

Avoiding the Unintended Consequences of Growth in Medical Care

How Might More Be Worse?

Elliott S. Fisher, MD, MPH

H. Gilbert Welch, MD, MPH

GROWTH IS A MAJOR FEATURE OF American medicine. Over the course of this century, the proportion of the economy devoted to medical care has more than quadrupled. TABLE 1 details this growth over the past 2 decades.¹ Price-adjusted spending on hospital and physician services has doubled, while spending on home health care has increased more than 10-fold. The number of physicians per capita has increased by 50%, while the number of cardiologists has doubled and the number of radiologists has increased 5-fold. Differences of a similar magnitude are found across US communities.²

Although medical care has obvious benefits, many assume that more medical care must lead to improved health and well-being. There are theoretical reasons, however, to believe that additional growth will be associated with progressively smaller returns (FIGURE 1). The law of diminishing returns also suggests that at some point additional growth will yield no benefit (the “flat of the curve”). And while the debate about where we sit on the curve is far from settled,³⁻⁸ the theory suggests that there is some point at which additional growth might actually produce harm.

Although harm may be more likely in theory, such a broad generalization provides little guidance. In this article we explore how harm might occur. Our aim

The United States has experienced dramatic growth in both the technical capabilities and share of resources devoted to medical care. While the benefits of more medical care are widely recognized, the possibility that harm may result from growth has received little attention. Because harm from more medical care is unexpected, findings of harm are discounted or ignored. We suggest that such findings may indicate a more general problem and deserve serious consideration. First, we delineate 2 levels of decision making where more medical care may be introduced: (1) decisions about whether or not to use a discrete diagnostic or therapeutic intervention and (2) decisions about whether to add system capacity, eg, the decision to purchase another scanner or employ another physician. Second, we explore how more medical care at either level may lead to harm. More diagnosis creates the potential for labeling and detection of pseudodisease—disease that would never become apparent to patients during their lifetime without testing. More treatment may lead to tampering, interventions to correct random rather than systematic variation, and lower treatment thresholds, where the risks outweigh the potential benefits. Because there are more diagnoses to treat and more treatments to provide, physicians may be more likely to make mistakes and to be distracted from the issues of greatest concern to their patients. Finally, we turn to the fundamental challenge—reducing the risk of harm from more medical care. We identify 4 ways in which inadequate information and improper reasoning may allow harmful practices to be adopted—a constrained model of disease, excessive extrapolation, a missing level of analysis, and the assumption that more is better.

JAMA. 1999;281:446-453

www.jama.com

is both to stimulate discussion and, where possible, to provide guidance about how to avoid the unintended adverse consequences of growth. The first section highlights 2 distinct levels at which more medical care may be introduced, both of which will require attention if we are to minimize the risks of harm. The second section focuses on the mechanisms whereby harm may occur. In the third section, we turn to the fundamental challenge—reducing the risk of harm from more medical care.

MORE MEDICAL CARE— 2 LEVELS OF ANALYSIS

More medical care can be introduced at 2 distinct levels: discrete and systemic. Discrete clinical decisions entail choices to adopt in practice a specific diagnostic

Author Affiliations: VA Outcomes Group, Department of Veterans Affairs Medical Center, White River Junction, Vt, and the Center for the Evaluative Clinical Sciences, Dartmouth Medical School, Hanover, NH. **Corresponding Author and Reprints:** Elliott S. Fisher, MD, MPH, Center for the Evaluative Clinical Sciences, Dartmouth Medical School, 7251 Strassenburgh Hall, Hanover, NH 03755-3863.

or therapeutic intervention. Such decisions can often be well specified in terms of the underlying pathophysiologic perturbation to be corrected and, in general, in terms of the expected outcomes. Physicians carry the major responsibility for decision making at this level; for example, whether a patient with gallstones and nonspecific abdominal pain should have a cholecystectomy or whether a patient with back pain should undergo magnetic resonance imaging (MRI).

More medical care can also be introduced, however, at the level of the system, by adding capital resources or increasing the workforce. Decisions that influence the broad deployment of resources are often shared but are largely driven by physicians' interests; for example, whether to build a new neonatal intensive care unit or whether to recruit an additional cardiologist to the practice. Such decisions will influence whether an otherwise healthy premature infant will be monitored in the intensive care unit or how soon a patient with congestive heart failure will return for follow-up. These decisions are rarely aimed at a single pathophysiologic disturbance and have broad and less predictable effects.

At both levels, however, the current cultural and legal environments exert tremendous pressure to do more, serving to reinforce what for some remains the operating assumption that it is "safer" to monitor the infant in the intensive care

unit, to obtain more information on the underlying anatomy, to see the patient with congestive heart failure earlier, or to remove the gallbladder now rather than later. This assumption, however, ignores the possibility of harm. Evidence suggests that harm may occur from interventions at either level.

Most randomized trials are intended to demonstrate the safety and efficacy of discrete clinical interventions; occasionally they provide evidence of harm. Three examples are provided in TABLE 2. A recent trial investigated the intensity of monitoring women at risk of preterm labor.⁹ More intensive surveillance had no effect on the primary outcomes (the incidence of births at <35 weeks, cervical dilatation at the time preterm labor was diagnosed, or neonatal outcomes), but did lead to significantly more unscheduled visits and greater use of prophylactic tocolytic drugs. The second trial considered extending treatment to patients with mild arrhythmias.¹⁰ Because the benefit of antiarrhythmic therapy for patients with severe ventricular arrhythmias was apparent, it was surprising that, when class 1-C agents were used to suppress largely asymptomatic arrhythmias, the result was a 2.5-fold increase in mortality. Finally, a randomized trial compared angioplasty with medical therapy for patients with coronary artery disease, most of whom had mild disease.¹¹ After 2 years, angioplasty had reduced symptoms only in the group with severe angina, yet doubled the risk

of nonfatal myocardial infarction (MI) or death overall.

Only a few studies have examined the system-level effects of more medical care. The most comprehensive was carried out more than 20 years ago, a randomized trial of the influence of co-payments on utilization and outcomes, which found that those provided free care received about 40% more care than those with co-payments.¹² Although more care appeared beneficial in high-risk subgroups, on average the group randomized to receive more care had no improvement in function and had more pain, more worry, and more restricted-activity days. The second study compared risk-adjusted outcomes among Medicare beneficiaries who experienced MIs in New York and Texas (where rates of post-MI angiography were about 50% higher than in New York).¹³ Residents of both states were equally likely to have left main or 3-vessel disease detected and treated, but Texans with mild disease were much more likely to undergo angiography and receive revascularization. After 2 years of follow-up, Texans had significantly lower exercise tolerance, more angina, and higher overall mortality. The third study compared routine follow-up after hospitalization to structured discharge procedures and more frequent follow-up visits.¹⁴ Although intended to reduce hospitalization, the greater intensity of surveillance resulted in a 36% increase in

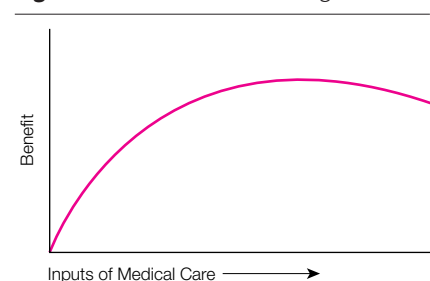
Table 1. Intensity of US Health Care: Growth Over Time and Variation Across Regions

	Growth in Capacity Over Time			Variation Across Regions, Ratio 90th/10th Percentile†
	1975	1995	Ratio 1995/1975	
Health spending per capita*				
Hospital services	\$667	\$1331	2.0	2.2
Physician services	\$304	\$767	2.5	2.0
Home health care	\$8	\$109	13.1	10.9
All personal health care	\$1454	\$3344	2.3	2.1
Physician workforce per 100 000				
All physicians	158	238	1.5	1.8
Cardiologists	2.3	5.0	2.2	5.2
Radiologists	0.9	4.6	5.3	3.0

*Spending data for 1975 and 1995 are for entire US population. Data for 1975 are expressed in terms of 1995 dollars based on the consumer price index, excluding health care.¹

†Data used to determine geographic variation in spending were for the Medicare population older than 65 years only. Variation across regions was quantified as the ratio of the 90th percentile to the 10th percentile of per capita spending or physician supply in US Hospital Service Areas.²

Figure 1. The Law of Diminishing Returns



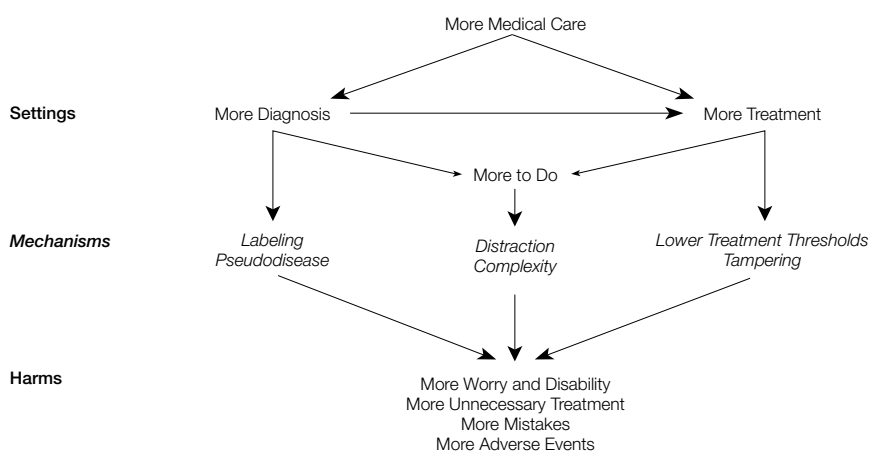
The first unit of input provides substantial benefits (imagine the first physician in a community), while additional units provide declining additional benefit (imagine the thousandth physician). Eventually, increasing inputs lead to no additional benefit (the "flat of the curve"). At some point, in theory, additional inputs lead to harm.

Table 2. Selected Studies in Which More Medical Care Led to Harm*

Input	Setting	Harm
Introduction of Specific Clinical Interventions		
Use of home monitoring	Randomized trial of 2422 pregnant women at risk for preterm labor ⁹	More monitoring caused more unscheduled visits ($P < .002$) and greater use of tocolytic drugs ($P < .01$) (which caused adverse effects in 7% of recipients), but did not change the rate of low-birth-weight infants
Addition of antiarrhythmic medications to routine care	Randomized trial of 1727 patients who had successful suppression of asymptomatic or mildly symptomatic ventricular arrhythmias following AMI ¹⁰	All-cause mortality was 2.5 times higher in the group assigned to encainide or flecainide compared with placebo ($P < .001$), and deaths due to arrhythmia were 3.6 times higher ($P < .001$)
Coronary angioplasty instead of medical management	Randomized trial of 1018 patients with CAD confirmed by angiography, 60% of whom had single-vessel disease and 47% of whom had mild or no angina ¹¹	Those randomized to PTCA had a relative risk of death or nonfatal MI of 1.92 over 2 y of follow-up ($P < .05$); an early benefit of reduced angina was found only in the subgroup with severe angina at baseline and was markedly attenuated by study end
Deployment of Additional Capacity		
Additional input of about 40% more care, both outpatient and inpatient services	Randomized trial of 3958 subjects (aged 14-61 y) comparing free care vs coinsurance ¹²	For population overall, more care led to greater levels of worry ($P = .05$), pain ($P = .05$), restricted activity days ($P < .01$), and no gain in functional status; subgroups at high initial risk (the poor) benefited
Population-based provision of more coronary angiography and revascularization	Longitudinal cohort study comparing population-based samples of patients with MI in Texas ($n = 1837$) and New York ($n = 1852$) ¹³	No difference in the detection and treatment of severe disease, but more mild disease detected and treated in Texas; Texans had lower survival and higher rates of angina
Addition of structured discharge procedures and a 70% increase in follow-up visits	Randomized trial of 1396 hospitalized patients with diabetes, heart failure, or COPD ¹⁴	Greater surveillance produced a 36% increase in readmission, the study's primary outcome ($P = .005$), and a trend toward higher mortality

*AMI indicates acute myocardial infarction; CAD, coronary artery disease; PTCA, percutaneous transluminal coronary angioplasty; and COPD, chronic obstructive pulmonary disease.

Figure 2. Pathways by Which More Medical Care May Lead to Harm



readmission rates and a trend toward increased mortality.

Some might conclude that these studies represent rare exceptions to a general rule of benefit from advances in technology and expansion of capacity. At the same time, they underscore what should be an obvious point: more medical care may at times be harmful. The next section provides a schematic overview of how harm may occur.

**MORE MEDICAL CARE—
A SCHEMATIC APPROACH TO
THE MECHANISMS OF HARM**

Whether because of the introduction of a new clinical practice or the expansion of local capacity, harm from more medical care is most likely to result from 1 of several common pathways. **FIGURE 2** outlines our framework for thinking about the potential mechanisms of harm.

More Diagnosis

There has been dramatic growth in the use and capabilities of diagnostic tests. From 1987 to 1993, the rate of coronary angiography performed on Medicare beneficiaries went up by 75%, and the rate of computed tomography (CT) or MRI scans of the lumbar spine doubled.¹⁵ Increasing diagnostic capability offers the opportunity to detect subtle abnormalities of physiology or anatomy long before they manifest the clinical signs or symptoms that would have provided the basis for a diagnosis in the past. Spiral CT can now detect hepatic lesions of 2 mm in diameter,¹⁶ a tenth the size of the lesions detectable through routine imaging studies in 1982.¹⁷ The advent of genetic testing promises that physicians will increasingly have the capacity to identify those individuals who may develop abnormalities in the future (those who are at risk).¹⁸

New diagnostic tests and lower diagnostic thresholds not only increase the observed prevalence of disease but also shift the spectrum of detected disease. Newly detected disease will in general include milder cases, subtler abnormalities, and smaller lesions and will tend to

be of lower grade and earlier stage. Symptoms will be less bothersome, sometimes absent. Even without treatment, this shift in the spectrum of detected disease will lead to an apparent improvement in patient outcomes.

And there is a lot of disease for us to find. Advanced diagnostic testing is 1 way to find more. For example, a quarter of young adults have knee abnormalities by MRI¹⁹; half have lumbar disk bulge (despite the absence of back pain).²⁰ **TABLE 3** shows the effect of selected tests in increasing how much disease is found.²¹⁻²⁴ Another way to find more is to purposely change the disease definition, such as has recently occurred when the American Diabetes Association lowered the threshold to diagnose diabetes.²⁵ The impact of changing disease definitions is also relevant for hyperlipidemia. **FIGURE 3** demonstrates that while 21.2% of US adults have “abnormal” levels of total cholesterol, given a threshold of 6.21 mmol/L (240 mg/dL), more than half, 51.2%, have abnormal levels, given a threshold of 5.17 mmol/L (200 mg/dL). The combination of enhanced capability and lower diagnostic thresholds means that in general the yield from diagnostic testing improves—providing immediate positive feedback for clinicians to pursue further testing. Even if disease burden is constant, observed prevalence will increase over time—providing further stimulus for more diagnostic testing.²⁷

Mechanisms of Harm: Labeling, Pseudodisease. As diagnostic testing increases, more patients who receive diagnoses will have no symptoms; but for the diagnostic test, they would not have known that they were sick. One way that this could lead to harm is through labeling: the effect of telling someone who feels well that he or she is sick. Although the effect of labeling is most familiar in the setting of hypertension,²⁸ it has also been reported in carriers following sickle cell screening,²⁹ among those screened for developmental delays,³⁰ and most recently among those diagnosed as having Lyme disease.³¹ A classic example involved Seattle schoolchildren who were labeled

as having “something” wrong with the heart.³² Although only 2 should have been limited in any way (both of whom were disabled by overt disease), 20% had been restricted in their physical activities and another 20% were “treated differently” by their parents because they believed their children had “heart troubles.”

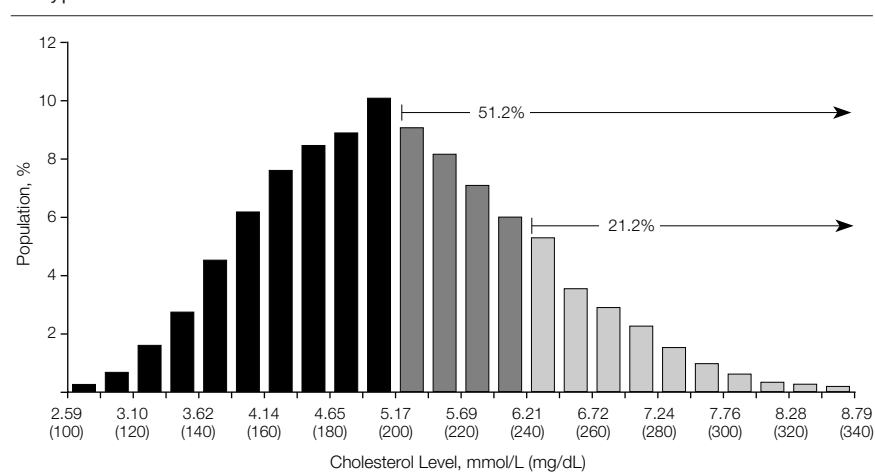
The smaller and increasingly subtle findings identified by advanced diagnostic technology create a second mechanism of harm: pseudodisease. Pseudodisease is disease that would never become apparent to patients during their lifetime without the diagnostic test. Pseudodisease is a function not only of the lesion (more likely with mild or slowly progressive disease), but also of the host (more likely as the probability of death

from other causes increases). Although it may be difficult or impossible to determine at the time of diagnosis precisely which cases constitute pseudodisease, there is accumulating evidence that pseudodisease is destined to become a common problem. It is increasingly clear that the population with an occult disease is many times larger than the population destined to become sick from it. Microscopic examination of specimens from individuals without known cancer nonetheless reveals a high prevalence of the disease: a third of adults have pathologic evidence of papillary carcinoma of the thyroid,³³ as many as 40% of women in their 40s may have ductal carcinoma in situ of the breast,³⁴ half of men in their 60s have adenocarcinoma of the prostate.³⁵

Table 3. Influence of Diagnostic Testing on Disease Prevalence

Disease Setting	Prevalence of Disease (%) Based on		Increase (Ratio)
	Traditional Test (Clinical Examination)	Advanced Tests (New Technology)	
Abdominal aortic aneurysm: 201 men with hypertension or coronary artery disease ²¹	2.5	9.0 (Ultrasound)	3.6
Thyroid nodule: 100 unselected patients ²²	21.0	67.0 (Ultrasound)	3.2
Deep venous thrombosis: 349 trauma patients ²³	0.9	57.6 (Duplex ultrasound)	64.0
Pulmonary embolus: 44 deep venous thrombosis patients ²⁴	15.9	52.3 (Ventilation-perfusion scan)	3.3

Figure 3. Effect of Changing Diagnostic Thresholds on the Prevalence of Hypercholesterolemia



If individuals with total serum cholesterol levels higher than 5.17 mmol/L (200 mg/dL) are defined as abnormal, more than half the US adult (≥ 17 years old) population is labeled as diseased. Data from National Center for Health Statistics.²⁶

Table 4. Situations in Which a Lower Treatment Threshold May Lead to Harm

Domain	Treatment Threshold Lowered So Treatment Is Provided
Spectrum of disease	To patients whose prognosis is good, regardless of treatment To patients whose prognosis is poor, regardless of treatment (futility)
Competing risks	When death is likely from other causes, even if treatment of the primary disease is successful
Patient values	To a patient who is not bothered by current symptoms
Risk of treatment	By providers whose outcomes are unknown or where higher risks may overwhelm the benefits

One manifestation of the problems caused by pseudodisease is the accumulating evidence that the findings associated with “disease” have become so subtle that physicians disagree about who has the diagnosis. Radiologists disagree about which women should be referred for breast biopsy³⁶ (and about which breast to biopsy³⁷); pathologists disagree about who has ductal carcinoma in situ^{38,39} and who has melanoma.⁴⁰ Pseudodisease leads not only to physician disagreement and unnecessary patient worry and disability, but also to unnecessary treatment.

More Treatment

Physicians rarely stop with a diagnosis. During the 1980s, both the number of surgical procedures and (price adjusted) spending on prescription drugs increased by more than 70%.¹ The number of revascularization procedures among the elderly with coronary artery disease recently increased by more than 2-fold,¹⁵ while the number of visits for children at which stimulants were prescribed to treat attention-deficit/hyperactivity disorder increased from 300 000 to 2.4 million.⁴¹ Our technical capacity to intervene is also increasing: new classes of pharmacological agents, advances in interventional radiology and cardiology, and widening applications for minimally invasive surgery.

When more care is provided, which patients receive the additional care? Patients with known chronic disease are

treated with greater intensity.⁴² Additional treatment is provided to patients who are close to death.⁴³ Treatment is also extended, however, to those with less severe disease. When more intensive care unit beds are available, the average severity of illness of those admitted declines.⁴⁴ When the numbers of hospitals and surgeons performing coronary artery bypass graft are doubled, the rate of treatment for 1- or 2-vessel disease (without involvement of the left anterior descending artery) increases 9-fold.⁴⁵

Mechanisms of Harm: Lower Treatment Thresholds, Tampering. TABLE 4 lists situations in which lower treatment thresholds may lead to harm. The spectrum of disease will have an important influence; the potential benefit of treatment is lower for people whose untreated prognosis is good. Treatment risk, however, is generally less responsive to severity of illness; those with mild disease still face substantial risks from interventions. For example, patients with less severe disease have less to gain from intervention for coronary artery disease⁴⁶ and carotid stenosis.^{47,48} If people are treated for inconsequential disease, the risks of treatment will exceed the benefits. At the other end of the spectrum, patients whose prognosis is poor regardless of therapy are also harmed when treatment of their disease is pursued.

Harm may occur in 3 other settings. The first is when a patient faces a substantial risk of death from competing causes, whether as a consequence of age or other disease. Even if these patients have true disease (rather than pseudodisease), the near-term risks of therapy can overwhelm any long-term benefits.⁴⁹ The second setting occurs when patients are prescribed treatments for symptoms that do not bother them. Measuring the level of exertion that brings on dyspnea, for example, is not the same as asking how bothered a patient is by shortness of breath.⁵⁰ There is evidence that patients are treated for symptoms that are not bothersome, even though doing so can offer no benefit and only risk.⁵¹ A third setting is when the distinction between efficacy (outcomes in ideal set-

tings) and effectiveness (outcomes in community practice) is clinically important. The net benefit of carotid endarterectomy, for example, is highly dependent on operative risk. Most hospitals, however, have operative risks that are substantially higher than those reported in the randomized trials that led to lower treatment thresholds.⁵² Even at 12 academic centers, a recent study estimated that operative risks for carotid endarterectomy outweighed the benefits in more than half of the patients.⁵³

More treatment may also result in “tampering,” a phenomenon long recognized in the statistical process control literature. Tampering occurs when an intervention is made to correct a deviation in a measure of system performance that reflects random variation (or noise) rather than systematic variation (a signal of a significant deviation). Intervening in response to random variation causes systems to become more unstable.⁵⁴ Although tampering has received little attention in the medical literature, many health care interventions are based on measurements subject to such variation. Acting on such variation creates more instability and a greater likelihood of unnecessary treatment and adverse events.⁵⁵ This phenomenon should be familiar to physicians in both the inpatient setting (for example, fiddling with ventilator settings or fluid balance) and the outpatient setting (for example, changing insulin dose or anti-hypertension medications). If pulmonary artery catheters turn out to be harmful,⁵⁶ 1 of the mechanisms could be from promoting tampering with normal physiologic variation.

More to Do

Wherever they work—the intensive care unit, the radiology department, or the outpatient clinic—physicians have more to do. Advanced diagnostic technology and more frequent testing lead to more diagnoses. And each diagnosis requires more attention, now that we are intervening to slow disease progression in such common conditions as hyperlipidemia, diabetes, and congestive heart failure. Less obvious is a potential cycle of

increasing intervention.²⁷ As more disease is found, apparent prevalence and incidence rise. At the same time, the spectrum of diagnosed disease includes increasing numbers of mild cases—leading apparent treatment outcomes to improve (via lead time and length biases), even if the efficacy of treatment is unchanged.⁵⁷ The apparent rise in disease burden combined with apparent improvements in outcome provide a powerful stimulus to do more.

Mechanisms of Harm: Distraction, Complexity. With more to do, the possibility that clinicians will miss something important becomes greater,⁵⁸⁻⁶⁰ a mechanism of harm we call distraction. Might distraction explain why so many patients fail to receive aspirin or β -blockers following MI?^{61,62} Might distraction explain why our patients complain that we spend less time listening to their concerns? As patients have more diagnoses and more treatment options, physicians will be increasingly challenged to prioritize correctly and may increasingly risk missing the forest for the trees.

An analogous mechanism operates at the level of the system: complexity.^{55,63} Consider the numerous steps required at each phase of common treatments: anticoagulation, including initiating, dispensing, monitoring, and adjusting therapy; or a laparoscopic cholecystectomy, including preparing the equipment, anesthetizing the patient, carrying out the procedure itself, and following up with postoperative care. Even if the probability of failure in any single step is low (eg, 1%), the probability of at least 1 failure rises with the number of steps (eg, with 10 steps the probability is 10%; with 100 steps the probability is 63%; with 1000 steps the probability exceeds 99.9%). As there is more to do, systems become more complex and mistakes are more likely. Evidence suggests that adverse events—many due to simple mistakes—are not rare. Adverse drug reactions occur in up to 10% of the elderly in the community⁶⁴ and 12% of all patients in the hospital⁶⁵; many lead to death.⁶⁶ Estimates of the overall rates of adverse events causing injury or death

range from 3.7% for patients hospitalized in New York State⁶⁷ to 16.6% among hospitalized patients in Australia.⁶⁸

MORE MEDICAL CARE—MINIMIZING THE RISK OF HARM

Although medicine is increasingly rooted in science, the practice of medicine will remain filled with uncertainty.⁶⁹ The fundamental task of physicians, therefore, seems destined to remain one of judgment: how best to apply scientific knowledge to the care of a specific individual such that the benefits are likely to outweigh the harms. There are 4 reasons why, however, our judgment may be impaired (TABLE 5).

1. *The way we think about disease is often inconsistent with the true state of nature.* Our patients generally understand disease as a dichotomy: they are either sick or well. Although clinicians recognize that illness severity ranges broadly (and that most chronic diseases develop over a prolonged period), we often fall back on a dichotomous model when deciding to treat: the data we have

on diagnostic tests (eg, sensitivity, specificity) assume a dichotomous model; the data we have on treatment often fail to describe how outcomes vary across the spectrum of disease. We must come to understand and make our decisions based on a more complex model that accounts for the full spectrum of disease. Our patients must come to understand that not all disease progresses inexorably to a feared outcome. To accommodate this complexity, however, we need data—data on the natural history of the increasingly mild and subtle abnormalities detected by advanced diagnostic technology and data on the benefits and harms of treatment.⁷⁰

2. *Findings drawn from narrowly framed studies are extrapolated broadly.* Understandably, treatments are often investigated under tightly controlled conditions: patients have well-defined disease (and are often at high risk for the outcome being measured) and receive a well-defined intervention (often from exceptionally experienced providers). The tendency is to extrapolate the findings to other settings: to lower-risk patients,

Table 5. Reducing the Risk of Harm From More Medical Care

Underlying Causes	Suggested Approaches
1. Constrained model of disease Harms of labeling and pseudodisease are exacerbated when disease spectrum is ignored Risks of treatment are misspecified when disease is treated as a dichotomous variable	Account for disease spectrum Provide data on the natural history of the increasingly mild disease detected by advanced diagnostic technology Evaluate the benefits and harms of treatment of mild disease
2. Excessive extrapolation From results in 1 group of patients (those eligible for trials) to others From results for 1 intervention (eg, treatment of hypertension with thiazides) to another (treatment with calcium channel agents) From results in centers that participated in clinical trials to the results in practice	Draw inferences with care Were results for patients of similar illness severity reported in the trials? Has benefit been proven for this specific intervention? Are the risks of treatment at your center known?
3. A missing level of analysis The impact of system resources on treatment and outcomes are rarely examined Quality improvement initiatives and other system changes are usually unevaluated	Evaluate the impact of the system Study the impact of changes in capacity (technical capabilities, capital resources, workforce) Hold quality improvement efforts and system change to rigorous standards of proof
4. We look for more to be better Evidence of harm from more medical care may be discounted or interpreted as absence of benefit Research focuses on identifying benefits of new technology	Acknowledge that it might not be Consider seriously the possibility that harm may occur from more medical care Pursue research on the potential harms of medical care and provide guidance on what not to do

to similar or related treatments, and to providers with unknown capabilities. To reduce the risk of harm, such inferences should be drawn with care and the burden of proof shifted to those who would promote the intervention.⁷¹

3. *Research efforts are missing a level of analysis, the system.* Although system changes are implemented daily, we know little about their effects. There are many ways to change the system and add capacity—increase visit frequency, hire ancillary health personnel, purchase new diagnostic technology, recruit a new surgeon. Such changes certainly influence costs; how they affect outcomes is unknown. Outcomes must be measured not only for discrete groups of patients, but also for the broader population that might be affected by the system change. Untoward population effects must be specifically sought. We do not know the impact of diagnostic testing on the many individuals who must be screened to find patients eligible for most treatments (about 25 people had to be screened, for example, to enroll 1 patient in the Asymptomatic Carotid Artery Stenosis trial⁴⁸). Even interventions intended to improve quality may have unexpected effects.¹⁴ System changes ought to receive the same level of scrutiny and be subject to the same burden of proof as any other intervention.

4. *We look for more to be better.* Data that are inconsistent with our underlying beliefs are often either rejected or ignored.⁷² It may not be happenstance that findings of harm in studies of specific technologies are perceived as isolated findings. It may not be happenstance that the 2 studies that suggested harm from increasing aggregate capacity were framed as showing an absence of benefit (by either the authors or the editors). The study of cardiovascular outcomes in New York and Texas, for example, reported significantly higher death rates and worse angina among residents of Texas; the abstract and conclusion, however, reported only that there was no benefit from the more invasive practice style.¹³ The trial of increased surveillance of patients with chronic disease found higher hospitalization rates but emphasized the greater

satisfaction from the more frequent contact in the intervention arm.¹⁴ Because harm from more medical care is unexpected, findings suggesting harm are discounted.

While the limited empirical basis for many widely adopted practices motivated the growth of outcomes research, the focus of our research enterprise remains the development of new technologies and treatments. Relatively little effort is devoted to evaluating current clinical practices, identifying their limitations, and advising clinicians about what not to do. Moreover, stakeholders in the increasingly market-driven US health care system have few incentives to explore the harms of the technologies from which they stand to profit.⁷³ As industry-sponsored research becomes more prevalent, the well-known problems of publication bias⁷⁴ and conflict of interest will only get worse,⁷⁵ making it even more difficult for findings of harm to surface. Public and philanthropic support for research will be essential to counterbalance those who stand to benefit from growth. To minimize the risk of harm, we must look for it.

Disclaimer: The views expressed herein do not necessarily represent the views of the US Department of Veterans Affairs or the US government.

Funding/Support: The work was supported by a grant from the Robert Wood Johnson Foundation.

Acknowledgment: We thank Donald Berwick, MD, Thomas Inui, MD, and Peter Mogielnicki, MD, for their thoughtful reviews of an early draft of the manuscript. We also acknowledge William Black, MD, and John Wennberg, MD, who have profoundly influenced our thinking on these issues.

REFERENCES

- National Center for Health Statistics. *Health, United States 1995*. Hyattsville, Md: US Public Health Service; 1996.
- Wennberg JE, Cooper MM. *The Dartmouth Atlas of Health Care in the United States*. Chicago, Ill: American Hospital Publishing Inc; 1998.
- Carlson RJ. *The End of Medicine*. New York, NY: Wiley; 1975.
- Illich I. *Medical Nemesis: The Expropriation of Health*. New York, NY: Random House; 1976.
- McKeown T. *The Role of Medicine: Dream, Mirage, or Nemesis?* Princeton, NJ: Princeton University Press; 1979.
- Beeson PB. Changes in medical therapy during the past half century. *Medicine*. 1980;59:79-99.
- Levine S, Feldman JJ, Elinson J. Does medicine do any good? In: Mechanic D, ed. *Handbook for Health, Health Care, and the Health Professions*. New York, NY: Free Press; 1983:394-404.
- Bunker JP, Frazier HS, Mosteller F. Improving health: measuring effects of medical care. *Milbank Q*. 1994; 72:225-258.
- Dyson DC, Danbe KH, Bamber JA, et al. Monitoring women at risk for preterm labor. *N Engl J Med*. 1998;338:15-19.
- The Cardiac Arrhythmia Suppression Trial (CAST) Investigators. Preliminary report: effect of encainide and flecainide on mortality in a randomized trial of arrhythmia suppression after myocardial infarction. *N Engl J Med*. 1989;321:406-412.
- RITA-2 trial participants. Coronary angioplasty versus medical therapy for angina: the second Randomised Intervention Treatment of Angina (RITA-2) trial. *Lancet*. 1997;350:461-468.
- Newhouse J. *Free for All*. Cambridge, Mass: Harvard University Press; 1993.
- Guadagnoli E, Hauptman PJ, Ayanian JZ, Pashos CL, McNeil BJ, Cleary PD. Variation in the use of cardiac procedures after acute myocardial infarction. *N Engl J Med*. 1995;333:573-578.
- Weinberger M, Oddone EZ, Henderson WG. Does increased access to primary care reduce hospital readmissions? *N Engl J Med*. 1996;334:1441-1447.
- Verrilli D, Welch HG. The impact of diagnostic testing on therapeutic interventions. *JAMA*. 1996;275: 1189-1191.
- Urban BA, Fishman EK, Kuhlman JE, Kawashima A, Hennessey JG, Siegelman SS. Detection of focal hepatic lesions with spiral CT: comparison of 4- and 8-mm interscan spacing. *AJR Am J Roentgenol*. 1993;160: 783-785.
- Tanasescu DE, Waxman AD, Drickman MV, et al. Liver scintigraphy in colon carcinoma: correlation with modified Duke pathological classification. *Radiology*. 1982;145:453-455.
- Welch HG, Burke W. Uncertainties in genetic testing for chronic disease. *JAMA*. 1998;280:1525-1528.
- Kornick J, Trefelner E, McCarthy S, Lange R, Lynch K, Jokl P. Meniscal abnormalities in the asymptomatic population at MR imaging. *Radiology*. 1990;177: 463-465.
- Jensen MC, Brant-Zawadzki MN, Obuchowski N, Modic MT, Malkasian D, Ross JS. Magnetic resonance imaging of the lumbar spine in people without back pain. *N Engl J Med*. 1994;331:69-73.
- Lederle FA, Walker JM, Reinke DB. Selective screening for abdominal aortic aneurysms with physical examination and ultrasound. *Arch Intern Med*. 1988;148:1753-1756.
- Ezzat S, Sarti DA, Cain DR, Braunstein GD. Thyroid incidentalomas: prevalence by palpation and ultrasonography. *Arch Intern Med*. 1994;154:1838-1840.
- Geerts WH, Code KI, Jay RM, Chen E, Szalai JP. A prospective study of venous thromboembolism after major trauma. *N Engl J Med*. 1994;331:1601-1606.
- Moser KM, Fedullo PF, LitteJohn JK, Crawford R. Frequent asymptomatic pulmonary embolism in patients with deep venous thrombosis. *JAMA*. 1994; 271:223-225.
- Report of the Expert Committee on the Diagnosis and Classification of Diabetes Mellitus. *Diabetes Care*. 1997;20:1183-1197.
- National Center for Health Statistics, US Department of Health and Human Services. *Third National Health and Nutrition Examination Survey, 1988-1994, NHANES III Household Adult and Laboratory Data Files [CD-ROM]*. Hyattsville, Md: Centers for Disease Control and Prevention; 1996. Public use data file documentation No. 76200.
- Black WC, Welch HG. Advances in diagnostic imaging and overestimations of disease prevalence and the benefits of therapy. *N Engl J Med*. 1993;328: 1237-1243.
- Birkett NJ, Evans CE, Haynes RB, et al. Hypertension control in two Canadian communities: evidence for better treatment and overlabelling. *J Hypertens*. 1986;4:369-374.

29. Hampton ML, Anderson J, Lavizzo BS, Bergman AB. Sickle cell nondisease. *AJDC*. 1974;128:58-61.
30. Cadman D, Chambers LW, Walter SD, Ferguson R, Johnston N, McNamee J. Evaluation of public health preschool child developmental screening: the process and outcomes of a community program. *Am J Public Health*. 1987;77:45-51.
31. Reid MC, Schoen RT, Evans J, Rosenberg JC, Horwitz RI. The consequences of overdiagnosis and overtreatment of Lyme disease: an observational study. *Ann Intern Med*. 1998;128:354-362.
32. Bergman AB, Stamm SJ. The morbidity of cardiac nondisease in schoolchildren. *N Engl J Med*. 1967;276:1008-1013.
33. Harach HR, Franssila KO, Wasenius VM. Occult papillary carcinoma of the thyroid: a "normal" finding in Finland: a systematic autopsy study. *Cancer*. 1985;56:531-538.
34. Nielsen M, Thomsen JL, Primdahl S, Dyreborg U, Andersen JA. Breast cancer and atypia among young and middle-aged women: a study of 110 medicolegal autopsies. *Br J Cancer*. 1987;56:814-819.
35. Montie JE, Wood DP Jr, Pontes E, Boyett JM, Levin HS. Adenocarcinoma of the prostate in cytoprostatectomy specimens removed for bladder cancer. *Cancer*. 1989;63:381-385.
36. Beam CA, Layde PM, Sullivan DC. Variability in the interpretation of screening mammograms by US radiologists. *Arch Intern Med*. 1996;156:209-213.
37. Elmore JG, Wells CK, Lee CH, Howard DH, Feinstein AR. Variability in radiologists' interpretations of mammograms. *N Engl J Med*. 1994;331:1493-1499.
38. Rosai J. Borderline epithelial lesions of the breast. *Am J Surg Pathol*. 1991;15:209-221.
39. Schnitt SJ, Connolly JL, Tavassoli FA, et al. Interobserver reproducibility in the diagnosis of ductal proliferative breast lesions using standardized criteria. *Am J Surg Pathol*. 1992;16:1133-1143.
40. Farmer ER, Gonin R, Hanna MP. Discordance in the histopathologic diagnosis of melanoma and melanocytic nevi between expert pathologists. *Hum Pathol*. 1996;27:528-531.
41. Pincus HA, Tanielian TL, Marcus SC, et al. Prescribing trends in psychotropic medications. *JAMA*. 1998;279:526-531.
42. Fisher ES, Wennberg JE, Stukel TA, Sharp S. Hospital readmission rates for cohorts of Medicare beneficiaries in Boston and New Haven. *N Engl J Med*. 1994;331:989-995.
43. Pritchard RS, Fisher ES, Teno JM, et al. Influence of patient preference and local health system characteristics on the place of death. *J Am Geriatr Soc*. 1998;46:1242-1250.
44. Strauss MJ, LoGerfo JP, Yeltatzie JA, Temkin N, Hudson LD. Rationing of intensive care unit services: an everyday occurrence. *JAMA*. 1986;255:1143-1146.
45. Tu JV, Naylor CD, Kumer D, DeBuono BA, McNeil BJ, Hannan EL. Coronary artery bypass graft surgery in Ontario and New York State: which rate is right. *Ann Intern Med*. 1997;126:13-19.
46. Yusuf S, Zucker D, Peduzzi P, et al. Effect of coronary artery bypass graft on survival: overview of 10-year results from randomized trials by the Coronary Artery Bypass Graft Surgery Trialists Collaboration. *Lancet*. 1994;344:563-570.
47. Collaborators NASCET. Beneficial effect of carotid endarterectomy for symptomatic patients with high grade stenosis. *N Engl J Med*. 1991;325:445-453.
48. Executive Committee for the Asymptomatic Atherosclerosis Study. Endarterectomy for asymptomatic carotid artery stenosis. *JAMA*. 1995;273:1421-1428.
49. Welch HG, Albertsen PC, Nease RF, Bubolz TA, Wasson JH. Estimating treatment benefits for the elderly: the effect of competing risks. *Ann Intern Med*. 1996;124:577-584.
50. Nease RF Jr, Kneeland T, O'Connor GT, et al, for the Ischemic Heart Disease Patient Outcomes Research Team. Variation in patient utilities for outcomes of the management of chronic stable angina: implications for clinical practice guidelines. *JAMA*. 1995;273:1185-1190.
51. Barry MJ, Fowler FJ Jr, Mulley AG Jr, Henderson JV Jr, Wennberg JE. Patient reactions to a program designed to facilitate patient participation in treatment decisions for benign prostatic hyperplasia. *Med Care*. 1995;33:771-782.
52. Wennberg DE, Lucas FL, Birkmeyer JD, Brendenbergh CE, Fisher ES. Variation in endarterectomy mortality in the Medicare population: trial hospitals, volume, and patient characteristics. *JAMA*. 1998;279:1278-1281.
53. Matchar DB, Oddone EZ, McCrory DC, et al. Influence of projected complication rates on estimated appropriate use rates for carotid endarterectomy. *Health Serv Res*. 1997;32:325-342.
54. Scherkenbach WW. *The Deming Route to Quality and Productivity*. Rockville, Md: Mercury Press; 1991.
55. Berwick DM. Controlling variation in health care: a consultation from Walter Shewhart. *Med Care*. 1991;29:1212-1225.
56. Connors AF, Speroff T, Dawson NV, et al. The effectiveness of right heart catheterization in the initial care of critically ill patients. *JAMA*. 1996;276:889-897.
57. Feinstein AR, Sosin DM, Wells CK. The Will Rogers phenomenon: stage migration and new diagnostic techniques as a source of misleading statistics for survival in cancer. *N Engl J Med*. 1985;312:1604-1608.
58. Redelmeier DA, Tan SH, Booth GL. The treatment of unrelated disorders in patients with chronic medical diseases. *N Engl J Med*. 1998;338:1516-1520.
59. Steinbrook R. Patients with multiple chronic conditions: how many medications are enough? *N Engl J Med*. 1998;338:1541-1542.
60. Kiefe CI, Funkhouser E, Fouad MN, May DS. Chronic disease as a barrier to breast and cervical cancer screening. *J Gen Intern Med*. 1998;13:357-365.
61. Krumholz HM, Radford MJ, Ellerbeck EF, et al. Aspirin in the treatment of acute myocardial infarction in elderly Medicare beneficiaries: patterns of use and outcomes. *Circulation*. 1995;92:2841-2847.
62. Soumerai SB, McLaughlin TJ, Spiegelman D, Hertzmark E, Thibault G, Goldman L. Adverse outcomes of underuse of β -blockers in elderly survivors of acute myocardial infarction. *JAMA*. 1997;277:115-121.
63. Leape LL. Error in medicine. *JAMA*. 1994;272:1851-1857.
64. Chrischilles EA, Segar ET, Wallace RB. Self reported adverse drug reactions and related resource use. *Ann Intern Med*. 1992;117:634-640.
65. Bates DW, Cullen DJ, Laird N, et al. Incidence of adverse drug events and potential adverse drug events: implications for prevention. *JAMA*. 1995;274:29-34.
66. Lazarou J, Pomeranz BH, Corey PN. Incidence of adverse drug reactions in hospitalized patients: a meta-analysis of prospective studies. *JAMA*. 1998;279:1200-1205.
67. Brennan TA, Leape LL, Laird NM, et al. Incidence of adverse events and negligence in hospitalized patients: results of the Harvard Medical Practice Study I. *N Engl J Med*. 1991;324:370-376.
68. Wilson RM, Runciman WB, Gibberd RW, Harrison BT, Newby L, Hamilton JD. The quality in Australian health care study. *Med J Aust*. 1995;163:458-471.
69. Eddy DM. The challenge. *JAMA*. 1990;263:287-290.
70. Albertsen PC, Hanley JA, Gleason DF, Barry MJ. Competing risk analysis of men aged 55 to 74 years at diagnosis managed conservatively for clinically localized prostate cancer. *JAMA*. 1998;280:975-980.
71. Sackett DL, Richardson WS, Rosenberg W, Haynes RB. *Evidence-Based Medicine: How to Practice and Teach EBM*. New York, NY: Churchill Livingstone; 1997.
72. Kuhn TS. *The Structure of Scientific Revolutions*. Chicago, Ill: University of Chicago Press; 1970.
73. Deyo RA, Psaty BM, Simon G, Wagner EH, Omenn GS. The messenger under attack: intimidation of researchers by special-interest groups. *N Engl J Med*. 1997;336:1176-1180.
74. Dickersin K. The existence of publication bias and risk factors for its occurrence. *JAMA*. 1990;263:1385-1389.
75. Kassirer JP, Angell M. Financial conflicts of interest in biomedical research. *N Engl J Med*. 1993;329:570-571.