The Savings Illusion — Why Clinical Quality Improvement Fails to Deliver Bottom-Line Results


It has become a core belief in U.S. health care that improving clinical quality will reduce health care costs. It seems intuitive that reducing readmissions, shortening lengths of stay, and building efficiency into clinical processes will reduce resource utilization and thereby lower costs. Certainly, evidence suggests that there is no association between high quality and high costs. Yet true bottom-line savings from improved clinical quality rarely materialize, and costs continue to climb. Manufacturing and service companies around the world have demonstrated the cost benefits of improving product quality and production efficiency. So why haven’t nearly two decades of work on improving health care quality had a measurable effect on health care costs?

The explanation lies in the cost structure of the typical health care setting. Its management and organization create a rigid cost structure that is relatively insensitive to small changes in patient volume, resource use, or the severity of patients’ health conditions. This fixed-cost dilemma leaves most health care costs insensitive to changes in volume and utilization, so clinical quality improvements typically create additional capacity rather than bottom-line savings. An examination of the different cost layers highlights the distinction between variable costs, such as supplies and medications, where reduced use produces true savings, and fixed costs, such as facilities and ancillary services, where the costs persist despite reduced use.

To better understand the cost structure of health care delivery, it can be useful to consider how different costs behave depending on the degree to which they are sensitive to changes in resource utilization. The four cost layers we have identified are defined in the table.

Clinical improvements that reduce layer 1 costs, such as those of supplies and medications, will generally create bottom-line savings, since these are the only truly variable clinical costs in a hospital. To generate savings by reducing use of the resources that account for layer 2 costs, the need for the resource must be reduced enough to allow elimination of a payable unit. For instance, a single nursing unit might have to...
discharge multiple patients before any savings in hourly nursing labor costs could be captured by allowing an hourly employee to go home early. Reducing layer 3 resources — those for equipment, operating-room time, or physicians' salaries, for example — almost always produces additional capacity without bottom-line savings. If an intervention reduces operating-room time by 15 minutes, the costs of the equipment and salaried staff required to run the operating room do not change. Nonclinical layer 4 costs are primarily fixed in the short run, but reducing administrative labor costs by achieving administrative efficiency will produce true savings in future operating cycles.

Because of these cost behaviors, quality-improvement efforts that reduce lengths of stay or re-admissions or increase radiology throughput do not create substantive bottom-line savings. They generally create capacity to treat additional patients. Similarly, efforts to expand the access of disadvantaged populations to primary care under the assumption that such access will be paid for through avoiding use of high-cost care sites — such as emergency departments — do not generate cost savings. The cost of staffing and equipping an emergency department does not change if there are small reductions in utilization. Indeed, improved access will increase health care costs if new physicians and staff are hired to serve new patients in primary care practices.

Although capacity creation does not generate bottom-line savings, it does create an opportunity to admit another patient and collect additional revenue. Because health care costs are relatively fixed and do not change much at the margin, the cost of admitting a new patient is remarkably low, making volume growth a highly profitable strategy. Volume growth also can give the appearance of reducing costs, since the cost per case decreases when the high fixed costs are spread over a larger number of patients, although total costs will probably continue to rise. Growing volume and increasing revenue, rather than creating true bottom-line savings, are typically at the core of the business case for high-quality care.3

Because of the rigid cost structures, incremental reductions in resource use are unlikely to generate cost savings for either a health care setting or the health care system. The most meaningful way to achieve savings is to focus on overall reductions in utilization rates for health care services and to eliminate the associated unnecessary capacity.

In a recent article, Kaplan and Porter argue that most health care costs are not fixed.4 Postulating that personnel costs can be adjusted and space reallocated on the basis of demand and patient mix, they suggest that cost behaviors are not responsible for the inability to generate cost savings, but “management inattention” is. Although we do not dispute this logic, its practical application is dependent on both procedure volume and the time horizon required for aligning resources with demand. High-volume procedures and treatments for which resource use can be standardized across the cycle of care and for which capacity can be readily adjusted to accommodate appropriate volume appear to be best suited to the aggressive cost management advocated by Kaplan and Porter. Presumably, lower-volume treatments and procedures would have to be consolidated regionally to be more amenable to effective cost management. Until that happens, the fixed-cost dilemma will remain an obstacle. Cost layering provides management with a framework for targeting chang-
es that will generate the most immediate savings.

Whereas quality improvement is producing significant benefits for patients, quality initiatives will continue to produce disappointing bottom-line savings as long as the capacity created is used to support growth in patient volume. As the U.S. health care system begins shifting its focus from volume to value, hospitals will need to adapt their cost structures and capacity to accommodate lower per capita utilization rates as well as reductions in the per-episode intensity of care.

Disclosure forms provided by the authors are available with the full text of this article at NEJM.org.

From the Dartmouth Institute for Health Policy and Clinical Practice, Lebanon, NH.

This article (10.1056/NEJMp1111662) was published on December 14, 2011, at NEJM.org.


Copyright © 2011 Massachusetts Medical Society.