Health Care Reform  
From Myth to Practice  
Richard A. Cooper, MD

As the United States enters the “posthealth care reform era,” physicians are faced with both the realities of practice and the myths of reform. The “Realities of Practice” are that the demand for health care will grow in direct proportion to economic expansion, the complexities of health care will grow in direct proportion to technological development, and the needs of patients will grow in inverse proportion to their education and social well being. More physicians will be required to provide a wider range of services to patients in widely differing economic circumstances.

The “Great Myth” is that the United States has too many specialists, which leads to too much health care spending, and that if spending everywhere could resemble spending in geographic regions where it is the lowest, health care expenditures would decrease by 30%; enough to finance health care reform.

Myth and reality could not be more different, but myth rather than reality has driven the process.

As is true for all myths, this one is rooted in elements of reality. Specialists and generalists alike are critical of what they do and seek improvement. Self-critical exercises, such as morbidity and mortality conferences, are one historic example, a prelude to the quality movement, with its bent toward regulation. Safer anesthetics and improvements in operating suites and critical care units further testify to this broad effort, aided by advances in diagnostic and monitoring equipment. Physicians are the most self-critical professionals on earth.

It is also true that physicians are imperfect, and some are less than imperfect. It is easy to find examples of waste, inefficiency, and poor judgment, or of unneeded care and defensive medicine. However, commitment, compassion, and self-sacrifice are more the rule. The system is also imperfect, with its redundancies, administrative complexity, regulatory zeal, and excessive litigation. Both the profession of medicine and the system of medical care are imperfect. That is why physicians meet to discuss and search for better ways to prepare for the realities of clinical practice.

The question is whether medicine’s imperfections result in the overuse of supply sensitive services provided by specialists who gravitate to regions where they can provide uninterrupted care, and therefore, the nation would be better-served by having fewer specialists, as those who led health care reform would have us believe.

THE WAR AGAINST THE SPECIALISTS: THE FIRST 100 YEARS

A war against specialists is not new. Concern about specialists has existed for more than a century. In 1895, Stedman observed that specialists were squeezing out family doctors “as vines do the big trees.”1 As expenditures for health care began to grow 50 years later, these concerns became entwined with parallel concerns that specialists were driving health care spending2,3; economists believed that the supply of surgeons determined the volume of surgery.4 This phenomenon was coupled with apprehension that, with overzealous expansion of medical schools in the 1970s, surpluses of specialists would soon exist, while the supply of internists and family physicians would languish,5 a prediction that was legitimized by the now-discredited reports of the Graduate Medical Education National Advisory Committee (GMENAC) in 19816 and the Council on Graduate Medical Education (COGME) in the early 1990s.7

The troika of too many specialists, too few primary care physicians, and too much spending became a fixture of health care planning, and its acceptance became the prerequisite for leadership positions in virtually every important organization. Doors to contrary views were closed.

During the Clinton Health Plan era, this sentiment culminated in a proposal to decrease the number of residents being trained by about 25% so that the total would be equal to 10% more than the number of US medical graduates, and half were to become primary care physicians; the 110%-50:50 rule. This proposal was supported by a broad base of statistical evidence, none of which proved to be valid, but it did not require statistics. It expressed “a social judgment, not a scientific judgment.”8 Proponents viewed it as part of the “battle for the soul of medicine,” which had become “Balkanized by specialists,”9 who were the “invisible drivers of health care costs.”10

Although efforts to insert the “110%-50:50” rule into the Clinton Health Plan failed, as did the Health Plan itself, proponents were successful in establishing a freeze on Medicare spending for graduate medical education several years later, and that persists to this day.11 Even more, the belief that specialists bore responsibility for the health care system’s woes was etched into the popular culture.

As the year 2000 approached, it was clear that the massive surplus of physicians that GMENAC and COGME had projected were not materializing, as I had predicted.12 Indeed, my colleagues and I demonstrated that economic and demographic trends were likely to result in profound shortages if training capacity was not increased.13 Although our “contrarian views” were embraced by Senator Moynihan and communicated to his Senate colleagues,14 they were rejected by most academic leaders15 yet history proves that we were right.

THE DARTMOUTH ERA

During the 1970s, Wennberg and Gittelsohn at Dartmouth began to codify what for most physicians was a common observation—the fact that different physicians did things differently, even in the same institution.16 They extended this observation to discover systematic differences between communities, which were not readily explained by the characteristics of the patients living in them. In due course, they constructed an Atlas composed of 306 “health care communities,” which they called hospital referral regions (HRRs).17 Each was a closed system where most of the patients received most of their care most of the time. This offered a way to compare health care across the nation, but although ingenious, their system was fatally flawed in 3 ways.
First, it was based entirely on Medicare data. The assumption was that data obtained through Medicare applied equally to other insurance groups, but as I demonstrated, that is not the case. Medicare expenditures per enrollee bear no resemblance to overall spending per capita, nor should they. Communities differ in the generosity of Medicaid and employer-sponsored insurance, as well as in the prevalence of uninsured. It is the composite revenues from all of these sources that pay for nurses and other personnel, financial facilities, fuel the processes of care, and influence outcomes.

Second, the Dartmouth group’s observations were limited to measuring expenditures or utilization, but not outcomes, leaving questions of value unanswered. They attempted to circumvent the problem of risk-adjustment by focusing their studies on patients during the last 2 years of life. Their Web site explains that “we focused only on patients who died so we could be sure that all patients were similarly ill. By definition, the prognosis was identical—all were dead. Therefore, variations cannot be explained by differences in the severity of patients’ illnesses.” Of course, as every physician knows, similarly dead is not similarly ill, a point more forcefully made by others. Vast differences exist in the complexities of disease among patients who have died.

Probably the most important factor associated with differences in complexity, costs and outcomes is income. As discussed later, socioeconomic status has a great deal to do with resource consumption—low-income patients consume more than double, and of course, income is geographic. Poor people live in poor neighborhoods and rich people live in rich neighborhoods, all within communities that may be rich or poor in states of varying wealth. Patients’ needs and burden of illness are strongly related to their economic status, whereas the availability of resources to meet those needs is related to broader measures of wealth, a phenomenon that I have termed “the affluence-poverty nexus.”

Finally, Dartmouth’s 306 HRRs vary enormously. Populations range from fewer than 200,000 to more than 5.0 million, and land mass varies from as little as 50 square miles to more than 50,000 square miles. Some include a single urban area and others extend the length of a state. Most important, average per capita income ranges widely, particularly in HRRs that encompass urban centers where affluence and poverty tend to coexist. Aggregating and averaging these varied income groups in HRRs of dissimilar population and area makes it impossible to understand the underlying connections between income, burden of disease, utilization, and outcomes. Simply stated, HRRs are not a valid unit of analysis.

THE 30% SOLUTION

Thus, the Dartmouth group faced a dilemma. They could measure inputs, albeit with the limitations of Medicare data and in ways that were plagued by aggregation and averaging, but they could not measure outcomes, certainly not by studying patients who had died. The solution was to reach into large national data sets that could not measure outcomes, certainly not by studying patients who were plagued by aggregation and averaging, but they had died. Therefore the 306 HRRs were collapsed into 5 “quintiles,” not based on geography, but on the average Medicare spending in each. As a result, geographically heterogeneous HRRs, each with heterogeneous mixtures of affluence and poverty, were aggregated, and the average outcomes in each were measured, and they were all the same—they were all “average.” That should not have been the case if the outcomes related directly to Medicare spending, but they did not. Outcomes measures, such as 5-year mortality, functional status, access to care, satisfaction with care, and process measures, such as mammography and the use of aspirin after acute myocardial infarction, all reflect community standards and relate to the total funds devoted to health care from all sources, not just Medicare, as discussed earlier.

The results were that “quintiles” with higher average Medicare spending had average outcomes that were no better than in “quintiles” where Medicare spending was lower. This, of course, was due to the fact that the average total spending per capita in each was the same. Nonetheless, Dartmouth researchers insist that Medicare spending is representative of overall health care spending, which is simply not possible. But if Medicare is not representative, the Dartmouth story evaporates, and so many people have bought into it that it is too good to give up. So it persists, despite the facts. As the King of Siam cautioned his son: “Tho’ a man may be in doubt of what he know, very quickly he will fight to prove that what he does not know is so.” And so they did. Dartmouth researchers proclaimed:

Regional differences in Medicare spending are due almost entirely to use of discretionary services that are sensitive to the local supply of physicians and hospital resources: more frequent physician visits, greater use of specialists, and greater use of hospital and intensive care unit as sites of care. Policymakers and purchasers concerned with resurgent growth in health care spending will need to focus on these supply-sensitive services. If the United States as a whole could safely achieve spending levels comparable to those of the lowest-spending regions, annual savings of up to 30% of Medicare expenditures could be achieved. Such savings could provide the resources to fund important new benefits or to extend the life of the Medicare Trust Fund for future retirees.

SELLING THE 30% SOLUTION

Over the next few years, the 30% solution was cited by the Dartmouth group and others who shared their perspective, and it was occasionally cited in the lay press. It was not until late in 2007 that, in a blitz-like fashion, it found voice in the popular culture. First, the book “Overreached,” by Shannon Brownlee, a Wennberg collaborator appeared. She asks, “Why would more doctors lead to worse care, and fewer doctors to better care?” She even quoted a Dartmouth researcher saying, “If we sent 30% of the doctors in this country to Africa, we might raise the level of health on both continents.” David Leonhart of the New York Times named Overreached the economics book of the year for 2007, noting that “we now spend between one fifth and one third of our health care dollars on care that does nothing to improve our health.”

Then in rapid succession Peter Orszag, director of the Congressional Budget Office and soon-to-be director of the Office of Management and Budget (and President Obama’s right hand on health care), published a 2-part paper in the New England Journal of Medicine, pointing out that “higher-spending regions do not have higher life expectancies or show significant improvement on other measures of health,” and later in the Wall Street Journal added that “if we can move our nation toward the practices of lower-cost areas, health-care costs could be reduced by 30%, about $700 billion a year.” This was followed by a JAMA Commentary written by members of the IOM’s Board on Health Care Services, who noted that “waste accounts for 30% to 50% of health care spending,” and by a New York Times editorial, which concluded that “if the entire nation could bring its costs down to match the lower-spending regions, the country could cut perhaps 20 to 30% from its health care bill, a tremendous saving.” This was a full year before the presidential election and more than 2 years before Health Care Reform was passed. It was irrefutable as the recipe for health care reform with no new taxes. However, it was simply a recipe for cake with rancid flour. In the last analysis, the Dartmouth Atlas has proven to be a map to nowhere.
A TALE OF 2 CITIES

If the problem with the Dartmouth group’s studies of geographic variation was excessive aggregation and averaging, it should be possible to understand the process by disaggregating these larger regions into smaller and more homogeneous units of analysis. To this end, my colleagues and I have examined health care utilization and have studied its relationship to income, using ZIP codes as our units of analysis, focusing on the Milwaukee and Los Angeles HRRs.

Because Milwaukee is so racially and economically segregated, its ZIP codes proved to be relatively homogeneous with respect to socioeconomic characteristics. Moreover, its low-income ZIP codes were clustered within a very narrow “poverty corridor.” The analyses of these data led to 2 important findings. First, that income and health care utilization are strongly correlated, with sharply higher utilization at the lowest incomes. Secondly, whereas per capita utilization of health care was approximately 30% greater in the Milwaukee HRR than in others in Wisconsin, this difference was entirely accounted for by the poverty corridor. Utilization in the corridor was almost double the rate in the rest of Milwaukee, but the rate in the rest of Milwaukee was no different than in other HRRs in the state. Clearly, at least in the Milwaukee HRR, excess utilization did not result from the overuse of supply sensitive services. It was due to the greater health care needs and higher utilization of Milwaukee’s low-income population.

Los Angeles presented a similar but more complex picture. Its population is 8-fold that of Milwaukee’s. Like Milwaukee, it has a dense “core” of profound poverty, which includes about 5% of its population, and this is surrounded by a larger “poverty zone,” encompassing another 20%, and small areas of poverty are scattered elsewhere. In contrast, almost 5% of the population resides in ZIP codes with an average household income of $100,000 or more. When the relationship between income and utilization was examined throughout the area, the pattern that was observed in Milwaukee emerged again: a steeply inverse relationship between household income and hospital utilization.

Health care utilization in the “core” was double the level observed in high-income ZIP codes, and utilization in the surrounding poverty zone was 65% greater. Health care utilization in Los Angeles overall was more than 30% greater than in its wealthiest ZIP codes.

Was variation in health care simply a manifestation of variation in the distribution of lower-income households? The easy way to answer this question was to confine our analysis to households with high average incomes, which we did in Los Angeles and 7 other populous counties in California (Fig. 1). Among these 8 counties, overall utilization was highest in Los Angeles, 65% higher than in the lowest county, Marin County, one of California’s wealthiest. Utilization in other counties was intermediate between both these extremes. However, when only ZIP codes with incomes more than $100,000 were examined, there was very little variation. If geographic variation was related to the intensity or style of physician practices, one would expect to see geographic differences among the highest-income segment, from which the most revenue could be extracted, but that was not the case.

The tale told by these 2 cities is that health care utilization is very closely tied to income, that when the urban poor are excluded, health care utilization among regions is quite similar; and that when only the wealthy are considered, there is no appreciable variation at all. Moreover, in both Wisconsin and California, the extra utilization by lower-income individuals accounts for approximately one-third of the total. The 30% solution is a myth, or more accurately, the 30% variation is a reflection of variation in poverty.

FIGURE 1. Variation in Hospital Days per 1,000 among Eight California Counties. Acute-care hospital days per 1,000 of population among ages 45–64 were measured at the ZIP code level in eight California counties, both for all ZIP codes and for those ZIP codes with a median household income greater than $100,000.

DARTMOUTH’S 30% SOLUTION AND PRESIDENT OBAMA’S 1.5% SOLUTION

A catch-phrase of health care reform was “bending the curve,” the notion that measures could be taken that could decrease the growth of health care spending. In 2009, the United States spent $2.5 trillion on health care, 17.2% of its gross domestic product (GDP). On the basis of projections made by the Council of Economic Advisors and the Congressional Budget Office, existing trends will cause health care spending to reach more than $6.75 trillion in 2025, 25% of the projected GDP. Adjusted for inflation and population growth, spending will increase 65% between now (2010) and then.

The 30% solution was presented as the magic bullet to flatten the curve and, furthermore, as proof that no additional physicians would be needed over the next decade. The President extracted promises from various health care groups to aid in slowing the growth of health care spending from a rate that is 2.5% greater than GDP growth, the historic trend, to 1.0% greater than GDP, a 1.5% decrement. This rate is implicit in the projections of future growth based on the health care reform legislation, although it is uncertain whether any of the measures in the bill will bend the curve. However, even if they do, the degree of bending projected will simply delay the time needed to reach a 65% growth in spending by a few years, possibly to 2030. Moreover, irrespective of these projections, the population will continue to grow and grow older.

Between 1990 and 2005, when health care spending also increased by 65%, the health care labor force increased by 52%. Based on this experience and on trends that link the growth of health care workers and the long-term growth of physician supply, it can be estimated that, by 2030, the per capita demand for health care workers will increase by another 50%, and the demand for physicians will increase by more than 25%.

Can these demands be met? Sadly, the answer is no (Fig. 2). Too much time has passed with no attention to the future needs for physicians. Any increases would require not only that the current caps on Medicare support for residencies be lifted, but that appropriate sites for training be identified and that training programs fulfill the necessary academic requirements. An initial goal might be to increase the number entry-level residents by 500 (2%) annually, yet even that would fail to meet near-term needs (Fig. 5). Doubling that to 1000 annually would do more but not enough. Health care reform failed to address even minimal increases in physician supply, and.
while it addressed the need for more primary care physicians by increasing Medicare and Medicaid reimbursement and offering some incentives for trainees, this would simply decrease the numbers of specialists who are also needed. It is a zero-sum game. That leaves the United States constrained to a supply of physicians that, in per capita terms, is roughly similar to the current level for decades to come.

THE LEGACY OF HEALTH CARE REFORM

Many hope that health care reform will lead to better, more accessible, and more equitable health care within narrower financial parameters and tighter regulations. Besides, there are too many specialists, and more primary care could reduce demand even further. And, so they tell us, if physicians in high cost areas would simply become as efficient as those in lower cost areas, everything will work out. There is no need for more physicians, certainly specialists. As Lewis Carroll’s Queen told Alice, “Sometimes I’ve believed as many as 6 impossible things before breakfast.”

The truth is that legislation cannot reduce the overall rate of health care spending. The amount spent in the United States and in every other developed country is a reflection of its economic capacity. The current recession has slowed health care spending, just as the boom of the 1990s accelerated it, both unrelated to “need.”

Over time, the nation does not spend more on health care than it can afford, at least for very long, nor does it spend less. When resources materialize, the existing reservoir of unmet need consumes them, and if that were not to occur, the discovery of still more beneficial services will. But the unmet need can never be met. The ability to provide beneficial services wherever needed cannot be satisfied, even it our rich nation.

The biggest driver of excess need is poverty. The added utilization of health care by low-income patients accounts for more than one-third of all health care in the United States, and the same is true in Britain. The United States cannot pour health care dollars into both its scientific successes and its social failures. We can only do more of the former if we can do less of the latter. In that respect, real health care reform is social reform—education, housing, nutrition, social support networks, and jobs.

Although Health Care Reform will assuredly increase access for the poor, the states will struggle to accommodate the added costs of that access, particularly Medicaid (which will cover half of the newly insured). But Health Care Reform legislation will decrease disproportionate share (DSH) payments to hospitals for the added costs of caring for Medicaid patients and others who are poor. Moreover, based on geographic differences in care that have been attributed to waste and inefficiency rather than to poverty, Health Care Reform reduces payments to hospitals with “excess” readmissions, which occur principally among poor patients. The legislation also provides bonus payments for regions with lower Medicare spending, which tend to have less poverty. A broad umbrella has been stretched across the poor, but it is thin and leaky.

Finally, there is virtually nothing in the legislation that will assure an adequate supply of physicians. Unused residencies will be made available in a restricted manner, favoring primary care and general surgery, but the number is small and the process complex. Even before health care reform, the nation was headed for a serious physician shortage, and reform has only made it worse. Without an adequate supply of qualified physicians, the fundamental goals of health care reform cannot be achieved. Indeed, the health of the nation will be at risk.

REFERENCES


