exhibits 1 and 2, a simple visual inspection suggests that you would have to add roughly ten specialists per capita to move up ten spots in the quality ranking, but you would only have to add one generalist per capita to move up the same ten spots. And, apparently, even the small effect of additional specialists on quality is statistically insignificant. The size of these relationships is far more useful for the policy debate than simple correlation coefficients, and once again Cooper’s data confirm our original findings.

■ **Disregards the importance of confounding factors.** Simple correlation coefficients (or slopes, for that matter) do not allow researchers to take other factors into account.

These confounding factors can be of vital importance in evaluating the effect of medical interventions on health, as illustrated by the recent reevaluation of the efficacy of hormone replacement therapy. Such confounders are equally important in observational studies such as Cooper’s and ours. Fortunately, there exists a well-established and widely used technique to gauge the importance of confounders: regression analysis.² Yet Cooper views our use of basic regression as newfangled and misleading.

Here is a simple illustration of the importance of considering multiple factors: the correlation between median house values and health care quality rank (where a lower number for rank indicates higher quality) is a statistically significant -0.24 .³ Could more expensive houses be good for health care quality? We can learn more about the magnitude of this relationship from a regression: a 10 percent increase in home value moves a state up in the quality rankings by 1.4 spots. It would be ridiculous to conclude from this, however, that we can improve health care quality by driving up home values. What else could be going on? States with higher-income residents have both higher home values and higher quality of care. If we run a regression looking at the effect on quality of both home value and income, the coefficient on home value drops by a factor of 10 and becomes statistically insignificant, while the coefficient on income remains large (a 10 percent increase in income is associated with a move up in the ranks of 3.4 spots) and statistically significant. If we just considered these two factors separately, we would not be able to tell which association was stronger. It is only by holding income constant that we see that home value has no independent effect on quality.

Cooper’s analysis suffers from exactly this deficiency—because the number of specialists and the number of generalists in a state are related, looking at them independently cannot tell you the relative contribution of each to quality. We thus run a regression where we include both the total number of physicians and the share who are generalists. We learn from this regression that although the total number

of physicians is important, the share who are generalists makes an independent contribution to quality. Put another way, if we compare two areas with the same total numbers of physicians, the one where a greater share are generalists will have higher quality (and lower spending). Cooper mischaracterizes this as “theoretical,” “hypothetical,” and “imputed.” Rather, we examine the empirical relationship (not a theoretical model) between actual physicians (not hypothetical ones) and the same quality rankings he uses (not imputed ones). Cooper’s characterization of our findings as “manipulating statistical ‘residuals’” in a way that is “needlessly misleading” suggests a misunderstanding of multiple regression: we do use “actual” specialists and generalists, just holding the total workforce constant. Every research paper that examines more than one factor at a time is using this same framework.

■ Fails to meet scientific standards.

There are other technical issues that we do not have space to address fully here, including the fact that Cooper does not report data means, whether his regressions are weighted based on state size, or even the regression coefficients that he graphs in his exhibits, making it impossible for serious researchers to evaluate his findings even using his criteria. We note as an aside that Cooper seems to have used rank as the independent explanatory variable, rather than as the dependent outcome of interest that his discussion suggests. The findings that are reported are often muddled. There are several statements—for example, “(log regression coefficient: $R^2 = 0.60$)” (p. w92) or “States with larger populations and greater population density have more physicians per capita, and, moreover, they have more physicians at each level of quality ranking” (p. w97)—that reflect confusion about statistics.

Concluding Observations

In our study, we found that the quality of care was lower in areas where spending was higher, and also lower in areas where specialists made up a greater share of the physician workforce. The evidence found in Cooper’s analysis only strengthens these findings. Of

course, it could be that some other omitted factor is an important part of the story—which is why one must be cautious in attributing causality to the observed relationships, and why it is valuable to use even more sophisticated statistical techniques (for example, we explore the importance of state “fixed effects” to account for any factors that are persistently different across states).

Cooper is interested in the relationship between the size of the physician workforce and quality—an important but different question from our investigation of the relationship between the composition of the workforce and quality. Unfortunately, Cooper’s analysis sheds light neither on our question nor on his own.

Health care quality is the result of many factors, and the supply of specialists and that of generalists have common determinants that may also affect quality directly. Serious attempts to measure the contribution of the size of the workforce to quality need to take this interconnectedness into account, but Cooper does not. Other prominent research (uncited by Cooper), however, uses more sophisticated methods to examine this important question and does not find a relationship between quality and the number of physicians.⁴ In light of these serious deficiencies, we learn little from Cooper’s analysis.

NOTES

1. R.A. Cooper, “States with More Physicians Have Better-Quality Health Care,” *Health Affairs* 28, no. 1 (2009): w91–w102 (published online 4 December 2008; 10.1377/hlthaff.28.1.w91).
2. G.U. Yule, “On the Theory of Correlation,” *Journal of the Royal Statistical Society* 60, no. 4 (1897): 812–854; and K. Pearson et al., “The Law of Ancestral Heredity,” *Biometrika* 2, no. 2 (1903): 211–228.
3. U.S. Census Bureau, “Historical Census of Housing Tables Home Values,” <http://www.census.gov/hhes/www/housing/census/historic/values.html> (accessed 4 November 2008).
4. D.C. Goodman et al., “The Relation between the Availability of Neonatal Intensive Care and Neonatal Mortality,” *New England Journal of Medicine* 346, no. 20 (2002): 1538–1544; and D.C. Goodman and E.S. Fisher, “Physician Workforce Crisis? Wrong Diagnosis, Wrong Prescription,” *New England Journal of Medicine* 358, no. 16 (2008): 1658–1661.