Every clinician-scientist admires the randomized, controlled trial (RCT) as the best-known method for studying the effects of a treatment on specified outcomes. A good RCT eliminates (by random assignment) any statistical linkages between the characteristics of the participants (for example, age, unmeasured comorbid conditions) and outcomes. Alas, we have no RCTs for many medical interventions. However, several valuable alternatives have emerged, including case-control analyses and analysis of regional variations in treatment patterns and medical outcomes (“small-area variations”). In this issue, Fisher and colleagues (1, 2) use small-area variation analysis to illuminate the effects of medical resource use on patients presenting with hip fracture, colorectal cancer, and acute myocardial infarction.

The authors characterized a region’s propensity to use medical resources based on the overall spending patterns in Medicare patients’ last half-year of life. These expenditure patterns depended mostly on inpatient treatment choices and specialist and subspecialist use, usually involving discretionary care, in contrast with relatively uniform patterns for “evidence-based medicine” interventions. This supports the premise that regional variations arise from disagreements about proper uses of medical interventions (3, 4).

Most important, the second of the two articles by Fisher and colleagues (2) demonstrates that patients in high-spending regions had no better survival (if anything, slightly worse) than those in lower-spending regions. Separate analyses of a sample of all Medicare beneficiaries further showed no gains in overall patient mortality, functional status, or patient satisfaction.

Small-area variation works best when geographic loca-
tion acts as a proxy for an RCT-like randomization, eliminating links between patients’ conditions and the therapy they receive, an approach often called natural randomization (5). But various problems can confound natural randomization, including migration (patients moving to where the medical treatment matches their medical conditions or preferences) and medical providers “following” unusually large concentrations of people with particular conditions (for example, cardiologists following elderly people to Florida).

Fisher and colleagues took several steps to ensure that they had achieved natural randomization. For example, they separately analyzed subsets of their sample who had moved from low-intensity regions, those who had not moved from high-intensity regions, and those who had not recently moved. They found the same outcomes in each subgroup. They checked for urban–rural differences, for differences in penetration of health maintenance organizations in the regions, for teaching status of hospitals in a region (often a proxy for more complicated patients), and for hospital volume (also related to patient severity and complexity). In no case did their results differ across subgroups, increasing our confidence that hidden or unmeasured patient preferences or medical complexity did not drive their results. They also used information about each study patient’s clinical conditions to assess whether underlying complications or complexities affected their results (they did not).

It is important to understand that Fisher and colleagues have not studied the effects of particular medical interventions on outcomes of care for patients with acute myocardial infarction, colorectal cancer, and hip fracture. Rather, they studied the outcomes of care for patients who lived in areas characterized by different medical styles and overall treatment intensities. Thus, their results do not show whether any specific treatment protocols are more or less effective. Rather, their study illuminates the question of whether reducing intensity of medical treatment overall would likely have important effects on patient outcomes. If their natural randomization is effective, then one can infer that patients in the areas with greater intensity gained little or no additional health (or, by Fisher and colleagues’ estimates, had actually slightly worse health outcomes), at least for the index diseases studied. Of course, if the small-area variation approach did not achieve natural randomization, then drawing that inference would risk a possible ecologic fallacy.

This question sits at the center of health policy debates: Can we afford to reduce the overall intensity of medical treatment without paying an overly large price in terms of reduced health? Fisher and colleagues show a world in which the productivity curve in medical care has become flat; no more health emerges when treatment occurs at a higher-than-average intensity. This is not to say that improved health care has not provided important gains for the American people, only that increasing intensity at the upper end of the intensity spectrum is unlikely to improve health outcomes, and indeed (again, subject to the proper caveats), resource use in those areas could be reduced with no decrement in health outcomes.

Obvious questions remain, even if the results of this study are accepted. Do important benefits remain unmeasured? Some studies suggest that using mortality alone may mask some valuable benefits. For example, regional comparisons of patients with heart disease showed no greater survival in Ontario, Canada, where coronary artery bypass grafting and angioplasty are performed infrequently, than in upstate New York, where these procedures are very common; however, the New York patients gained important improvements in other health-related areas, such as chest pain and shortness of breath (6). Fisher and colleagues also measured patient satisfaction (finding no differences across regions), some components of which should correlate with unmeasured quality-of-life variables.

This study’s national scope, its careful control of possibly confounding variables, and its underlying design make it possibly the most compelling yet showing that increased treatment intensity does not bring with it commensurate gains in health. Our health care market has many features guaranteed to increase treatment intensity. This study adds greatly to the value of asking, “What’s enough, what’s too much?”

Charles E. Phelps, PhD
University of Rochester
Rochester, NY 14627

Requests for Single Reprints: Charles E. Phelps, PhD, 200 Wallis Hall, University of Rochester, Rochester, NY 14627; e-mail, charles.phelps@rochester.edu.


References

© 2003 American College of Physicians–American Society of Internal Medicine

18 February 2003 Annals of Internal Medicine Volume 138 • Number 4 349