

Is Health Care Ready for Six Sigma Quality?*

MARK R. CHASSIN

Mount Sinai School of Medicine, New York

CONCERNS ABOUT THE QUALITY OF HEALTH CARE in the United States have recently emerged from many different quarters. A presidential commission concluded that quality problems often cause impaired health (President's Advisory Commission 1998). The Institute of Medicine's National Roundtable on Health Care Quality concluded that "serious and widespread quality problems exist throughout American medicine" (Chassin, Galvin, and the National Roundtable 1998). A recent article in *Consumer Reports* highlighted quality problems for a wider audience (Lieberman 1998). Finally, health care quality is a prominent and recurring topic in the nationwide debates about the perceived adverse effects of managed care (Miller 1997).

This article explores critical underlying causes of quality problems, discusses some of the most salient obstacles to improvement, and suggests the components of an effective strategy to increase the pace and scope of quality improvement in the delivery system.

*The views expressed in this article are solely those of the author and not necessarily those of the members of the Institute of Medicine's National Roundtable on Health Care Quality or its sponsoring organizations.

The Milbank Quarterly, Vol. 76, No. 4, 1998
© 1998 Milbank Memorial Fund. Published by Blackwell Publishers,
350 Main Street, Malden, MA 02148, USA, and 108 Cowley Road,
Oxford OX4 1JF, UK.

Defining and Measuring the Quality of Health Care

The Institute of Medicine's definition of quality has proved of enduring usefulness: "Quality is the extent to which health services for individuals and populations increase the likelihood of desired health outcomes and are consistent with current professional knowledge" (Institute of Medicine 1990). Many reliable and valid measures of quality have been developed, building on this definition. In general, valid quality measures assess either processes (diagnostic or therapeutic interventions) or outcomes (health states that people experience). Process measures are valid quality measures when their relation to important health outcomes has been proved. The frequency with which heart attack survivors receive beta blockers is a valid quality measure because these medications improve survival in this clinical situation.

For a health outcome to be a valid quality measure, it must be related conclusively to a process or group of processes that can be modified to improve the outcome. Thus, the number of babies born with HIV infection is a valid measure of quality of care because treatment in the prepartum period with zidovudine has been proved to reduce the transmission of infection from mother to infant. Cardiogenic shock, on the other hand, has not been proved to respond to specific treatment regimens; therefore, deaths from that cause are not valid measures of health care quality.

The Six Sigma Challenge

Many careful research studies have used valid measures of quality to investigate the nature and magnitude of specific quality problems. Quality problems may be classified into three categories: overuse, underuse, and misuse (Chassin 1991; Chassin, Galvin, and the National Roundtable 1998). As the research literature makes clear, quality problems of all three varieties abound in American medicine (Schuster, McGlynn, and Brook, this issue; President's Advisory Commission 1998). The majority of these problems are not rare, unpredictable, or inevitable concomitants of the delivery of complex, modern health care. Rather, they are frighteningly common, often predictable, and frequently pre-

ventable. Viewed by those companies that have committed themselves to the most advanced applications of industrial quality management, the magnitude of the failures or quality defects in the provision of health care must seem stupefying. A few examples will highlight this contrast. Motorola and General Electric, among others, have set reliability goals for the manufacture of their products and services that they describe as the quest for Six Sigma Quality.

Motorola invented this strategy, which is named for a statistical measure of variation (the standard deviation of a normal distribution). Simply put, adopting the goal of Six Sigma quality means setting tolerance limits for defective products at such high levels that fewer than 3.4 defects occur per million units (or opportunities). These limits are set to include all observations within 6 standard deviations of the mean. Setting tolerance limits at lower levels of sigma results in higher rates of defects (table 1). Advocates of this approach to quality claim that it works just as well in service industries as in manufacturing (Harry 1998). A defect rate may be defined in whatever terms are sensible for the process that is being improved. It may refer to the number of parts (per million produced) for an aircraft engine that fail to meet all the mechanical specifications for inclusion in the finished product. A defect may also be defined as the number of telephone calls from customers that go unanswered after three (or four or five) rings (per million calls). In health care, defects might be defined as the number of two-year-olds who are not completely immunized (per million two-year-olds in the population). Another might be the number of pregnant women failing to receive prenatal care in the first trimester (per million pregnancies). A third might be the number of patients with clinical depression who are not diagnosed or well treated (per million patients with depression).

Simply setting the goal of reducing defects to 3.4 per million (or fewer) does not guarantee that it will be achieved. Allied-Signal, which began its Six Sigma program in 1994, claims that most of its manufacturing plants operate in the range of 3.5 to 4 sigma and that three model factories have already achieved 6 sigma levels of quality (Jackson 1997). General Electric improved from 3 to 3.5 sigma in the first 22 months of its program, reducing the average frequency of defects from 67,000 to 23,000 per million (Harry 1998). Although this strategy was first applied to manufacturing processes, Motorola, General Electric, and others have extended its applications to aspects of their businesses that involve direct customer services (Behara, Fontenot, and Gresham 1995;

TABLE 1
Selected Health Care Quality Problems Viewed as Defects per Million
Compared with Quality Performance in Selected Industries

Sigma level ^a	Defects per million opportunities	Selected health care examples ^b	Selected industrial examples ^c
6	3.4	—	Allied-Signal: 3 model factories
		—	Publishing: one misspelled word in all the books in a small library
	5.4	Deaths caused by anesthesia during surgery	—
	10–16	—	2 Siebe plants in Italy and United Kingdom, making temperature controls for refrigerators
5	230	—	Airline fatalities
4	6,210	—	Airline baggage handling Restaurant billing
	10,000	1% of hospitalized patients injured by negligence	—
3	66,800	—	Publishing: 7.6 misspelled words per page in a book
	210,000	21% of ambulatory antibiotics for colds	—
2	308,000	—	—
	580,000	58% of patients with depression not detected or treated adequately	—
1	690,000	—	—
	790,000	79% of eligible heart attack survivors fail to receive beta blockers	—

^aFor sigma levels of defects, see Walmsley (1997).

^bFor a discussion of health care examples, see text.

^cFor industrial examples, see Behara, Fontenot, and Gresham (1995) and Jackson (1997).

Walmsley 1997). One of General Electric's divisions reporting results in such areas produces industrial diamonds. In the two years since it began its Six Sigma program, on-time deliveries improved by 85 percent and billing mistakes fell by 87 percent (Melymuka 1998).

How does health care stack up? Table 1 shows the level of defects per million corresponding to different sigma levels and displays examples of health care quality studies that have documented the incidence of specific problems. One study showed that only 21 percent of eligible elderly heart attack survivors were taking beta blockers following their illness, a treatment that has been shown to save lives (Soumerai et al. 1997). This amounts to a defect rate of 790,000 per million, or less than 1 sigma. In that study, patients who did not receive the drugs experienced a 75 percent higher death rate than those who did. Another study showed that 58 percent of patients with clinical depression were either poorly evaluated or inadequately treated, a defect rate of 580,000 per million (Wells et al. 1989). Recent reports have documented that 21 percent of all antibiotics prescribed to ambulatory patients are used to treat colds and other viral respiratory infections, conditions for which they are useless (Gonzales, Steiner, and Sande 1997; Nyquist et al. 1998). This one inappropriate practice represents a defect rate of 210,000 per million uses of ambulatory antibiotics. The Harvard Medical Practice Study estimated that hospitalized patients were injured because of negligence in about 1 percent of all admissions (Brennan et al. 1991), a figure that was characterized as comfortingly low by some observers when the study appeared. To Motorola, however, these failures translate into 10,000 defects per million, 3,000 times worse than the Six Sigma goal.

The one health care specialty that has reduced serious defects to rates that are close to 3.4 per million is surgical anesthesia. In the 1970s and 1980s, deaths related to anesthesia occurred at rates of 1 in 10,000 to 20,000—or 25 to 50 per million (Ross and Tinker 1994). Through a variety of mechanisms, including improved monitoring techniques, the development and widespread adoption of practice guidelines, and other systematic approaches to reducing errors, it is now estimated that anesthesia deaths may now be as rare as 5 per million cases (Lunn and Devlin 1987; Eichhorn 1989; Orkin 1993).

If the performance of certain high-reliability industries, whose standards of excellence we take for granted, suddenly deteriorated to the level of most health care services, some astounding results would occur.

At a defect rate of 20 percent, which occurs in the use of antibiotics for colds, the credit card industry would make daily mistakes on nine million transactions; banks would deposit 36 million checks in the wrong accounts every day; and deaths from airplane crashes would increase one thousandfold.

Underlying Causes of Quality Problems

Why do we have such serious quality problems in health care? Some have suggested lack of information as the main reason (Wennberg, Barnes, and Zubkoff 1982; Wennberg 1987). If only physicians knew the latest scientific evidence concerning the effectiveness or ineffectiveness of specific interventions (and if only there was more of it), the right things would get done more often (and the wrong things would be more often avoided). I believe that the underlying forces at work are far more complex and much more difficult to remediate. In addition, the fundamental causes differ, depending on which class of quality problem one considers.

Overuse

Providing a health service when its risk of harm exceeds its potential benefit constitutes overuse. Perhaps the most frequently cited causative factor in discussions of overuse is fee-for-service (FFS) payment. The evidence clearly implicates payment incentives as an important cause of increased utilization. The Medical Outcomes Study showed FFS payment to be independently associated with increased utilization after adjusting for differences among patients (Greenfield et al. 1992). Miller and Luft's review of studies evaluating the impact of managed care revealed that physicians in managed care plans used an average of 22 percent fewer expensive procedures, tests, and treatments of various sorts than those in FFS plans (Miller and Luft 1994). Studies show that physicians with X-ray diagnostic facilities in their own offices or ownership interests in physical therapy or radiation therapy facilities use more services than their counterparts who do not have such arrangements (Hillman et al. 1992; Mitchell and Scott 1992; Mitchell and Sunshine 1992). Although the evidence linking FFS payment to in-

creased utilization may be clear, surprisingly the link to overuse has not been proved. No study has established formal overuse or appropriateness criteria for specific procedures and then compared the extent of inappropriate care or overuse in FFS arrangements with other types of financing.

A less well appreciated, but probably more important, factor leading to overuse is enthusiasm—the degree to which physicians and other purveyors of specific health services become passionate advocates for the services they provide, instead of objective caregivers, whose recommendations are governed strictly by scientific evidence of efficacy. The influence of this force is profound (Chassin 1993a). It is often rooted in the pleasure experienced by physician specialists after they have skillfully performed a particular procedure or technical task. Nicod and Scherrer (1992) supply their own insight into this phenomenon with a description of how much fun it is to perform coronary angioplasty. Their letter details the “very questionable” conjunction of FFS reimbursement, the fun of performing this procedure, and the insidious effect of the self-referral process that permits the cardiologist who assesses the patient’s condition and recommends the procedure to be the one who then performs it.

This phenomenon helps explain how inappropriate, or unnecessary, services actually occur. Very few physicians would knowingly and repeatedly provide services under circumstances that they believed would be more likely to result in harm than in a health benefit. Enthusiasts believe they are doing good for patients, often despite considerable evidence and a consensus to the contrary. This misplaced zeal also partially explains why overuse is so resistant to information-based approaches to solution. Enthusiasts are believers; they are not uncertain about what they do. On the contrary, they are convinced that their services provide important benefits. Medical evidence is rarely so unambiguous that enthusiasts are unable to find a justification for continuing to believe in the effectiveness of what they do. Thus, it is perhaps not so surprising that a Rand study found that 16 percent of hysterectomies performed in a group of seven managed care plans were inappropriate, at a rate ranging from 10 percent to 27 percent among plans (Bernstein et al. 1993). This level of overuse occurred despite the financial incentive for the plans to reduce utilization. Changing payment incentives alone is not likely to dampen the kind of enthusiasm that leads to overuse.

Another cause of overuse is related to the way patient referrals to specialists frequently occur. When primary care physicians refer patients for specialty care, they often do so in the expectation that a particular diagnostic or therapeutic procedure will be performed (e.g., coronary angiography, upper gastrointestinal tract endoscopy, knee arthroscopy). Specialists are under some pressure in this situation to function like technicians: that is, to perform the requested procedure, instead of conducting a thorough and independent assessment of the necessity for the intervention. Specialists in both FFS and managed care settings must be careful not to alienate their sources of referrals. They may thus be tempted to perform the requested procedure even if the indications are not clearly appropriate. Primary care physicians do not do their patients a service when they abdicate responsibility for this decision making and leave it entirely to the specialist.

Various social and cultural factors add to the complex reasons for overuse. Americans are activists. We expect that our doctors will “do something” when we present with symptoms of illness. This expectation frequently takes the form of pressures from patients for specific interventions (e.g., antibiotics for colds). Ubiquitous advertising creates the presumption that the slightest malady may be easily relieved with the right pill or potion. Whether the advertising created the expectation, or the cultural bias drove the advertising, the effect is the same. An American patient does not care to hear these words of advice: “Your symptoms are very likely to be innocuous; you don’t need any blood tests or X-rays, consultants or prescriptions. You can do a few simple things yourself at home to cope with this discomfort.” As difficult a message as this is for American patients to hear, it is equally difficult for American doctors to deliver. How much easier it is to write the expected antibiotic prescription for the viral respiratory infection, or to order the laboratory test or X-ray, instead of taking the time to explain why watchful waiting is the safer, more appropriate, strategy.

Related to this cultural propensity is our national infatuation with technology. Scientific discoveries are routinely reported with exaggerated claims in the media, and patients expect that the latest machine or pill or surgical procedure will be used to treat their conditions. Questions about safety, about appropriateness, about how experienced an individual physician might be with a technique that requires lengthy training are almost never asked. This infatuation is aggravated by the large financial incentive faced by developers of new drugs and treat-

ments to speed their products to market and to minimize the process of evaluation. For interventions like coronary angiography and angioplasty that face no regulatory requirement for demonstrating efficacy prior to their use, researchers struggle to produce rigorous evidence of effectiveness decades after they have swept through the delivery system. It is hardly surprising, therefore, when editorial writers conclude that physicians treat heart attacks and unstable angina far too aggressively in the United States, “despite the absence of scientific support for such an approach . . . : A substantial number of patients with acute coronary syndromes [heart attack and unstable angina] undergo coronary angiography and revascularization without a clear indication” (Lange and Hillis 1998).

Reinforcing these social and cultural proclivities is physicians’ fear of the malpractice lawsuit. Despite scant evidence documenting that this factor plays a significant role in producing overuse of services, there is little doubt of its impact, at least in augmenting the other, previously described factors. Malpractice is a convenient scapegoat for physicians whenever the subject of overuse is discussed. Despite the fact that almost all malpractice suits are initiated because patients failed to receive needed diagnostic or therapeutic interventions (underuse) or because of a negligent act committed during the performance of an appropriate intervention (misuse), many physicians appear to believe nevertheless that their risk of suit can be diminished by performing unnecessary services (overuse), a practice that has been termed “defensive medicine.”

Overuse is a particularly intractable problem because of its array of underlying causes, coupled with the general public ignorance of the fact that it is a quality problem that harms millions of Americans. A multifaceted strategy will be required to have a measurable impact on it. It is unrealistic to believe that public education, physician education, or alterations in financial incentives by themselves will solve this complex problem.

Underuse

Failing to provide an effective service when it would have produced favorable outcomes constitutes underuse. Problems of underuse result from a different group of factors, including financial barriers such as lack of insurance, the imposition of copayments and deductibles, and benefit packages that do not, for example, cover preventive care. Un-

deruse problems are harshly exacerbated by the barriers to care that face the uninsured. Lack of insurance has been demonstrated to increase the risk of death and disability and to result in the worsening of chronic disease (Lurie et al. 1986; Franks, Clancy, and Gold 1993; Franks et al. 1993). The Rand Health Insurance Experiment documented that patient copayments led to underuse of preventive care services and to a lower proportion of patients whose hypertension was adequately controlled (Lurie et al 1987; Keeler et al. 1985). Capitation payment, at least theoretically, encourages underuse, just as surely as FFS payment encourages overuse. Although the research literature is far from ideal, comparative studies of populations served by FFS arrangements and those enrolled in capitated health plans show about the same levels of underuse for a variety of therapeutic services (Wells et al. 1989; Udvarhelyi et al. 1991; Retchin and Preston 1991). Although managed care plans may provide preventive services somewhat more often than their FFS counterparts, the level of underuse in both settings is considerable (Dudley et al., this issue).

Another important, but less often recognized, reason for underuse is the rapid and recent accumulation of an enormous amount of information about what works and what does not to produce good outcomes in health care. Today, it is just not possible for an individual physician to hold in his or her head all that is necessary to practice good, evidence-based medicine. This, however, is a very recent phenomenon. Doctors did not always have a large amount of scientific information on which to base treatment decisions. It is perhaps difficult to recall that the science producing these data is a creation only of the last 50 years. The randomized controlled trial (RCT) was established as the research method of choice to produce this evidence just after World War II (Cochrane 1972). The first publication reporting the results of an RCT appeared in 1952; it described the benefits of different treatment regimens for tuberculosis (Daniels and Hill 1952).

A rough index of the rapidity with which medical knowledge about efficacy and effectiveness has accumulated can be obtained by examining the number of articles from RCTs that have been published in the peer-reviewed clinical research literature. The data in figure 1, which were collected from the automated database, Medline (as searched on June 1, 1998), illustrate the staggering rate of this rise between 1966 and 1995: from just over 100 articles per year to nearly 10,000 annually. The first five years of this thirty-year period account for

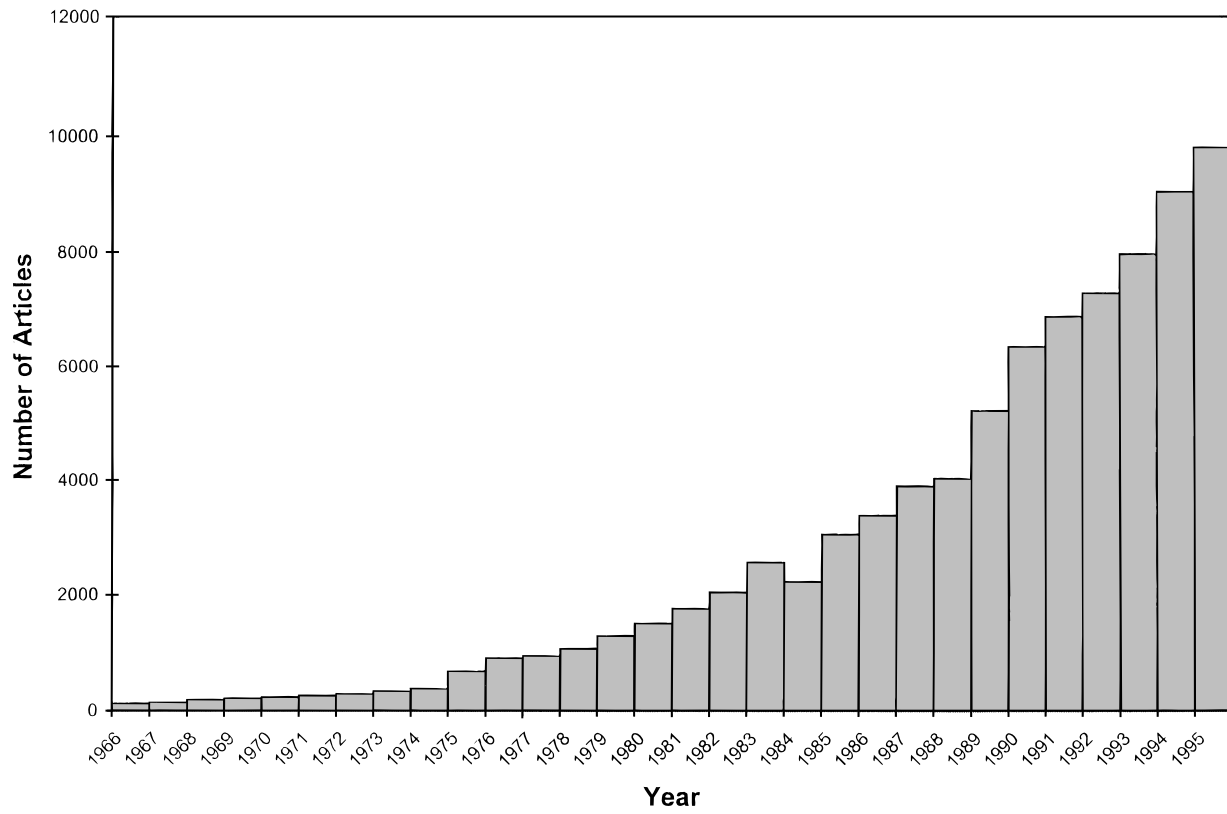


FIG. 1. Articles published from randomized controlled trials: 1966 to 1995.

only 1 percent of all the articles; the last five years account for half (49 percent).

The volume and complexity of this information pose several problems for physicians and medical practice. The sheer number of specific interventions that good care requires is beyond the ability of any unaided human being to recall and act on effectively. Yet the dominant modes of practice still expect this impossible degree of accomplishment. The complexity of these studies has also evolved considerably from the time when a new treatment was matched in a simple test against a placebo. A host of issues has now rendered many of these studies very difficult for the average physician to assimilate. These include issues of patient selection and the biases that process may introduce, multivariate statistical analytic methods; absolute versus relative risk reduction reporting; and questions about how to combine the results of many studies of the same or related clinical topics. Our capacity to summarize these data in usable forms and make them readily available to physicians when needed is minuscule compared with the magnitude of the challenge. Among other factors is the slow rate at which sophisticated information systems have been adopted by hospitals, medical groups, and other providers. Such systems could facilitate the routine follow-up of patients who miss scheduled visits or tests or who could benefit from periodic reminders to persevere with a complicated course of treatment. Yet another cause of underuse is a misunderstanding by physicians and others of when specific interventions are contraindicated. Thus, childhood immunization opportunities are missed when toddlers have colds, and some heart attack survivors fail to receive beta blockers because they are smokers (but do not have significant asthma or chronic lung disease).

Misuse

Avoidable complications of appropriate health care define misuse. We know less about the causes of misuse problems than about the other two classes of quality problems. The Harvard Medical Practice Study, which examined a representative sample of 30,000 hospital admissions occurring in New York State in 1984, was the largest attempt to examine errors in the care of hospitalized patients. These researchers found that patient injuries due to negligent care were related to errors in diagnosis in 22 percent of cases, to mishaps related to noninvasive, non-drug-related treatment in 21 percent, mistakes in medication use in 12 per-

cent, technical complications of surgery in 8 percent, and surgical wound infections in 6 percent (Leape et al. 1991).

We have just begun to systematically investigate preventable complications of health care. This has not been a popular area of research. We have spent considerable resources in quality assurance activities that have been largely reactive; they take a bad patient outcome and then search for an individual to blame. However, when systematic analyses of preventable complications have been performed, they revealed that faulty systems of care are responsible for error more often than individuals. For example, when Leape and his colleagues studied medication errors and the patient injuries that resulted from them, they found errors related to the poor functioning of 15 different specific systems within the larger system of ordering and delivering medications to patients. They found that errors in the dissemination of knowledge about drugs to doctors and nurses were responsible for 29 percent of all errors that led to preventable or potential injuries due to medication use; 12 percent of all errors were due to mistakes in dose and identity checking; and 11 percent were due to mistakes stemming from the lack of specific patient information (Leape et al. 1995).

Researchers at LDS Hospital in Salt Lake City, Utah, capitalized on the fact that several specific kinds of errors causing injuries related to medications accounted for a large share of the total. Among the most important of these factors was the failure to adjust dosages for decreased renal function, age, body mass, or liver function (James 1997). They designed a series of computer programs to assist physicians in improving their antibiotic prescribing practices. The impact of these programs was dramatic: they reduced the frequency of antibiotic-related medication injuries by 30 percent, decreased the mortality of antibiotic-treated patients by 27 percent, and reduced the costs per antibiotic-treated patient by 58 percent (Pestotnik et al. 1996).

Like many underuse problems, large numbers of preventable complications in health care appear to arise from our construction of health care delivery systems. We have created systems that depend upon idealized standards of performance that require individual physicians, nurses, and pharmacists to perform tasks at levels of perfection that cannot be achieved by human beings. High-reliability industries create systems that either prevent or anticipate and compensate for the errors that normal humans make. The delivery of high-quality health care has become extremely complex, in part because of the rapid accumulation of new knowledge

and the creation of new, effective treatments. Health care should adopt the lessons about avoiding or minimizing the impact of human error that have been learned by high-reliability industries facing similar kinds of problems. .

Meeting the Challenge of Health Care Quality Improvement

The challenges of quality improvement in health care are easy to identify but difficult to achieve. Simply put, we must:

1. always provide effective care to those who could benefit
2. always avoid providing ineffective services
3. eliminate all preventable complications of health services (Chassin, Galvin, and the National Roundtable 1998)

If we achieve these objectives at Six Sigma levels of reliability, we will improve health dramatically and ensure that we maximize the value of our massive yearly expenditures on health care. Fundamental changes must occur before we can even imagine achieving these objectives.

Medical Education and Training

Although the following comments concern the education and training of physicians, they also apply to most health professionals. For many decades we have employed a model of educating physicians that emerged in the nineteenth century. We assume we know, and can impart during medical school, a finite body of facts that all medical students must know. Following medical school, we rely on an apprenticeship method, in which each succeeding cohort of residents is taught primarily by the one just ahead of it. We rely on “master clinicians”—the teaching attending physicians—to impart their clinical wisdom patient by patient, as the young physicians-in-training gather closely around the bedside imbibing the heady brew. This process trains physicians to be the only important decision makers, rather than to function as a part of a team of caregivers. As such, physicians are expected to know the right answers to all questions about clinical care and to perform their work perfectly.

The premises on which this model were constructed are now so clearly faulty that they require a drastic overhaul. There is, first of all, no finite and immutable body of facts to impart to medical students. Physicians face a far more complicated reality of constantly changing and increasing medical knowledge. What they require to practice effective, high-quality medicine is not an encyclopedic memory but, rather, the skills to acquire the specific pieces of information that are necessary to make clinical decisions when they need to make them. Physicians also require the analytic skills necessary to review critically reports of clinical research data and claims of efficacy and effectiveness. Finally, physicians must learn that other health care professionals play crucial roles in the delivery of high-quality care—a team approach is far more effective than individuals acting alone. Learning how to function as an effective member of the health care team should be another goal of medical education. These essential skills cannot be taught in the context of a system that pretends that the “master clinician,” whom young physicians are to emulate, keeps all requisite knowledge in his or her head and uses that knowledge to make perfect decisions in every clinical situation.

A more effective educational model would recognize that the knowledge base required to practice high-quality medicine is growing at unprecedented speed (see fig. 1). This phenomenon requires that physicians be taught where to find the new knowledge that will emerge year by year after they graduate, how to evaluate its significance, and how to decide in what form to incorporate it into their practice. In essence, physicians must be trained to become lifetime learners. The days in which physicians could recall the clinical teachings of their instructors from many years previously, apply them to their current patients, and expect to practice good medicine are gone.

Finally, and perhaps most important, physicians must be committed not only to learning throughout their careers—most have recognized this obligation in one form or another for many years—but also to the challenge of continuously assessing and improving the health care they provide. This commitment must begin in medical school with experiences, shared by students, residents, and attending physicians, in which questions are formulated from the evidence base of medicine (e.g., when are inhaled steroids the most effective treatment for asthma?), are applied to develop treatment programs for individual patients (let’s make sure Peter goes home with a steroid inhaler and is trained in how to use it), and become the basis for studying the quality of the entire practice

(let's see if we prescribed this class of drugs for all of our asthma patients who would benefit). Medical schools should teach these skills to students, who then observe their daily application when they take part in patient care, and residency programs should reinforce the practices by incorporating quality measurement and improvement projects into the regular patient care routine.

Some medical schools are beginning to alter their approach to medical education. Some are reducing lecture time in favor of small-group teaching because this mode of study is popular with students. Changing how the curriculum is taught may be beneficial, but pedagogical adjustments alone will not solve the problem. Medical education must emerge from the old, and now bankrupt, model of experts teaching facts, to a new model in which facilitators train young physicians in the skills they will need for a lifetime of knowledge acquisition, analysis, and continuous quality measurement and improvement.

Health Care Delivery System Change

Crucial as these changes in education and training are, they will not be effective in putting health care on the road to Six Sigma quality unless we support these quality-driven professionals with new systems of care. How can health care learn from the airline industry, Motorola, General Electric, and other high-reliability performers? We could take a major step forward by abandoning the expectation that physicians, nurses, and other clinicians will perform perfectly. We have much to learn from the disciplines of cognitive psychology and human factors research about the causes of human error and how errors can be prevented or blocked from causing harm. Leape's analysis of the implications of applying these lessons to medicine is especially insightful (Leape 1994). Gaba (1989) has discussed their usefulness in anesthesia. The airline industry reacted to studies in the 1970s showing that more than 70 percent of crashes involved some measure of human error by developing standardized training programs to train aircraft crew in better teamwork (Helmreich 1997).

Clinicians should welcome support systems that facilitate the conduct of their clinical care. Some applications of clinical practice guidelines represent rudimentary beginnings. Done well, practice guidelines should contain all the necessary elements of routine care for most individuals with a particular condition. They should prompt the physician

to consider what specific characteristics of each individual patient might warrant departures from the guidelines. Implemented well, such systems save clinicians time. They should be assisted by computerized systems that, among other functions, can catalogue past histories, check orders for medications against measures of liver and renal function, and schedule reminders for screening tests.

Many physicians are said to resist clinical practice guidelines, fearing that they will lead to “cookbook medicine,” or the imposition of rigid rules on clinical practice that fail to account for individual patient variation. In the context of quality improvement, however, guidelines should be used, not in this way, but rather as prompts for physicians to consider all the elements of routine care (e.g., order pneumococcal vaccination for the elderly patient admitted with heart failure) and as a means to eliminate some of the more tedious tasks that often give rise to human error (e.g., adjusting the medication dose for weight or renal function) while still allowing departures for sound clinical reasons (e.g., a rare comorbid condition not contemplated by the guidelines). They should be part of the continuous improvement of systems of care. Guidelines will not be perfect at the outset: systems that use them must be constructed so that experience can be applied to improve the guidelines, just as the guidelines indicate where care delivery can be improved (James 1993; Chassin 1993b).

Clinicians will find such systems immensely helpful if they are trained in quality measurement and quality improvement and accept as a condition of medical practice that they will always be engaged in measuring and improving their clinical practice. It is impossible to imagine an airline pilot objecting to “cookbook flying” when asked to use a pre-flight checklist. It is equally unimaginable that a pilot would not have data instantly available to tell him exactly where his plane is at every instant and how that position compares with the flight plan. Pilots are no more prone to error than doctors or nurses; the systems in which they perform their jobs, however, are constructed to prevent error and to anticipate and compensate for the inevitable errors that are not averted. Some medical systems recognize this generic problem. Two nurses, for example, usually check vital information before a blood transfusion is administered. What we have failed to appreciate in health care, however, is the extent to which this philosophy, rigorously applied, can reduce the frequency of error from the extraordinarily high rates we observe in health care today to the minuscule levels of Six Sigma.

Thus, we require the development of hundreds, perhaps thousands, of new systems to guide the provision of virtually every aspect of health care. The more complicated the care, the more in need it is of new systems. Nor should such systems be designed to correct only misuse problems, where the focus is today. Rather, they must address all quality problems. Consider a woman with early-stage breast cancer, a condition with a five-year survival of greater than 90 percent if high-quality care is provided. At present, in most instances, the woman first receives at one location the mammogram that identifies the suspicious lesion. She is then referred to a surgeon, who performs a lumpectomy at another location; she goes on to visit a radiation oncologist, who performs radiation therapy at still another site; and she finally sees a medical oncologist, who administers multidrug chemotherapy at yet another, separate, location. Her care throughout is followed by a primary care physician from a separate office. Nor can she be confident that crucial information will flow reliably from one stage of the process to the next about exactly what was done, what was found, and what was planned. Indeed, the patient is often expected to be the information courier in these settings.

Few, if any, organized delivery systems can claim to measure and assure high quality effectively at each critical point in the provision of breast cancer care. Some important questions are usually unaddressed:

- How completely and objectively is information provided to women so they can choose the treatment plan (mastectomy versus lumpectomy plus radiation) that is best for them?
- How well was the initial surgery done? Were tumor margins clearly marked and examined pathologically to ascertain whether re-excision was necessary?
- Was a timely referral made to radiation therapy and was the full course of treatment completed?
- Was the patient evaluated for tamoxifen or multidrug chemotherapy, or both?
- Was treatment begun and carried out in a timely manner?

This is only a sample of the crucial issues that must be addressed to discover whether women are receiving the full benefit of high-quality breast cancer care. Few health plans, hospitals, or other organized systems of care have conducted this kind of thorough assessment. It cannot be surprising, therefore, when research shows that many women fail to

receive effective care, or that quality varies dramatically among settings of care. In New York, for example, women who received their initial breast cancer surgery at hospitals performing fewer than ten breast cancer operations per year were 60 percent more likely to be dead five years later than women who received their initial procedures at hospitals performing more than 150 such operations annually (Roohan et al. 1998). These survival rate differences, which were adjusted for differences in age, stage of disease, and comorbidity, cannot be the result of the immediate effects of surgery on mortality. Instead, they suggest failures at every step of the complex pathway that women must traverse to receive the full benefits of effective breast cancer care. The new systems of care that we need must focus on attaining the Six Sigma goal of quality by measuring and improving the care delivered at each step for every condition we can treat effectively.

Moving Health Care toward Six Sigma Quality

If the Institute of Medicine Roundtable's call for "urgent action" to improve health care quality is taken seriously, how can we move the entire health care delivery system toward improving quality? What public policy approach is most likely to produce this outcome? This was the charge taken up by the IOM conference on quality improvement, which stimulated the articles appearing in this volume. In my view, the conference and the papers it spawned demonstrate that no single approach is likely to succeed. The conference assessed the effectiveness of total quality management, marketplace competition, regulation, and payment incentives and found each strategy both promising and wanting at the same time. An integrated strategy is called for, one that takes advantage of the strengths of each approach and compensates for its weaknesses (Chassin 1997).

There is, however, a prerequisite for any such public policy strategy: increased public awareness of just how much in need of improvement our health care system is. At present, the level of public understanding of this issue is low. Health plans, hospitals, and medical groups will not make the kind of investments required to create the new systems of care we need unless consumers demand better performance from their health care providers. Employers will not press their suppliers of health care for

higher quality unless their employees insist upon it. Legislators will not appropriate the necessary public funds or engage in effective regulation unless they perceive a groundswell of public support. Thus, an immediate, large-scale effort is required to educate the public and leading representative organizations about the deficiencies in health care quality. Knowledgeable academics and other health care organizations, consumer organizations, and civic bodies, such as foundations whose mandates are to address health care issues, should seize the initiative in this educational campaign.

Each of the major strategies analyzed by the authors contributing to this collection of articles has an important role in carrying out health care quality improvement. Arguably, the greatest improvement in quality can be achieved through professionally promoted and orchestrated activities that are conducted at sites of care (hospitals, medical groups, and nursing homes). As the other contributors have described, however, such efforts are notably sparse. Many have cost reduction, rather than quality improvement, as their primary objective. Few have succeeded in enlisting the enthusiastic support of the majority of clinicians (Chassin 1996). No strategy to reach the Six Sigma goal of quality can hope to succeed without intensifying these efforts. Medical professionals should become leading advocates for making quality a central focus of every provider's practice. Quality improvement efforts should focus on measuring and improving the three major categories of problems: overuse, underuse, and misuse. Leading institutions should publicize which problems they are measuring, the action they are taking to improve, and the evidence of their achievement. Consumers, purchasers, and benefits consultants should be pressuring providers at all levels of the delivery system for this kind of information.

Jane Sisk, writing elsewhere in this volume, points out that marketplace competition has yet to prove itself an important force for improving quality. Purchasers until now have largely sought low-cost health care (Hibbard et al. 1997; Bailit 1997), rather than care of high quality. When businesses adopt the Six Sigma strategy, they do so because it makes business sense. They believe that the requirement of a considerable initial investment will be rewarded by a substantial return on that investment, measured in increased profitability and market share (Walmsley 1997). In order to turn competition into an important force in quality improvement in health care, we must create the market conditions (which include the regulatory environment and the state of con-

sumer knowledge) that motivate health plans and providers to pursue quality improvement because it is in their business interest to do so.

Increased demands on employers from their employees, resulting from heightened public awareness of quality problems, could be an important ingredient in bringing about this change. Employers might then insist more vigorously that health plans and providers fully disclose their quality improvement activities and document the results. This approach to the use of purchasing power is analogous to the industrial quality model of working together with suppliers over the long term to improve the quality of their products. A public that is more aware of extant quality problems might also begin to use these new data to select their plans and providers based on quality performance.

Government must invest in producing the public goods that will facilitate professionally driven continuous improvement. These include additional research in quality measurement and improvement methods, increased support for the development of specific applications of available research to facilitate targeted improvement activities, and ensuring the dissemination of this research at the lowest possible cost. Government at the federal and state levels should also consider creative regulation that could facilitate local quality improvement (e.g., payment incentives in public health care programs based on excellence in quality) and reconsider regulatory approaches that would improve quality directly. One example would be to reduce the hazards faced by millions of consumers when they receive complex treatments at facilities that perform too few of the procedures involved to achieve optimal outcomes. Cardiac surgery, neonatal intensive care, and complex vascular surgery are just some of the treatments for which high volume has been demonstrated to be associated with better outcomes (Hannan et al. 1989; 1992; Phibbs et al. 1996). Prohibiting low-volume programs in these and other areas could improve outcomes directly by increasing the proportion of patients treated at high-volume institutions. A recent study showed a doubling of death rates following coronary bypass surgery at California hospitals that performed fewer than 100 cases per year compared with those performing 500 or more (Grumbach et al. 1995). These authors calculated that all of California's very-low-volume (fewer than 100 cases per year) cardiac surgery hospitals could be eliminated with little patient inconvenience. After such a change, the proportion of California's population living within five miles of a hospital performing this surgery would decrease from 59.1 percent to 54.3 percent, whereas

the proportion living within 25 miles of such a facility would be unchanged (91 percent).

Government can also serve as a neutral focal point for the collection and public dissemination of data on quality of care. The New York State program of collecting and publishing risk-adjusted data on mortality following coronary artery bypass surgery by hospital and surgeon would have been impossible to create or sustain without the leadership of the state health department. In that program, data are collected prospectively on every patient undergoing cardiac surgery in New York State, using common data instruments and specifications developed by the health department, under the auspices of an advisory committee consisting principally of leaders in cardiology, cardiac surgery, and medicine. These data are forwarded to the health department, which audits their accuracy, analyzes them based on a sophisticated logistic-regression risk-adjustment model, and publishes the data annually. Most important, the health department uses the availability of these comparative performance data to galvanize local efforts to improve. Many publications in the peer-reviewed clinical literature have documented various aspects of this program. Improvements were achieved because the comparative performance data prodded individual hospitals and cardiac surgery programs to improve. The data were not used by managed care companies or employers to guide purchasing decisions (Hannan et al. 1994; Chassin, Hannan, and DeBuono 1996). Recent data from the Medicare program demonstrates that, during the initial five years of this program, New York achieved the lowest surgical mortality for coronary artery bypass surgery of any state in the nation and had the third most rapid rate of decline in surgical mortality (Peterson et al. 1998). This experience represents one model for carrying out statewide or regional programs to improve quality of care, one that respected private organizations could well emulate.

All these efforts will require the expenditure of much time, energy, and resources. Private investment for adequate information systems and quality measurement and improvement applications lead the list. Public investment through the kinds of supporting governmental activities described above must accompany private efforts. Such investments, intelligently made, can produce large returns in improved health. These activities need not increase aggregate health care costs. Solving overuse and misuse problems reduces health care costs and improves quality at the same time (Chassin, Galvin, and the National Roundtable 1998).

Available evidence suggests that problems in these two areas are so extensive that solving them might allow us to save enough that we could afford to correct our underuse problems, including providing health insurance to those who currently lack it.

Is Six Sigma Quality Possible in Health Care?

Can health care approach the near perfection of Six Sigma in actual practice? Are human systems so different from others in which Six Sigma has been achieved or attempted that high levels of reliability are unattainable? Perhaps. General Electric has begun to apply the same Six Sigma methods that worked to improve its manufacturing processes to its other, more service-oriented businesses (Walmsley 1997). We believe, however, that asking this question must not be an excuse for failing to embark on the journey. Health care now frequently produces defects at rates as high as 500,000 per million—as exemplified in failures to recognize and treat clinical depression (Wells et al. 1989) or control hypertension (Udvarhelyi et al. 1991). Enough examples of improvement exist to conclude that we can do much better. We can learn a good deal from industries that are working toward the Six Sigma goal. Let's try it in health care and see how close we can get.

References

- Bailit, M. 1997. Ominous Signs and Portents: A Purchaser's View of Health Care Market Trends. *Health Affairs* 16(6):85–8.
- Behara, R.S., G.F. Fontenot, and A. Gresham. 1995. Customer Satisfaction Measurement and Analysis Using Six Sigma. *International Journal of Quality and Reliability Management* 12(3):9–18.
- Bernstein, S.J., E.A. McGlynn, A.L. Siu, et al. 1993. The Appropriateness of Hysterectomy. *Journal of the American Medical Association* 269:2398–402.
- Brennan, T.A., L.L. Leape, N.M. Laird, et al. 1991. Incidence of Adverse Events and Negligence in Hospitalized Patients. Results of the Harvard Medical Practice Study I. *New England Journal of Medicine* 324:370–6.
- Chassin, M.R. 1991. Quality of Care—Time to Act. *Journal of the American Medical Association* 266:3472–3.

- . 1993a. Explaining Geographic Variations—The Enthusiasm Hypothesis. *Medical Care* 31(5):YS37–44.
- . 1993b. Improving Quality of Care with Practice Guidelines. *Frontiers of Health Services Management* 10(1):40–4.
- . 1996. Improving the Quality of Care. *New England Journal of Medicine* 335:1061–4.
- . 1997. Assessing Strategies for Quality Improvement. *Health Affairs* 16(3):152–61.
- Chassin, M.R., R.W. Galvin, and the National Roundtable on Health Care Quality. 1998. The Urgent Need to Improve Health Care Quality. *Journal of the American Medical Association* 280:1000–5.
- Chassin, M.R., E.L. Hannan, and B.A. DeBuono. 1996. Benefits and Hazards of Reporting Medical Outcomes Publicly. *New England Journal of Medicine* 334:394–8.
- Cochrane, A.L. 1972. *Effectiveness and Efficiency, Random Reflections on Health Services*. London: Nuffield Provincial Hospitals Trust.
- Daniels, M., and A. B. Hill. 1952. Chemotherapy of Pulmonary Tuberculosis in Young Adults. An Analysis of the Combined Results of Three Medical Research Council Trials. *British Medical Journal* (May 31):1162–8.
- Eichhorn, J.H. 1989. Prevention of Intraoperative Anesthesia Accidents and Related Severe Injury through Safety Monitoring. *Anesthesiology* 70:572–7.
- Franks, P., C.M. Clancy, and M.R. Gold. 1993. Health Insurance and Mortality: Evidence from a National Cohort. *Journal of the American Medical Association* 270:737–41.
- Franks, P., C.M. Clancy, M.R. Gold, et al. 1993. Health Insurance and Subjective Health Status: Data From the 1987 National Medical Expenditure Survey. *American Journal of Public Health* 83:1295–9.
- Gaba, D.M. 1989. Human Error in Anesthetic Mishaps. *International Anesthesiology Clinics* 23(3):137–47.
- Gonzales, R., J.F. Steiner, and M.A. Sande. 1997. Antibiotic Prescribing for Adults with Colds, Upper Respiratory Tract Infections, and Bronchitis by Ambulatory Care Physicians. *Journal of the American Medical Association* 278:901–4.
- Greenfield, S., E.C. Nelson, M. Zubkoff, et. al. 1992. Variations in Resource Utilization among Medical Specialties and Systems of Care. *Journal of the American Medical Association* 267(12):1624–30.
- Grumbach, K., G.M. Anderson, H.S. Luft, et al. 1995. Regionalization of Cardiac Surgery in the United States and Canada. *Journal of the American Medical Association* 274:1282–8.
- Hannan, E.L., H. Kilburn, Jr., J.F. O'Donnell, et al. 1992. A Longitudinal Analysis of the Relationship between In-Hospital Mortality

- in New York State and the Volume of Abdominal Aortic Aneurysm Surgeries Performed. *Health Services Research* 27(4):517–42.
- Hannan, E.L., H. Kilburn, Jr., M. Racz, et. al. 1994. Improving the Outcomes of Coronary Artery Bypass Surgery in New York State. *Journal of the American Medical Association* 271:761–6.
- Hannan, E.L., J.F. O'Donnell, H. Kilburn, et. al. 1989. Investigation of the Relationship between Volume and Mortality for Surgical Procedures Performed in New York State Hospitals. *Journal of the American Medical Association* 262:503–10.
- Harry, M.J. 1998. Six Sigma: A Breakthrough Strategy for Profitability. *Quality Progress* 31(5):60–4.
- Helmreich, R.L. 1997. Managing Human Error in Aviation. *Scientific American* (May):62–7.
- Hibbard, J.H., J.J. Jewett, M.W. Legnini, et al. 1997. Choosing A Health Plan, Do Large Employers Use the Data? *Health Affairs* 16(6):172–80.
- Hillman, B.J., G.T. Olson, P.E. Griffith, et. al. 1992. Physicians' Utilization and Charges for Outpatient Diagnostic Imaging in a Medicare Population. *Journal of the American Medical Association* 268:2050–4.
- Institute of Medicine. 1990. *Medicare: A Strategy for Quality Assurance*. Vol.1. Washington, D.C.: National Academy Press.
- Jackson, T. 1997. A Black Belt in Quality: Tony Jackson Reports on the Unforgiving Demands of "Six Sigma" Process Controls. *Financial Times* (London) (Feb. 27):11.
- James, B.C. 1993. Implementing Practice Guidelines through Clinical Quality Improvement. *Frontiers of Health Services Management* 10(1):3–37.
- James, B.C. 1997. Every Defect a Treasure: Learning from Adverse Events in Hospitals. *Medical Journal of Australia* 166:484–7.
- Keeler, E.B., R.H. Brook, G.A. Goldberg, et al. 1985. How Free Care Reduced Hypertension in the Health Insurance Experiment. *Journal of the American Medical Association* 254:1926–31.
- Lange, R.A., L.D. Hillis. 1998. Use and Overuse of Angiography and Revascularization for Acute Coronary Syndromes. *New England Journal of Medicine* 338:1838–9.
- Leape, L.L. 1994. Error in Medicine. *Journal of the American Medical Association* 272:1351–7.
- Leape, L.L., D.W. Bates, D.J. Cullen, et al. 1995. Systems Analysis of Adverse Drug Events. *Journal of the American Medical Association* 274:35–43.
- Leape, L.L., T.A. Brennan, N. Laird, et al. 1991. The Nature of Adverse

- Events in Hospitalized Patients. Results of the Harvard Medical Practice Study II. *New England Journal of Medicine* 324:377–84.
- Lieberman, T. 1998. In Search of Quality Health Care. *Consumer Reports* (October):35–40.
- Lunn, J.N., and H.B. Devlin. 1987. Lessons From the Confidential Enquiry into Perioperative Deaths in Three NHS Regions. *Lancet* (Dec. 12):1384–7.
- Lurie, N., W.G. Manning, C. Peterson, et al. 1987. Preventive Care: Do We Practice What We Preach? *American Journal of Public Health* 77:801–4.
- Lurie, N., N.B. Ward, M.F. Shapiro, et. al. 1986. Special Report: Termination of Medi-Cal Benefits. A Follow-Up Study One Year Later. *New England Journal of Medicine* 314:1266–8.
- Melymuka, K. 1998. GE's Quality Gamble. *Computerworld* 32(June 8).
- Miller, R.H., and H.S. Luft. 1994. Managed Care Plan Performance since 1980. A Literature Analysis. *Journal of the American Medical Association* 271:1512–19.
- Miller, T.E. 1997. Managed Care Regulation in the Laboratory of the States. *Journal of the American Medical Association* 278:1102–9.
- Mitchell, J.M., and E. Scott. 1992. Physician Ownership of Physical Therapy Services. *Journal of the American Medical Association* 268:2055–9.
- Mitchell, J.M., and J.H. Sunshine. 1992. Consequences of Physicians' Ownership of Health Care Facilities—Joint Ventures in Radiation Therapy. *New England Journal of Medicine* 327:1497–501.
- Nicod, P., and U. Scherrer. 1992. Money, Fun and Angioplasty. *Annals of Internal Medicine* 116:779.
- Nyquist, A., R. Gonzales, J.F. Steiner, et al. 1998. Antibiotic Prescribing for Children with Colds, Upper Respiratory Tract Infections, and Bronchitis. *Journal of the American Medical Association* 279:875–7.
- Orkin, F.W. 1993. Patient Monitoring during Anesthesia as an Exercise in Technology Assessment. In *Monitoring in Anesthesia*, 3rd ed., eds. L.J. Saidman and N.T. Smith. London, U.K.: Butterworth-Heinemann.
- Pestotnik, S.L., D.C. Classen, R.S. Evans, et al. 1996. Implementing Antibiotic Practice Guidelines through Computer-Assisted Decision Support: Clinical and Financial Outcomes. *Annals of Internal Medicine* 124:884–90.
- Peterson, E.D., E.R. DeLong, J.G. Jollis, et al. 1998. The Effects of New York's Bypass Surgery Provider Profiling on Access to Care and Patient Outcomes in the Elderly. *Journal of the American College of Cardiology* 32:993–9.

- Phibbs, C.S., J.M. Bronstein, E. Buxton, et al. 1996. The Effects of Patient Volume and Level of Care at the Hospital of Birth on Neonatal Mortality. *Journal of the American Medical Association* 276: 1054–9.
- President's Advisory Commission on Consumer Protection and Quality in the Health Care Industry. 1998. *Quality First: Better Health Care for All Americans*. Washington, D.C.
- Retchin, S.M., and J. Preston. 1991. Effects of Cost Containment on the Care of Elderly Diabetics. *Archives of Internal Medicine* 151:2244–8.
- Roohan, P.J., N.A. Bickell, M.S. Baptiste, et al. 1998. Hospital Volume Differences and Five Year Survival from Breast Cancer. *American Journal of Public Health* 88:454–7.
- Ross, A.F., and J.H. Tinker. 1994. Anesthesia Risk. In *Anesthesia*. 4th ed., ed. R.D. Miller. New York: Churchill–Livingston.
- Soumerai, S.B., T.J. McLaughlin, D. Spiegelman, et al. 1997. Adverse Outcomes of Underuse of B-Blockers in Elderly Survivors of Acute Myocardial Infarction. *Journal of the American Medical Association* 277:115–21.
- Udvarhelyi, I.S., K. Jennison, R.S. Phillips, et al. 1991. Comparison of the Quality of Ambulatory Care for Fee-for-Service and Prepaid Patients. *Annals of Internal Medicine* 115(5):394–400.
- Walmsley, A. 1997. Six Sigma Enigma. *Globe and Mail Report on Business Magazine* (October).
- Wells, K.B., R.D. Hays, M.A. Burnam, et al. 1989. Detection of Depressive Disorder for Patients Receiving Prepaid or Fee-for-Service Care. *Journal of the American Medical Association* 262:3298–302.
- Wennberg, J.E. 1987. The Paradox of Appropriate Care. *Journal of the American Medical Association* 258:2568.
- Wennberg, J.E., B.A. Barnes, and M. Zubkoff. 1982. Professional Uncertainty and the Problem of Supplier-Induced Demand. *Social Science and Medicine* 16:811–24.

Address correspondence to: Mark R. Chassin, MD, MPP, MPH, Professor and Chairman, Department of Health Policy, Mount Sinai Medical Center, One Gustave L. Levy Place, Box 1077, New York, NY 10029 (e-mail: mark_chassin@smtplink.mssm.edu).