Conceptualizing, Estimating, and Reforming Fraud, Waste, and Abuse in Healthcare Spending

Jerry L. Mashaw†
Theodore R. Marmor††

The elimination of "waste, fraud, and abuse" from American medicine is not a quick or easy solution to the challenge of rising medical care costs. Although there are clearly some savings available, particularly in the area of administration, they likely amount to far less than many policymakers hope. Many of the policy options available for reducing excesses face significant political hurdles or involve value judgments about non-quantifiable issues involving quality of care. Other alternatives seem as likely to bar necessary medical care as to eliminate abuses. In some instances, particularly those involving consumer fraud, a crackdown may merely shift costs without saving any money. Finally, a number of the suggestions for trimming waste and abuse involve ethical and moral judgments that Americans have yet to acknowledge and, in any event, may not wish to sign over to the government or any other third party.

I. The Overview ......................................................... 456
   A. Pork or Prosciutto? ............................................ 456
   B. The Dead and Dying ......................................... 457
   C. Managed Care: Eliminating Waste or Reducing Quality? 457
   D. Technology: Curse or Cure? ................................. 458
   E. Administration: Cutting Paperwork Will Save Money .... 458
   F. Fraud, Malpractice, and Cost-Sharing ..................... 459

II. Some Basic Concepts and Puzzles ................................. 460
   A. Distinguishing Fraud, from Waste, from Abuse .......... 463

† Sterling Professor of Law, Yale Law School.
†† Professor of Public Policy and Management, Yale School of Management. For invaluable research assistance and critical commentary, the authors wish to thank Guy Fish, Robert Gribbon, and John Pakutka.

Copyright © 1994 by the Yale Journal on Regulation
I. The Overview

A. Pork or Prosciutto?

In most economic transactions defining “waste” or “abuse” involves relatively simple cost-effectiveness or cost-benefit calculations. But in medical care this task is greatly complicated by the use of averages to make judgments about individual cases, and by the difficulties in estimating both costs and benefits.

In medical care, benefit calculations are fraught with both scientific and ethical uncertainty, in part because the “benefits” of medical care are not particularly well-defined. For example, one recent clinical study concluded that ultrasound examinations during pregnancy have no overall health benefit, but that judgment rests on the presumption that a patient’s peace of mind has no value.

Moreover, as demonstrated by the reaction to Oregon’s decision to establish a priority list for Medicaid procedures, assigning specific values to medical procedures raises ethical questions that American society has yet to resolve. Legislators often say of appropriations battles that “one man’s pork is another man’s prosciutto.” Similarly, one person’s wasteful medicine may be another’s miracle cure.
Fundamentally, the issue of medical waste requires a societal decision about whether we prefer to risk harmful undertreatment or wasteful overtreatment. It is easy to oppose waste. It is a far harder task to do so when the alternative is denial of necessary medical services.

B. The Dead and Dying

The ethical questions surrounding potential overuse of medical care services are perhaps most difficult in relation to those patients, often elderly, who appear near death. Numerous studies document the disproportionate amount of medical care expended on those in the final months of life. High intensity medical interventions for this group often do little more than prolong a low quality of life, while adding pain, suffering, and expense. By the same token, one would expect younger and healthier patients to consume fewer resources than the elderly and the critically ill.

More significantly, research suggests that we know less about probable outcomes than we might imagine. Studies show that the largest portion of medical care outlays goes to those who defy the averages, specifically patients not expected to die but who do, and those expected to die but who in fact survive. It is difficult to imagine a system that could neatly root out “waste” of this nature, and the question of who will decide what care we eliminate is not a frequent subject of political debate.

C. Managed Care: Eliminating Waste or Reducing Quality?

Many experts and the Clinton Administration presume that “managed care” is the best response to the overutilization of services and the excessive administrative costs that supposedly account for much of the waste and abuse in American medicine.

Some forms of managed care, such as HMOs, do indeed have the potential to reduce costs significantly. The HMO is both the provider of services and the insurer. It cannot collect additional fees if it exceeds its budget, and it therefore has an enormous incentive to keep care within its budget. Studies suggest that, in some instances, HMOs may reduce hospital use by 40% and total spending by 25%. However, the need to align services with a preset budget may also lead to queuing and creates incentives to bar access to specialists and expensive medical technology.

Alternative forms of managed care appear to have less impact on costs. Pre-certification and concurrent review of hospitalizations appear to generate savings of about 8% on in-patient hospital costs and 4-5% on overall medical expenditures. Other forms of managed care, such as individual practice
associations (IPAs) and preferred provider organizations (PPOs), seem to have little effect on costs.

What’s less clear is whether the cost savings are purchased at the expense of technical quality and patient satisfaction. Surveys show HMO patients are less satisfied with their care than non-HMO patients and are particularly unhappy about excessive waiting time and being shuttled from physician to physician because of medical staff turnover or changes in their own medical needs. The $200 billion question is whether HMOs and other managed care arrangements that constrain costs are eliminating waste or merely limiting access to needed services.

D. Technology: Curse or Cure?

The American fondness for wanting “the best money can buy” seems to apply to medicine, and many commentators believe it accounts for a significant share of the nation’s rising health costs. There is no doubt that American medical facilities possess more high-technology equipment than their counterparts in the rest of the world and that this equipment is used at a much higher rate. Once the technology has been acquired, it will be put to use. There is little doubt that acquisition and use of such equipment adds short-term costs. But it is less clear that overutilization of technology is motivated simply by the quest for profit. Rather, American culture promotes aggressive intervention and places a premium on professionals’ being at the cutting edge of their fields. The reliance on third-party payments to finance medical care strengthens patients’ own bias towards using whatever methods are available when their health is at stake. And fee-for-service payments to physicians further support technological overkill.

The United States needs to put in place incentives for cost-reducing technology to balance the high rewards for innovation that increases costs. But the debate over medical technology, like so much else in the medical care puzzle, also turns in part on value judgments about costs and benefits. Some argue strongly that high technology medicine is high quality medicine, not waste. If technology in fact increases the quality of care, it becomes far harder to cut waste by discouraging the use of cutting-edge equipment and medication.

E. Administration: Cutting Paperwork Will Save Money

There are obviously big savings available from cutting paperwork in insurance claims, billing, underwriting, processing, and marketing. Government programs such as Medicare and Medicaid spend just $29 per person on administration, compared to $150 per person in the private system, and per
Healthcare Fraud

capita costs in the United States are six times higher than they are under Canada’s single-payer national health insurance system.

Administrative savings will require significant changes in the way America’s private insurance system operates, with the greatest savings achieved by a system that eliminates private insurance and replaces it with a wholly public program. The managed competition approach favored by the Clinton administration could also achieve savings by bringing uniformity to claims processing and billing. It would not save as much as a single-payer system, however, because insurance companies would continue to expend large sums on marketing.

F. Fraud, Malpractice, and Cost-Sharing

Clearly, the elimination of outright fraud and abuses such as “unbundling” and “upcoding” would generate genuine savings. But it is unclear how much money is at stake or precisely how to eliminate such practices. There is no empirical evidence to support the popular notion that fraud and abuse add about 10% to U.S. medical costs. Moreover, the degree of monitoring necessary to eliminate all fraud would affect privacy, professional autonomy, and other rights Americans hold dear.

It is also not at all certain that the practice of defensive medicine as a means to limit legal liability for malpractice significantly adds to medical costs. While the American Medical Association believes fear of liability may add $12-13 billion in costs each year, there is at least as much reason to believe fear of litigation holds down medical costs by discouraging risky, and expensive, medical intervention. Further, there is substantial evidence that the current system massively undercompensates most victims of medical negligence. A reform that limits the exposure of medical providers, while also addressing the undercompensation issue may produce little, if any, net savings.

Finally, some analysts argue that medical costs could be reduced significantly by stripping away some of the insulation of third-party payments and requiring individuals to share directly the expense of medical treatment. Cost-sharing, through either higher deductibles or co-payments, does appear to reduce use of medical services. However, utilization of service appears more closely tied to economic status than to medical needs. Cost-sharing is just as likely to discourage appropriate medical treatment as inappropriate treatment. Cost-sharing saves dollars, but it also eliminates beneficial use of medical resources—an exchange that may represent an undesirable public policy outcome.
II. Some Basic Concepts and Puzzles

Many seem to believe that the elimination of waste, fraud, and abuse in American medicine would, virtually by itself, finance universal health insurance. Excessive administrative costs, unnecessary procedures, inadequate consumer vigilance, unavailability of preventive care, and skyrocketing malpractice premia all suggest an extraordinarily wasteful allocation of the enormous resources American society devotes to medical care. Charts 1 and 2 on the following pages show graphically the seemingly inexorable rise in U.S. health expenditures as a percentage of the gross domestic product, and the striking disparity between the U.S. health expenditures and those of other OECD countries.

Why is medical care in America so inefficient? Why do we not read similar stories about waste, fraud, and abuse in the consumer electronics industry, or the market for processed food, or the college education system? In most markets, we expect providers who are fraudulent, abusive, or wasteful to be driven mercilessly from business by service-conscious and cost-conscious consumers, or put behind bars. Consumers who waste resources forgo the purchase of other desired goods, and thus have every incentive to smarten up.

Why is medical care different? Commentators have suggested three ways in which medical care differs from other services and have noted some probable consequences of these differences. First, the uncertainty of the need for medical care, coupled with the devastatingly high cost of treatment, creates a vigorous demand for health insurance. Less than one quarter of healthcare payments are “out of pocket.” Second, the community’s interest, both in the prevention of the spread of contagious diseases and in compassion for those unable to pay for care, leads to substantial public subsidies for medical care and health-related research and development. In other words, we want sick or injured patients to receive care regardless of ability to pay, and we want research to constantly expand the frontiers of medical science. Finally, the uncertain and asymmetric nature of medical information leads to licensure of health professionals and reliance on these fiduciary providers to guide consumer choice. We really don’t know what to buy, and must trust the seller to tell us.

These differences between medical and other markets give rise to many of the areas in which waste, fraud, and abuse are thought to occur: restrictive licensure, supplier-induced demand, overuse of specialists and high-tech medicine, excessive administrative costs, skyrocketing malpractice premia and defensive medicine, herculean efforts to prolong life, and outright fraud. Unfortunately, there are few neat ways of surgically striking out (or measuring, for that matter) waste, fraud, and abuse in medical care. Chemotherapy, or perhaps carpet bombing, is a better analogy: wiping out the bad means risking

Sources: OECD, 1991 and Congressional Budget Office

*(1995 and 2000 data projected)
Total Health Expenditures as a Percentage of Gross Domestic Product (1990)

Source: OECD, 1991
injury to the good. Solutions such as increased patient cost-sharing, managed care, supply controls, and public rationing will all run this risk.

Nevertheless, reform is important. Wherever possible, waste, fraud, and abuse should be removed. But not throwing the proverbial baby out with the bathwater requires attending to two tasks. One is achieving a more precise understanding of just how fraud, abuse, and waste manifest themselves in American medicine. The other is clearly comprehending the myriad contexts in which these terms are employed. Indeed, clarifying the importance of fraud, waste, and abuse requires not only that we give determinate meaning to those terms, but also that we ask questions such as “by whom” and “in relation to what.” Finally, if an analysis of fraud, waste, and abuse is to clarify policy debates about the reform of American medical care and its financing, analysts must appreciate the uncertainties of medical practice and the scarcity of important evaluative data.

Accordingly, our analysis begins with a simple characterization of fraud, waste, and abuse. It then illustrates (1) the differing implications these terms have for cost reduction depending upon whose conduct is said to be fraudulent, wasteful, or abusive and (2) the varying policy contexts within which fraud, waste, and abuse can be identified and perhaps eliminated. The discussion then provides and comments on existing estimates of unnecessary costs resulting from these three causes. As will become clear, any serious analysis must highlight the uncertainties that surround the measurement of fraud, waste, and abuse in American medical care.

A. Distinguishing Fraud, from Waste, from Abuse

Fraud is in many ways the simplest concept to define. Classic fraud involves the misrepresentation of relevant facts combined with the detrimental reliance on that misrepresentation by another party to the transaction. Many applications of the fraud concept are straightforward. Medical care professionals, for example, commit a fraud when they misrepresent the nature or the scope of the services they provide to increase the level of compensation they receive—whether from an individual or a third party payor (such as a public or private insurer). Some misrepresentations fall more clearly within the scope of “fraud” than others. Billing for procedures never performed or patients never treated is clearly fraudulent, as is hiding a mother’s assets to qualify her for Medicaid reimbursement of nursing home costs.

In contrast, “abuse” is often used to describe instances of profligate spending on medical care that is at the margin of fraudulent practice, in which the conduct in question is permitted by the existing arrangements for medical care provision and financing. The borderline between “fraud” and “abuse” is as a consequence inherently hazy. Some instances of abuse are probably frauds
in euphemistic dress. Consider, for example, the psychiatrist who bills Medicaid for a “consultation” because on her way to another appointment, she spent a minute speaking with an institutionalized patient in the corridor, and inquired about the patient’s condition. The psychiatrist will have observed the patient’s manner (momentarily) and she may even later have dictated a note for the patient’s record. But this casual encounter is hardly what a sensible person would consider a psychiatric consultation. A jury might or might not convict this doctor of fraud, but the conduct is surely abusive.

Indeed, the possibilities for abusive behavior by providers are virtually limitless. Insisting on multiple visits where one is sufficient is an oft-told tale of medical abuse. Unbundling procedures into separate activities that are billed independently at a higher total cost is also thought to be common. Pursuing every diagnostic procedure conceivably relevant to a worker’s compensation case, as an accommodation to patients (or their attorneys) seeking to bolster recoveries, is another form of overutilization of American medical care that many believe to be abusive. Common to all of these cases is some plausible medical explanation for the procedure. Abuse lies in the motivation imputed to the provider, in the suspicion that costs were increased for purposes related, at best, only tangentially to producing a better medical result. But, strictly speaking, these are not instances of fraud. The tests or procedures or examinations were done, and the appropriate billing codes were used to seek reimbursement.

Consumer abuse shows a similar pattern. Patients who purchase prescription analgesics because they are covered by a pharmaceutical insurance plan inflate medical care costs unnecessarily and abusively when they know that a cheaper non-prescription analgesic would be as effective. Compliant physicians are accomplices to this abuse. Patients who transfer assets or who limit their hours of work in order to maintain Medicaid benefits may similarly be regarded as abusing the system. Note, however, that such “eligibility” abuses may or may not significantly affect aggregate health expenditures. They may be “cost-shifting” rather than “cost-enhancing” abuses, a topic to which we return below.

Abuse merges almost imperceptibly into waste. The surgeon who requires both a six-week and a six-month postoperative examination may be abusing the system. If he does not believe that a six-month examination is necessary, even if the patient or the insurer is willing to pay for it, the surgeon is clearly abusing the system. On the other hand, the physician may simply have different beliefs than some of his colleagues about the efficacy of, or the need for, a follow-up examination. One can properly describe his conduct as wasteful only if by some objective standard the costs of this extra follow-up visit do not equal the benefits that flow from it. An objective basis for identifying waste might be a carefully controlled test comparing medical outcomes with and without

464
six-month postoperative examinations. Or, it might follow from the dominant beliefs of the medical community as evidenced by the behavior of the majority of surgeons.

The criteria for identifying waste, whatever the methodology, are of at least two types: "cost-effectiveness" and "cost-benefit" tests. Cost-ineffective medical interventions are simply ones that produce, on average, no better medical results than some cheaper treatment. A procedure that fails the cost-effectiveness test will also be comparatively less cost-beneficial than alternative procedures—it's ratio of costs to benefits will be lower. But interventions may also fail a cost-benefit test if, whatever their superiority to other modes of treatment, they nevertheless fail to produce benefits that on average exceed their costs.

A moment's reflection on these standard definitions, however, reveals why methods of identifying waste are likely to be highly controversial. The urologist who resects a hypertrophic prostate, without first attempting to relieve the symptoms through available medications, is probably wasteful by both cost-effectiveness and cost-benefit standards. The same symptomatic relief might have been achieved at significantly less cost through pharmaceutical rather than surgical intervention. If similar results are available at lesser cost, the cost-benefit ratio of prescribing pharmaceuticals is lower than that of engaging in surgery.

The controversial nature of these judgments arises from two sources. One is the problem of using averages to make judgments about individual cases; the other is the difficulty of estimating both the costs and the benefits of treatment. With respect to the prostatectomy example, it might well be cheaper in the long run to do a surgical resection for a fifty-year-old patient than to keep him on Hytrin for thirty years—and perhaps have to do a resection anyway if the drug loses its efficacy over time. The resection might also remove precancerous tissue that would have eventually made surgery necessary. But cost-effectiveness, on average, tells us little about waste in particular cases. Where patient circumstances are widely diverse rather than closely similar, an apparently straightforward cost-effectiveness analysis based on averages may be quite misleading.

Cost-benefit analyses of waste are even more treacherous. Even where costs are relatively straightforward, benefits calculations are often fraught with both scientific and ethical uncertainty. In part, this is because the "benefits" of medical care are not well-defined. Effects on mortality or morbidity rates may be relatively clear, but these gross indicators often have little discriminatory power. Everything from complete inaction to major surgery may produce similar statistical correlates. Finer-grained evaluative criteria are needed, but appraising the effects of medical interventions on physical and mental comfort, or on functional capacity, is both laborious and uncertain. The current
controversy over ultrasound examinations during pregnancy is a good example of this point. A recently reported clinical trial found no medical benefits. Eminent obstetricians immediately denounced the report as clinically useless, however scientific its methodology, because it failed to place any value on the patients' peace of mind during pregnancy.

Moreover, the assignment of dollar values to even the more clear-cut morbidity/mortality estimates raises serious ethical questions. Disputes over the recent Oregon priorities list for Medicaid procedures reveal deep divisions about the value to be placed on prolonging life and relieving symptomatic discomforts or functional limitations. As legislators are wont to say about criticisms of the appropriations process, "one man's pork is another man's prosciutto." With respect to medical procedure waste, one person's squandering of medical resources may be another's miracle cure.

This is not to conclude that we lack clear instances of waste. For example, the administrative expense of billing and accounting for individual payors is wasteful when clearinghouse procedures, simplified medical reimbursement forms, or single-payor systems would eliminate a substantial portion of these expenses. And, it is difficult to believe that the extraordinary rates of diagnostic testing for worker's compensation claims—at nine times the average rate for injuries and illnesses of persons outside worker's compensation—represent some special medical necessity in the worker's compensation system. The same might be said of the dramatically higher utilization rates of physician-owned diagnostic testing equipment in comparison to the rates where there is no self-referral. For policy analysis, however, the important point to understand is that "waste" comes in even more guises than "abuse." Therefore, it is important to focus on the form of waste and how it is measured. Where the conception of waste is problematic, public policy confronts questions about what risks it wishes to run—whether it prefers to run risks of harmful undertreatment or of wasteful overtreatment—more than simple judgments about whether and where to eliminate waste. Constructing incentives for providers or patients to use the exact amount of medical treatment that is cost-beneficial or cost-effective is an unrealistic ambition. Being "against waste" is easy. Deciding how to eliminate it when the alternative risk is underprovision of medical services is much more challenging.

B. By Whom and about What

Analyzing the impact of fraud, waste, and abuse in medical care expenditures must move beyond sorting out differences among these concepts and clarifying the difficulties of their application. The questions of who makes use decisions, and about what, are also highly relevant to whether real savings
Healthcare Fraud

are available from a concerted attempt to reduce fraud, waste, and abuse from the medical system. The following examples illustrate this point.

Provider fraud almost always generates unnecessary medical expenditures. In the obvious case, payment was made for something that wasn’t done. By contrast, consumer fraud often concerns insurance eligibility rules. The question then is who pays, not whether care was received and payment made (or left uncompensated notwithstanding its real resource cost). Consumer fraud may also generate medical services that ought to be eliminated. Where consumers would forgo treatment if they had to pay for it—not because of lack of funds, but because of the uncertainty of perceived benefits in relation to cost—a fraud that shifts the treatment cost to a third party produces unnecessary expenditure. But frequently the only question is whether the cost shows up in a state or federal budget, the budget of a private insurance company or a self-insured employer, or the uncompensated-care/bad-debt ledger of a medical provider. Hence, while the total cost of provider fraud in the system could probably be eliminated, any figure representing consumer fraud on Medicaid, Medicare, or private insurers vastly overstates the amount that fraud is adding to medical costs.

There is a similar distinction between provider and consumer abuse. However, it is trickier to measure the costs of provider abuse than provider fraud. If doctors in hospitals were prevented from disaggregating procedures to maintain or increase their incomes, they might nevertheless have the political and economic power to recoup some or all of their losses by renegotiating the fees for the aggregates. If so, figures representing total billing abuse significantly overstate the real savings available from its elimination.

The overestimation of savings through the elimination of abuse are nowhere better illustrated than in the widespread belief that there are huge savings to be gained from reforming our malpractice liability system. Contrary to popular belief, the so-called “defensive medicine problem” is extremely complex. The medical profession itself regularly blames overutilization of diagnostic procedures and overcautious scheduling of repetitive medical encounters on the fear of malpractice charges. But it is extremely difficult to gauge the separate effects of this claimed anxiety from the effects of the general culture of medical training and practice in the United States—a culture that promotes aggressive intervention and extensive use of expensive technology. Not only is the “malpractice-made-me-do-it” claim self-serving, but it may also be objectively irrational. There is very little evidence that physician defensiveness effectively protects against lawsuits.

Finally, any malpractice reform should take account of the substantial evidence that the current malpractice system massively undercompensates the victims of medical negligence. It is hard to imagine a reform that eliminates “pot of gold” malpractice litigation but does not simultaneously establish a
worker's compensation-like system for compensating—at lower levels—the many victims of medical negligence who now receive nothing through the tort system. In short, there is little reason to believe that malpractice reform will cause the total costs of compensation to go down or that increased compensation costs will be more than offset by a decline in defensive medicine.

The foregoing discussion is not an argument against medical malpractice reform. It is hard to say a good word for a “compensation” system that provides neither effective deterrence nor adequate compensation—and that may be driving up total medical costs as well as distorting career and treatment decisions by physicians. Nevertheless, the belief that medical malpractice reform will sharply reduce health expenditures could be a pipe dream.

A similar analysis can be made of the claim that the American medical system now provides a wasteful model of “emergency room medicine” for uninsured or underinsured persons who use the emergency room as a substitute for primary care. There is surely merit to the idea that universally-available primary and preventive care would eliminate many needless emergency room visits at considerable savings. Nevertheless, eliminating this “waste” through universal insurance for primary and preventive care may cause aggregate expenditures to rise because of large increases in the number of primary care encounters. Even so, such a change may be desirable. One cannot conclude, however, that the elimination of this “waste” will reduce the necessity for increased levels of financing for medical care. There are clear social gains from producing a healthier and happier population through a more cost-effective medical system. But the effects on aggregate medical care spending from eliminating this category of “waste” may be the opposite of the effects of efforts directed at eliminating paperwork or administrative and accounting waste.

C. Data, Data, Who's Got the Data?

A sensible analysis of the potential savings from reducing fraud, waste, and abuse in American medicine must make subtle distinctions and ask philosophical questions such as whether the reduced anxiety of those who pay medically unnecessary visits to the doctor should be counted as a benefit. But our analysis must also confront the absence of systematic data on, and the prevalence of misconceptions about, many of the features of fraud, waste, and abuse that are relevant to policy analysis. The prior discussion of malpractice litigation and defensive medicine is an obvious case in point, but many other commonly-accepted claims are equally problematic.

1. For a broader discussion and examination of some empirical studies on this matter, see LOUISE B. RUSSELL, IS PREVENTION BETTER THAN CURE? (1986).
Healthcare Fraud

For example, we actually know quite little about the benefits of HMO practice versus fee-for-service medicine. Only one existing study of the differences between HMO and fee-for-service costs could be described as scientifically rigorous. That study is now over a decade old and makes comparisons between systems of provision that may no longer exist in the same forms. Other data on this question are highly impressionistic, to be generous. Hence, the estimates that are currently bandied about concerning the savings that will result from forcing people into HMOs are unrealistic. HMO savings—usually attributed to a gatekeeper who economizes care or to a capitation method of payments—have been realized with respect to self-selecting or special populations. We have no idea whether the same results can be achieved when the policy is applied to the whole population. Although the argument for some savings is plausible, we lack the empirical basis to conclude that the incentives to economize that seem to be built into HMO practice result in behaviors that dramatically reduce medical care costs when applied to the population at large.

A similar controversy surrounds the managed care revolution of the past decade. There is evidence that managed care eliminates unnecessary treatment. There is also evidence that managed care is sufficiently costly to providers, insurers, and patients and that the savings are all but swallowed up by increased administrative expense—some of which is not counted because it shows up in patient time and consternation rather than on the budgets of providers and insurers. We also know very little about whether managed care is saving rather than shifting costs and what its effects are on the distribution of care in the population. If managed care shifts the availability of care in the direction of those most capable of “working the system,” it may exacerbate the inequities of current medical care provision without having any significant impact on aggregate costs.

On closer examination, there are serious questions about how much fraud one wants to eliminate from contemporary medical care. The degree of monitoring that would eliminate all fraud would affect privacy, professional autonomy, and other values in ways that would surely make it unattractive. The IRS, for example, could recapture eleven to fifteen dollars for every additional dollar spent more aggressively auditing American tax returns. Congress does not appropriate funds for this intensified auditing because of political considerations. Similar considerations limit the degree to which any class of health reforms could eliminate fraud.

Without further discussion of informational uncertainties and philosophical ambiguities, some basic lessons emerge from our analysis of fraud, waste, and abuse in American medicine. First, we must take great caution when using the

2. "Wildly unscientific" would be a more pejorative description of the data often cited.
relevant terms: they are both emotionally charged and highly ambiguous. Second, the relationship between the triumvirate of fraud, waste, and abuse and aggregate spending for American medical care is complicated. The extent of “savings” will depend in large part on whether the analysis concentrates on the budget of some particular payor or instead on national or jurisdictional aggregates. Moreover, prediction of savings from the elimination of fraud, waste, and abuse must take into account the likely increases or decreases in the cost of solving various fraud, waste, and abuse problems. Third, even after careful conceptual analysis, data limitations may not permit reliable estimates of the savings to be had from particular reforms. The degree of reliability will vary depending on the reform. Not only will deriving precise estimates be impractical, but the potential range will also vary by orders of magnitude of the double or triple digit variety. Finally, because these complexities are not part of the current general public debate, discussing them in greater detail would contribute significantly to a more realistic appraisal of reform proposals in American medicine. It is to that task that we now turn.

III. What Do We Know?

Fraud, waste, and abuse in American medical care encompass a wide range of practices. Rather than seeking further taxonomic clarity, the discussion that follows analyzes some prime candidates for cost savings suggested by the fraud, waste, and abuse literature. It provides estimates of savings where available, and comments on the persuasiveness of the current data. As should become increasingly clear, existing limits on our understanding make us all, at best, sophisticated consumers of often conflicting rhetoric. Nevertheless, we do know that some things are clearly wrong with current arrangements, we can estimate their costs within a reasonable range, and we have persuasive evidence relevant to fixing them. These simple truths should not be lost in the chaos of claims and counterclaims about particular aspects of medical fraud, waste, and abuse.

A. Over- and Undersupply of Relevant Professionals

One of the most confusing aspects of the medical waste issue is the debate about the relationship between the supply of medical professionals (or medical technologies) and the use of medical services. It is an article of faith in the economics profession, particularly the field of microeconomics, that licensure provisions decrease the supply of available physicians and increase their cost. In the realm of theory, this effect is not necessarily a “waste,” but rather a misallocation of medical resources. Total expenditures for medical care may not rise, but medical care will be undersupplied and the prices for each episode of medical care provision will be too high. In a world where substantial
populations are underserved by the medical care community and where access to care is barred by cost, public policies that restrict supply and increase costs are unattractive.

The theoretical concerns of microeconomists are echoed by many medical providers who do not have favored professional status and are therefore wholly or partially excluded from the medical care market: nurse-practitioners, physician's assistants, social workers, and so on. Nurse-midwives and chiropractors may be the most vocal of these groups. So what do we know about the way in which licensure affects the price and quantity of medical care services delivered?

The empirical studies of this phenomenon are so varied and contradictory that they are difficult to characterize. The best summary we can make goes something like this: interstate differences in the stringency of licensing requirements tend to inhibit professional mobility and drive up fees in the most restrictive states. The magnitudes of these effects, however, differ substantially across professions. In the field of optometry, the most restrictive licensure provisions may increase the cost of access to eyeglasses by as much as 33%. By contrast, restrictions on advertising for prescription drugs raise prices by approximately 5%, and restrictions on the use of other professional personnel may raise prices by even less.

The magnitude of these cost increases is subject to question, but the direction of the effect is constant across the empirical studies: restrictive practices contribute to higher fees and higher practitioner incomes. There is very little evidence that licensure provisions have any significant effect on the quality of medical services. Similarly, professional review organizations seem to have little effect on the competence of practitioners. Hence, the whole structure of licensing may impose costs and no benefit. The only activities that seem to have a clear impact on physician competence are the output review processes that are used by hospitals (which have an incentive to engage in such monitoring because of their own possible liability for negligence).

The problem is that hospital requirements tend to be higher than the standard requirements for licensure. Hence, to the extent that hospital monitoring is effective in weeding out incompetent practitioners, it also has the effect of further restricting supply and increasing costs. Current credentialing

4. Id. at 393 (citing JOHN F. CADY, DRUGS ON THE MARKET (1975)).
5. Id. at 387 (citing FRANK A. SLOAN & BRUCE STEINWALD, HOSPITAL LABOR MARKETS, (1980)) (licensure requirements raise wages of medical personnel by 2% to 13%).
6. For a summary of research on the effects of licensure on quality, see id. at 395.
7. For a review of empirical studies on the effects of credentialing practices on the quality of care, see id. at 397-400.
8. Id. at 402-03.
practices by state licensing boards seem to be increasing costs while having very little effect on the quality of patient care. Yet the possible directions for reform are not clear. To eliminate this "waste" by imposing the higher standards of competence that are customary for admission to practice in hospitals might well increase both the quality of care and aggregate medical costs. The benefits might not exceed the costs. On the other hand, deregulation of medical practice, would be difficult politically, not only because of interest group pressures, but also because of patients' general beliefs in the efficacy of credentialing.

In short, we may be wasting money on licensing systems, but we still would like additional credentialing or monitoring activities, which would increase overall costs. The need for better quality assurance is not simply a reflection of irrational patient fears. Studies dating back to the early 1960s reveal shockingly high incidences of preventable deaths or invalid diagnoses in random samples of hospital-admitted patients. A recent and widely noted Harvard Medical School study, conducted in New York hospitals, estimated the percentage of avoidable adverse effects on patients at 7%. That figure is consistent with studies in the mid-1970s of physician-inflicted injuries in hospital settings. Studies of hospital and out-patient care that included invalid diagnoses as well as preventable injury and death among adverse effects have found much worse performance by physicians. Estimates of inadequate care in these studies range from 29% to 65%.

Once again, credentialing may not be the remedy for this medical practice problem. In tests comparing the diagnostic capacity of technicians and pathologists, for example, medical technicians failed to detect evidence of carcinoma in pap smears 30% of the time, whereas 37% of the time pathologists failed. Credentialing increases costs, but may be worth those costs if done in ways that actually affect practice. Current licensure provisions may be extremely wasteful, but the proper direction of reform may be towards more stringent monitoring that would further increase costs, with a simultaneous increase in the quality of care.

Strangely, medical economics provides an equally plausible and exactly contrary theory concerning the relationship between supply restrictions and overall costs. That theory holds that the cost increases in medical care are in substantial part the result of an oversupply of medical care personnel. The claim is that medical care violates the basic principle of microeconomics that supply does not create its own demand. The explanation for this is that much decisionmaking about the use of medical care is made not by patients but by

---
9. Id. at 395.
11. See Gaumer, supra note 3, at 395.
12. L. Lamotte, Validation of Performance Measures, Address to the American Board of Medical Specialties Conference on Extending the Validity of Certification (March 24, 1976).
Healthcare Fraud

providers. Moreover, patients have little incentive to investigate the physician's recommendations when third-party insurers pay for treatment.

The controversy over this phenomenon of supplier-induced demand is extremely heated, but the light in the empirical literature is rather faint. Pioneering studies in the late 1970s suggest that there is a 0.3% increase in surgery rates with each 1% increase in the supply of surgeons. Refined methods and better data suggest that a 1% increase in surgeons increases surgery rates only .08% (one-third of the rate previously found), but that the effects are differentiated. Elective surgery rates increase .12%, whereas non-elective surgery does not change in response to increases in the supply of surgeons. Prices, on the other hand, increase 0.9% for each 1% increase in surgeons. This vivid contradiction of neoclassical economic theory is perhaps explained by the fact that surgeons compete for patients largely on the basis of perceived quality, including amenities, rather than price.

Critiques and rebuttal studies suggest that there is in fact no statistically significant increase in demand for services with increases in physician supply, but these studies have been largely debunked. The consensus is now that increasing supply does increase demand, but effects are not large and must be disaggregated by the type of service being supplied.

A recent study of trends in health expenditures and the availability of physicians confirms the view that physician-induced demand is not a major cause of the rise in healthcare costs in America. Growth in real medical care dollars expended per capita as well as the growth of medical expenditures as a percentage of GNP and of GNP per capita in the aggregate has been steady since 1930. On the other hand, the number of physicians per person has increased at a much lower rate, and during some decades has decreased. Hence, while the total number of medical professionals may correlate with the overall level of expenditure, the correlation between the growth of the number of medical professionals (defined in virtually all studies as physicians) and the growth in health expenditures is weak.

A more plausible theory about waste in the system related to the supply of medical professionals holds that we do not have too many physicians, or more generally, too many medical professionals. Rather, we have too many highly trained or highly specialized ones. This trend toward specialization skews

14. Id. at 302-05.
our medical system toward invasive and expensive interventions and away from primary and preventive care and care-giving at a low-tech level for those with chronic illnesses and impairments.

Comparison of the rates of specialization in the United States with those of other developed countries certainly suggests that we are very top-heavy. Twenty-five to 50% of physicians in most developed nations claim some board-certified specialty, whereas in the United States the number is 70%, and with recent graduates, well above 80%. Critics of this imbalance posit that we necessarily stint on primary and preventive care that would both improve health and reduce costs by preventing additional disease or illness. Conversely, the superabundance of specialists leads to the diagnosis of “within specialty” illnesses and the use, at high cost, of the specialty tools that are increasingly available. If many of these highly specialized and costly interventions are unnecessary or of limited utility, our system is creating considerable waste in the form of misallocation of professional resources.

There appear to be no reliable estimates of the amount of “waste” this skewing toward specialization produces in the United States. It is clear, however, that the solution is not a simple one. The economic rewards to specialization must be limited and the culture of medical education and medical practice changed. Both remedies are politically difficult, and the second involves a multi-generational effort.

The difficulties posed by the specialization issue illustrate the dispute over the use of non-physicians as primary caregivers. To be sure, it makes sense that routine care can be provided by persons not enormously overtrained for that activity. On the other hand, if primary care is the gateway into the medical system, it is the locus of critical decisionmaking about the future paths down which a patient travels. In short, much primary care involves diagnosis, the most difficult part of medical practice. For that task one presumably wants the most highly trained and experienced professionals available to the system. Any reorganization toward de-professionalization or reduction in specialists must proceed cautiously with these two quite different aspects of primary care firmly in view. Because that is true, estimating the savings from this sort of waste reduction is virtually impossible.

B. Oversupply and Overutilization of High-Technology Medicine

Closely connected to specialization is the vexing issue of the use and development of medical technology. Many commentators attribute a significant portion of the rise in American medical care costs to the development and

Healthcare Fraud

utilization of high-technology medical devices. There are four principal causes of overutilization of high technology medicine. First, high technology diagnosis and curative intervention tend to be associated with specialized practice. A country with a high proportion of specialists will have a high proportion of doctors accustomed to using high technology medicine. Second, blatant self-interest may be involved. The specialist who recommends a particular diagnosis or treatment may own the relevant equipment or be a participant in its ownership. Third, competition among doctors, and particularly among hospitals, may turn on the use of high technology equipment. Hence, hospitals and doctors will want to acquire such equipment and, once acquired, they will want to use it in order to amortize its cost. The use of high technology is in this sense simply an aspect of competition among healthcare providers. A fourth issue is the practice of defensive medicine, a response to fears of malpractice liability. That issue will be discussed later because it relates to overtreatment generally, not just to overuse of high technology medicine.

The first cause of overutilization of high technology medicine, specialization, need not detain us. Presumably it will be cured by redistribution of medical professionals between generalists and specialists.

There is little doubt that self-referring physicians use more of certain sorts of facilities than do independent-referring physicians. A recent study of the California Workers' Compensation system found that self-referral increased the cost of medical care for workers in the system for each of the three types of services studied: physical therapy, psychometric tests, and MRI scans. The costs increased, however, for different reasons. There were higher total costs for physical therapy because more cases were referred, but costs per case were lower. By contrast, more psychometric tests were done, and the costs per test were higher. Finally, an independent evaluation of the appropriateness of the scans revealed that self-referring physicians were 38% more likely than independent-referring physicians to be judged to have referred a patient inappropriately for a scan. This study, and similar ones, indicate that complete elimination of self-referral would reduce the overall cost of medical care. The crucial question, of course, is how much of the cost due to self-referral is "waste." For example, many believe that there is too little emphasis in American medicine system on rehabilitative medicine and on the treatment of mental illness. This could mean that the higher number of referrals for physical therapy in the California study and the higher number of psychometric tests were entirely appropriate. There was, after all, no control on the appropriateness of these referrals. The self-referring physicians may have been getting it right,


while the independent-referring physicians were getting it wrong. The MRI results are less ambiguous. Assuming that the control was correct, there was at least 38% more waste in the referral for MRIs by self-referring physicians.

The MRI is a good example of oversupply, the third principal culprit behind the overutilization of high technology medicine in the United States. There are more MRI machines in California than in the whole of Canada. Indeed, after adjustment for population characteristics, Californians undergo 51% more MRI procedures than the US national average.\textsuperscript{21} This figure is generally believed to be the result of supply-induced demand in California. The MRI machines have been put in place, and must be amortized. Hence their owners, whether self-referring physicians or independent hospitals and clinics, will tend to use them or to create incentives for their use. Indeed, in the California workers' compensation study, fully one-third of the MRI scans requested, both by self- and independent-referring physicians, were judged inappropriate. At $1000 per scan, this overuse adds significant costs to the medical care system.

It is indeed striking, when looking at foreign medical systems, to realize that virtually every other advanced country uses highly developed controls on investment in capital equipment, including high-technology equipment like MRI scanners. Most students of the success other countries experienced in constraining cost escalation ascribe a significant portion of that success to controls over capital investment, including investment in new technology.

Once again, however, there are benefits as well as costs to allowing some overcapacity. Overcapacity reduces queues, and queuing is not costless. Moreover, the knowledge that one can deploy developed equipment feeds technological innovation. Nevertheless, it seems clear that the United States has failed to construct incentives that sufficiently reward cost-reducing technological innovation in an attempt to counter the strong incentives to develop procedures or devices that increase both the level of technological sophistication and the cost of medical care.

Estimates of the excessive costs of the oversupply and overutilization of high technology medicine range from the enormous to the insignificant. Indeed, some commentators seem to ascribe most of the excess costs in the American medical system to the overuse and oversupply of high technology medicine. Studies that take this view reach their conclusion by ascribing virtually any otherwise unexplained rise in costs to new medical technology. In short, new technology becomes a residual that picks up all of the supposed excess expenses that cannot be allocated elsewhere.

This position is not crazy, just not well-justified. International comparisons provide the principal evidence for technology driven cost increases. For

\textsuperscript{21} Id. at 1506.
example, a recent study by Robert Evans finds that, per million people, the former West Germany has 0.7 open-heart surgical units, Canada has 1.2, and the United States has 3.3. Likewise, there are 3.7 MRI machines per million people in the United States compared with 0.9 in the former West Germany and 0.5 in Canada.

Moreover, we know that the constraints on the development and deployment of new and expensive technologies are weak in the United States. Our medical culture inculcates the value of being at the cutting edge of one’s field, and that often entails the use of the newest and often the most expensive techniques. Third-party payment systems of the sort that we now employ provide very little restraint on this cultural bias in the medical profession.

Nevertheless, one can go too far in ascribing excess costs to the American system’s penchant for the use of high technology medicine. High technology medicine is also sometimes high quality medicine, not wasteful medicine. Some claim that the failure to take account of the quality improvements involved in much of our technology causes medical inflation to be overstated in the United States. This argument challenges the equation of high technology with high growth in costs and waste at a fundamental level. It claims that the advancing technology is continuously raising the quality of American medical care, and that such quality improvement goes unrecognized. From this perspective, the growth in medical care costs ascribed to technological innovation is misplaced. Use of costlier new technologies is viewed as inflationary only because we have failed to recognize that we are getting higher quality medical care than we were in the past.

For now, perhaps the best that can be said is that the theory that supply of high technology medical equipment creates its own demand has some support. Controlling or reallocating that supply may be a significant aspect of cost containment in reforming American healthcare. As yet, however, no one has any good estimates of how much high technology medicine is too much or, how much additional value is being supplied to patients through its use.

C. Expenditures on the Dying

A particularly troubling area of possible overuse of medical care services is care devoted to dying, often elderly, patients. High intensity intervention in this area at best prolongs low-quality life, but more likely adds pain, suffering, and needless expense to the final days (or years, in the case of the chronically ill or brain-dead accident victim) of a patient’s life. Dr. Kevorkian’s suicide machine and healthcare proxies represent the potential for removing such malicious waste from our medical care system on a case-by-case basis.
What do we know about the extent of such waste? Health expenditure per capita does increase with age, and a large portion of Medicare funding is spent annually on a small percentage of its clientele. This pattern lends credence to the belief that we are skewing our resources not only toward the elderly and disabled, but also toward the worst-off of that group.

Unpacking such data, however, gives reason for pause. First, we would expect younger, healthier people to consume fewer resources than the elderly, and the sick elderly to consume more resources than their healthy counterparts. It is not an indictment of the healthcare system that it allocates resources to caring for the sick. Second, most of the Medicare recipients who consume a large amount of resources in a given year either get better or die, a fact that debunks the notion of widespread waste of the long-term, life-support variety. Third, from ages 65 to 85, medical care costs are inversely related to age, suggesting that practice patterns already reflect reduced intervention intensity as age increases. Finally, studies show that of the large amounts of resources devoted to the elderly, most are spent on two sub-groups: those not expected to die who do, and those expected to die who survive. This pattern suggests that case-by-case rationing is impossible.

It is difficult to imagine a system that neatly roots out waste of this nature. Any systematic reform of the way medical providers are paid or the malpractice code they face will, of course, influence the decisions made about care for the dying and disabled. “Who will say no” is not, however, a frequent subject of public political debate.

D. Excess Demand

Wasteful use of medical resources through excess supply and the demand induced by that supply finds a counterpart in explanations of waste that focus primarily on the demand side of the medical care equation. Excess demand takes two forms: excess demand for insurance and excess demand for medical goods and services.

There may be an excess demand for health insurance because of the way in which health insurance is supplied to most Americans. Employer-based health insurance receives a substantial tax subsidy through both its deductibility to the employer and the failure of the Internal Revenue Code to count health insurance benefits as income to the employee. As a consequence, it is in the interest of both the employer and the employee to substitute health benefits for

cash payments. The existence of this excess insurance may result in excess demand for medical goods and services because consumers who are fully-insured do not face the costs of use at the point of service. This latter effect is the second form of excess demand: an excess demand for medical goods and services themselves.

Other aspects of American life may also contribute to excess demand. Perhaps chief among these is the necessity of pursuing medical treatment or diagnosis in order to make claims for injuries either under workers' compensation or in the tort system. As mentioned above, those with workers' compensation claims may receive as much as nine times the amount of medical treatment for the same condition as those without such claims.24

There are two broad ways of addressing excess demand. The first is to attack the causes of excess insurance. One obvious remedy is reform of the Internal Revenue Code. In addition, managed competition proposals purport to address this problem by forcing consumers to choose among insurance plans that have different service and institutional characteristics as well as different price tags. Presumably consumers will be cost conscious in their selection of an insurance plan and will subsequently feel some of the economic effects of medical care use because they are no longer overinsured.

The second way to address excess demand is to decrease demand for goods and services at the point of service. The managed competition scheme relies principally on moving most patients and providers into HMO-style medical practice. Within the HMO, a primary care physician serves as a gatekeeper for consumers. The cost-saving potential of this gatekeeping function is unclear. The results of one experiment in the early 1980s suggest that gatekeeping alone has no effect on the use and cost of care.25 On the other hand, a RAND study in the early 1970s found 28% cost savings for the HMO group over the control group.26 No similarly-controlled experiment has been done in recent years, and because conditions of medical practice outside of HMOs have changed significantly, these results may no longer be valid.

In addition, recent surveys reveal that HMO subscribers are less satisfied with their care than are non-HMO patients.27 There is particular complaint about waiting time—a cost that does not show up in the provider's budget—and about the frustration of being shuttled from physician to physician as employees change and the needs of the patient alter.

Very few Americans have voluntarily chosen HMO practice, and HMO cost increases over the last two decades have paralleled those of other insurers.

26. Id. at 8 (citing Willard G. Manning et al., A Controlled Trial of the Effect of a Prepaid Group Practice on Use of Services, 310 NEW ENG. J. MED. 1505 (1984)).
or self-insured employers. Hence, the efficacy of relying on differential pricing to force HMO use, and on HMOs to constrain the costs of medical care, is unproven. We will provide further analysis of this cost-saving function in the later discussion of managed care.

The alternative and more direct means of dampening demand for medical care services at the point of service is through copayments or deductibles. The major study of cost-sharing arrangements was conducted by the RAND Corporation between 1974 and 1982. The findings of that study are clear. First, cost-sharing reduces the probability that an individual will seek care for any particular medical condition. Unhappily, the deterrent effect is strongest for the poor, especially poor children, who were at least 40% less likely to obtain care in the RAND experiment. Second, cost-sharing deterred enrollees from obtaining both appropriate and inappropriate medical care. Income, rather than health condition, determined the deterrent effect of cost-sharing systems, and patients in the cost-sharing plans had worse outcomes for specific treatable conditions, such as hypertension. Finally, the RAND study found that cost-sharing may prevent or delay entry into the medical system during a particular episode of illness. Cost-sharing does not, however, change the course of medical care once the patient has entered the system. During any given episode, the enrollees who had to share costs received the same services and medications as those who did not.

The general conclusion in the cost-sharing debate seems to be that cost-sharing saves dollars, but it is not selective about those savings. Cost-sharing cuts out wasteful use of medical resources as well as beneficial use of medical resources. Moreover, there are perverse distributional effects to cost-sharing at the point of medical service.

Providers and consumers must know much more about the likely efficacy of care, if consumers are to exercise sensible judgments when faced with cost-sharing. In addition, cost-sharing must be carefully tailored—and thus administratively more complex—to eliminate the pure income effects, if it is to be a method for selectively attacking waste rather than simply limiting access.

Finally, the experience of other countries suggests that the form of cost-sharing is highly important. Deductibles are difficult to administer, and they impose large administrative costs. Virtually all advanced societies that provide universal care with cost-sharing have abandoned deductibles in favor of


copayments. The task, then, is to target copayments in ways that eliminate waste rather than needed care.

E. Administrative Costs

Administrative costs is one area in which there is an ample body of data to support claims and substantial agreement concerning needed reforms. Yet even here, one must beware of assumptions that underlie claims of waste. In this area, most of the discussion has compared the United States' administrative costs with those of universal healthcare plans elsewhere. There has been particular attention to the Canadian system, which has a single payer in each province who insures the whole population and negotiates fees and budgets with providers. The General Accounting Office calculates that the United States would save over $65 billion a year in insurance overhead and administrative costs to doctors and hospitals if we adopted an analog of the Canadian system.31 In some sense, it is appropriate to view this $65 billion per year as pure waste. After all, if an alternative system design would eliminate the costs with no loss of medical services or information, then the system should change.

Nor is Canada unique. A look at the OECD entries reveals that in 1990, expenditures for medical administration approximated $150 per person in the United States, with our closest competitor being the former West German Republic, where per capita costs are $102.32 No other country reaches triple digits, and nations such as Canada and Australia expend only $23 and $38 per capita, respectively. Perhaps more significantly, the gap between U.S. and foreign administrative burdens has been growing, as the data for the U.S. and Canada in Chart 3 illustrate. Major savings can be achieved by reducing administrative costs in the American system.

It is important, however, to be somewhat more discriminating about categories of administrative costs than are some commentators. To put the matter slightly differently, when one speaks of administrative costs as "waste," one should be careful to note that these "wasteful" expenditures relate almost exclusively to the marketing, claims, processing, billing, underwriting, and regulatory-compliance arena associated with competitive private insurance. The reason that aggregate administrative costs are lower elsewhere is usually the absence or virtual absence of private insurance.

Germany's high administrative costs, for example, result from the continued existence of over 700 insurance funds which are regulated in order

to produce a uniform national system. In the U.S., whereas wasteful administrative costs account for perhaps 20% of American expenditures in privately insured healthcare services, the administrative costs of the public schemes, such as Medicare and Medicaid, are quite low. The per capita administrative costs for public systems is $29 per person, whereas costs in the private system are $150.33 That administrative costs for public insurance are less than private insurance administrative costs by a factor of five lends credence to the Canadian-based prediction of a 20% savings if America moves to a single-payer system.

These numbers also confirm that not all administrative costs can be squeezed out of any medical care system; medical care must be administered. Moreover, it is possible that certain aspects of the system are currently underadministered. We may not be spending enough, for example, on recordkeeping that would permit useful comparisons of outcomes or on the accurate monitoring of physician and hospital performance. The current system may also underfund healthcare planning and some aspects of health education.

A useful taxonomy of administrative costs is provided in Table 1. A shift to a single-payer system would have major effects on the administrative costs associated with the first and third rows in the table, but might have little impact on the costs associated with rows two and four. Managed competition has significantly less power to reduce administrative costs than does a single-payer plan. Selling and marketing costs would remain under managed competition. In fact, attempts to provide good information to consumers might actually increase costs related to the choice of an insurance plan. Nevertheless, there should be some economies as small insurers drop out of the private insurance market, because that is the sector in which administrative costs are highest.

F. Malpractice

Four factors have combined to make the malpractice debate particularly acrimonious: rapid growth in malpractice insurance premiums, physician misperception of the incidence of malpractice litigation, the attention given to large malpractice awards, and the conflict between two powerful interest groups—physicians and trial lawyers. As is often the case with acrimonious debate, the facts and figures thrown about by the participants require careful handling.

It is reasonably well-established that malpractice premiums have risen rapidly during the past two decades, and they now represent approximately 3.7% of physician charges.34 However, averages can be deceiving. While the


483
### Exhibit 1
Administrative Costs, By Function And Sector Of The U.S. Health Care System

<table>
<thead>
<tr>
<th>Function/Component</th>
<th>Health Insurance</th>
<th>Hospitals</th>
<th>Nursing Homes</th>
<th>Physicians</th>
<th>Firms</th>
<th>Consumers/Individuals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transaction-related</td>
<td>Claims processing</td>
<td>Admitting, billing</td>
<td>Admitting, billing</td>
<td>Billing</td>
<td>Tracking employee hires/ terminations</td>
<td>Submitting claims</td>
</tr>
<tr>
<td>Benefits management</td>
<td>Statistical analysis, quality assurance, plan design</td>
<td>Management Information Systems</td>
<td>Management Information Systems</td>
<td>Management Information Systems</td>
<td>Internal analysis</td>
<td>Tracking expenses eligible for reimbursement</td>
</tr>
<tr>
<td>Selling and marketing</td>
<td>Underwriting, risk, premiums, advertising</td>
<td>Strategic planning, advertising</td>
<td>Strategic planning</td>
<td>Advertising</td>
<td>Flexible benefit programs</td>
<td>Search costs</td>
</tr>
<tr>
<td>Regulatory compliance</td>
<td>Premium taxes, reserve requirements</td>
<td>Waste management</td>
<td>Discharge planning</td>
<td>Licensing requirements</td>
<td>Filing summary plan descriptions, COBRA obligations</td>
<td>Mandated benefit laws</td>
</tr>
</tbody>
</table>

*COBRA is the Consolidated Omnibus Budget Reconciliation Act of 1985, which includes provisions for continuation of coverage when an employee leaves a firm.*
average physician may pay $7,300 per year for malpractice insurance, the cost in some sub-specialities can approach $200,000 per year.\textsuperscript{35} Hence, a cost which is modest in the context of overall medical costs (less than 1%), can have significant skewing effects on the price of care in particular specialties, and on the willingness of physicians to enter or remain in those specialties.

Leaving the realm of malpractice insurance premia, the estimates of the costs of malpractice in American medicine become very soft. Estimates range between $1 billion and $25 billion per year, but tend to hover around $15 billion. A study by the American Medical Association (AMA) in the early 1980s published that figure, which was then revised downward to $12 billion to $14 billion in a later study by the AMA Center for Health Policy Research.\textsuperscript{36}

Although the AMA is not a disinterested party, its studies were not irresponsible. They were, however, seriously flawed methodologically in ways that are difficult to remedy. The findings are based upon the reports of physicians who were asked to estimate either practice changes or the increased costs that they incur due to fear of malpractice suits. Such an approach invites doctors to focus on malpractice fears as the motivation for their practice style. In the face of rapidly rising premiums, doctors have an incentive to overstate the costs of procedures as well. In contrast, a more recent poll asked doctors to list all of the factors that influence their practice routines or standards of care. This study revealed that the possibility of malpractice litigation ranked lower as a causal factor than a number of other influences such as continuing medical education, medical journals, and consultation with peers.

Because physicians are influenced by many factors in designing their practice standards, it is unlikely that we will be able to determine the true contribution of the fear of malpractice suits to the wasteful use of diagnostic techniques or medical interventions. Moreover, it is possible that malpractice litigation is producing no aggregate waste at all. The most recent and careful study of medical malpractice found no positive correlation between the incidence of malpractice litigation and iatrogenic injury and death, and that the fear of malpractice causes physicians to forgo some interventions.\textsuperscript{37} In short, the net effect of the fear of malpractice may be economically positive: it prevents iatrogenic injury and death as well as the direct costs of certain risky interventions.

No one who examines the American malpractice compensation system believes that it is performing well. While it is something of a straw man in the debate about costs and medical waste, the current malpractice compensation

\textsuperscript{35} Paul C. Weiler et al., A Measure of Malpractice: Medical Injury, Malpractice Litigation and Patient Compensation 2 (1993).
\textsuperscript{36} Roger A. Reynolds et al., The Cost of Medical Professional Liability, 257 JAMA 2776 (1987).
\textsuperscript{37} Paul C. Weiler et al., Proposal for Medical Liability Reform, 267 JAMA 2335 (1992).
system is admittedly ineffective. Recent studies suggest that only a small percentage of persons injured by medical negligence ever receive compensation, that who receives compensation and in what amount is something of a lottery, and that the administrative costs of the system (payments for litigation and lawyers' fees) are extremely high in relation to the amount of compensation received by plaintiffs. These effects, and the possibility of skewing the willingness of doctors to enter or remain in important fields like pediatrics, gynecology and anesthesiology, should inform any malpractice reform.

Recent experience suggests, however, that reform is politically difficult. The image of legislators preventing injured parties from recovering from negligent physicians is not terribly attractive to either legislators or the general public. Moreover, the trial lawyers' associations are quite powerful at both the state and national level. Fifteen states have begun reform with the modest experiment of requiring arbitration prior to filing a lawsuit. There is no evidence to date that this change reduces costs.

The state of Maine has begun a more ambitious program under which physicians are provided a defense to any malpractice action if they follow guidelines developed by groups of physicians in various specialties for dealing with particular conditions. The anecdotal evidence suggests that Maine physicians generally like the experimental arrangements and that the initiative has reduced the use of certain procedures, but that it has also increased the use of others. There are currently no cost estimates on the savings that may be achieved through the Maine approach.

A much more ambitious no-fault system has been proposed, but has been enacted only in Virginia and Florida, and only for one type of injury—neurological damage to infants during delivery. Such a scheme builds on the workers' compensation model and compensates victims whose injuries are caused by medical treatment, whether or not there is any demonstration of fault on the part of any particular physician or institution. Recoveries are more certain, but are limited to such items as medical costs, a percentage of lost wages, and predetermined amounts for particular residual impairments or death. Institutions and physicians could be experience-rated and their insurance costs adjusted accordingly. This adjustment would provide a financial incentive for better or more careful medical practice, and would focus remedial attention on those physicians or institutions that have higher risks of causing injuries to patients.

The no-fault scheme has much to recommend it, but it probably would not yield a reduction in the overall costs of medical care. Because a much larger

38. Id.
39. Edward Felsenthal, Maine Limits Liability for Doctors Who Meet Treatment Guidelines, WALL.
percentage of patients experiencing adverse effects would be compensated under such a scheme—albeit at much lower amounts—the total cost of the system might increase. There are currently no good cost estimates for this sort of major malpractice reform.

G. Straightforward Fraud and Abuse

We have previously provided a taxonomy of fraud and abuse, but it would be useful to provide some examples of the sorts of fraud and abuse that have been uncovered in recent years. Two relatively common forms of fraud and/or abuse are “upcoding” and “unbundling.” A recent example of upcoding was discovered in California, where hospital staff routinely upgraded the codes on patient file jackets, switching primary and secondary diagnoses to substitute more costly procedures and services for those that were actually administered to the patients. Because these actions were willful they were clearly fraudulent. On the other hand, there may be a large number of judgment calls in which hospitals or physicians having a choice between principal diagnoses will have an incentive to list the one that has a higher prepayment schedule. Estimates of the magnitude of upcoding or “DRG Creep” in American medicine range from 0% to 0.5% of healthcare expenditures ($0 to $4.2 billion in 1992), but are somewhat overstated because they ignore accidental and systematic “downcoding.” In short, contrary to the conventional wisdom based on a few horror stories, upcoding is not a big ticket item.

Fraudulent unbundling occurred when a group of Massachusetts anesthesiologists billed Medicare separately for the insertion of intravenous lines and catheters, although those procedures had already been reimbursed as part of billings for overall anesthesia services. Once again, fraudulent unbundling has many analogues of an abusive sort. Often a physician or hospital need not bill twice, merely separately, in order to increase the fees. No estimation of the systemwide magnitude of unbundling exists. Whether financially trivial or tremendous, however, unbundling is professionally damaging and worthy of reform attention.

Some of the largest instances of fraud in the medical system have involved even more blatant activities. For example, a physician and his sons billed Medicaid during the period of 1980 to 1987 for 400,000 phantom visits by Medicaid patients. They had programmed the diagnostic center’s computer to generate phony claims and back-up charts for as many as 12,000 fictitious visits a month. Another flagrant offense took place when a New York pharmacy bought prescriptions from Medicaid recipients that it did not fill, but for which

it billed the state. This pharmacy, in addition, simply fabricated prescriptions that were purportedly based on physician telephone requests.

As these cases illustrate, the opportunities for fraud and abuse in the medical care system are quite significant. Moreover, the complexity of the medical care system renders fraud and abuse easier for those willing to engage in it, and its decentralization adds to the monitoring problems of federal, state and private officials. In the Medicare context, for example, the federal government has agreed to allow a large variety of billing practices to accommodate local physician and hospital custom. The resulting complexity, combined with the relative stinginess of Congress in funding monitoring efforts by HCFA, helps to perpetuate undetected Medicare fraud and abuse.

That there is opportunity to engage in fraud and abuse and that monitoring resources are inadequate is generally accepted by students of the American medical care system. Few agree on how much this behavior costs and what to do about it. The extent of fraud and abuse in American medicine remains undetermined. A consensus has developed in the healthcare industry that fraud and abuse contribute roughly 10% to the total healthcare budget of the United States. If current spending is approximately 800 billion dollars, then fraud and abuse are said to account for 80 billion of that spending. This number has become an article of faith, and no one really knows what the correct number is. An estimate derived by Consumer Reports is presented in Table 2 on the following page.

Solutions to the problem of fraud and abuse are fairly straightforward. Increased funds for monitoring, system simplification, and elimination of physician self-referral are potential remedies. Each of these approaches is currently being implemented to some degree. But no one is carefully studying the effects of these initiatives. Of course, given that the baseline—the current costs of fraud and abuse—has not been established with any certainty, monitoring the effects of initiatives to reduce the level of fraud and abuse is difficult.

With respect to the issue of fraud and abuse, political judgment is critical. Legislatures and other officials have to decide how much they are willing to appropriate in an attempt to eliminate fraud and abuse, how invasive they are prepared to be with respect to monitoring, and how far they are willing to go to simplify the system and to change provider incentives. The effects of these measures on the level of fraud and abuse will not be easily quantified.

---

42. United States General Accounting Office estimate.
### Table 2
Consumer Reports Estimate of Medical Care System Waste, Fraud, and Abuse (1991)

<table>
<thead>
<tr>
<th>Area</th>
<th>Amount</th>
</tr>
</thead>
<tbody>
<tr>
<td>Administration</td>
<td>$ 70 Billion</td>
</tr>
<tr>
<td>Unnecessary Care</td>
<td>130 Billion</td>
</tr>
<tr>
<td>Fraud</td>
<td>52 Billion</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>$ 252 Billion</strong></td>
</tr>
</tbody>
</table>

Source: Consumer Reports, 1991
H. Managed Care: Palliative or Panacea?

If doctors and hospitals provide unnecessary care and use too many costly high-tech procedures; if patients, because of third party payments or another cause, demand excessive attention from the medical profession; if administrative costs eat up too much of the healthcare budget, and if opportunities for fraud and abuse are ubiquitous in the system; what is the solution to the waste, fraud and abuse problem? Increasingly, the mantra intoned by public and private healthcare officials and by numerous analysts is “managed care.”

Because it enjoys such support among practitioners, managed care deserves its own discussion as an answer to the waste, fraud and abuse problem. What is managed care? What do we know about its impact on expenditures? And how likely is it that existing managed care methods will have beneficial effects in a reformed American medical care system?

For purposes of cost containment, analysts usually group managed care arrangements into several categories. Group and staff model HMOs, as previously discussed, have been shown to produce one-time cost savings of significant amounts. The studies are dated, however, and it is not clear that the same percentage of savings over other arrangements would be achieved by newer HMOs.

A universal system of managing care based upon the group and staff model HMO would be extremely difficult to implement. First, these forms of HMOs cannot be established in sparsely populated areas where both residents and providers tend to be widely dispersed. For some substantial fraction of the population this is not a feasible form of managing care. Second, it is not clear that most Americans for whom HMOs are feasible can be persuaded to join them unless there is no real alternative. In 27 states, less than 2% of the population is in a group or staff HMO and in 35 states, less than 5% of the population are HMO members. In short, moving to the HMO managed care model involves shifting huge numbers of providers and consumers into an organizational arrangement that has been available for several decades, but that has not attracted the majority of Americans.

In any event, HMOs purportedly achieve cost savings by managing care through primary care physician gatekeepers. As previously discussed, this elimination of self-referral to hospital care or to specialists has substantial savings effects in comparison to fee-for-service providers without managed care. It is not clear, however, that the management of care is the key to HMO cost constraint. The real key may be the capitated nature of HMO financing. The HMO is both the provider of services and the insurer. It cannot collect

43. Goldberg, supra note 27.

490
additional funds should it overrun its budget, and therefore has a strong incentive to keep the provision of care within the budget that has been provided. This produces the waiting and shuttling that make HMOs unattractive to many.

A second form of managed care involves precertification and concurrent review of all hospitalization. This type of care management is considerably less effective than the staff or group HMO model in reducing costs. Whereas the HMO may reduce hospital use by 40% and total spending by 25%, precertification and concurrent utilization review has been found to reduce inpatient hospital costs by only 8% and overall medical expenditures by 4 to 5%.

While studies are inconclusive, other forms of managed care (such as IPAs, network model HMOs, point-of-service plans, and preferred provider organizations) may have little impact on overall medical care costs. Some studies have found small effects, but the amounts are negligible and may be offset by other costs. These other forms of managed care provide ample opportunities for providers to shift costs to others. As a consequence, there may be no real saving in the aggregate, even if the population enrolled in a managed care arrangement has reduced its overall medical care costs.

In short, group and staff model HMOs have a greater capacity than unconstrained or unmanaged care systems to save substantial amounts of medical expenditures. Pre-certification and concurrent review for hospitalization have similar but lesser effects on hospitalization and overall medical care costs, while all other forms of managed care have no demonstrable effect.

Applying effective managed care techniques to the entire population has the potential for significant savings. As of 1990, only about 80 million Americans were in programs that used an effective form of managed care, while nearly 180 million were in either ineffective managed care systems or arrangements that involved no managed care at all.

Assuming that effective managed care could approach the upper measure of savings of the group and staff HMO model, or 25% of current expenditures, the $200 billion question is whether the managed care arrangements that constrain costs are eliminating waste or merely eliminating access to needed services. This question leads to our final topic: How can we distinguish elimination of waste from barriers to access?

44. CBO, supra note 25, at 8.
45. Id. at 11-12. Some studies have found slightly higher savings, but it is not clear whether these studies controlled for out-of-hospital medical care for the same patients.
46. See id. at 8.
47. Verdon S. Staines, Potential Impact of Managed Care on National Health Spending, HEALTH AFF. (Supp. 1993) at 251.
1. Output Measures

Many of the standard interventions now used by American physicians have never been subjected to serious analysis of their effectiveness. Recent research on variation in medical practice styles across localities in the United States reveals dramatic differences in the rates at which particular diagnostic procedures are used or curative interventions made. Lack of testing of the efficacy of procedures, combined with an enormous disparity in their use, suggest that physician judgments about interventions or diagnostic tests are informed more by custom, professional bias, or individual economic incentive than by the efficacy of the care provided. If those are the bases for physician judgments, it follows that there is substantial waste in the system.

More research is underway on the efficacy of particular medical interventions, but a variety of technical difficulties in gathering and analyzing data exist. We are gaining knowledge about the efficacy of various treatments, but we are learning slowly and not without controversy.

Serious research in the area of medical interventions is extremely expensive, in terms of both dollars and time. “Gold standard” research results—the double-blind clinical trial with an untreated control group—are available for certain procedures, but the approach needed to produce them is notoriously expensive and takes years to complete. This type of research is routinely required for the approval of new chemical entities in the pharmaceuticals market, and accounts for 8 to 10 year delays and $100 million cost increases in the marketing of new chemical entities. (By the same token, of course, this sort of testing keeps many ineffective and dangerous preparations off the market.)

More importantly, it is unethical to use the gold standard test for many medical interventions. It is ethically problematic to assign patients to an existing, but perhaps less effective, form of treatment in order to get a control group for another treatment that might be more efficacious. Patients come to doctors to receive care, not to be used as guinea pigs in medical experiments. And, of course, doctors have an ethical obligation to give patients the best care that, in their judgment, can be provided.

Because of these problems, effectiveness studies must use panel techniques to evaluate the efficacy of various medical practices. The basic idea of these techniques is to get a cross section of experts in the field and have them evaluate the appropriateness or inappropriateness of particular forms of treatment. There is a significant amount of use of this method of research,

49. See, e.g., Albert L. Siu et al., Choosing Quality of Care Measures Based on Expected Impact and Improvement, 27 HEALTH SERV. RES. 619 (1992).
Healthcare Fraud

including that which provides the foundation for the Maine experiment previously described, in which following practice guidelines laid down by expert panels provides a defense against malpractice actions.

Nevertheless, the panel technique is controversial. Not all physicians are convinced by a decision of the panel that examines the medical records and reaches a consensus about the efficacy or inefficacy of various interventions. Moreover, it is difficult to find two patients who are identical in terms of their diagnoses, other physical and psychological circumstances, and available environmental supports. Reviews of files cannot capture all of the variables potentially relevant to treatment decisions. Hence, a panel judgment about what constitutes good treatment in the average case can be second-guessed every day of the week by individual physicians treating individual, highly peculiar patients. For these reasons, many are skeptical that we will ever have a reasonably complete set of output measures covering most medical practices, which could be used to distinguish between care that is beneficial and care that should be consigned to the dustbin of "waste."

From the available evidence, translating outcomes research into improved care is fraught with difficulty. In one study of American hospitals that had implemented outcomes research programs, only 4 of 31 reported improvements in quality of care as a result. Systematic reform will have to provide incentives to providers to adopt practice patterns that challenge not only their prior experiences, but also the maintenance of their current status and income levels.

Conclusion

Eliminating waste, fraud, and abuse in American medicine is not a panacea. Some believe that the United States, by dealing with waste, fraud, and abuse, can sharply reduce overall medical costs while preserving (or even improving) access to care. Unhappily, as this Essay shows, this belief is unproven and deeply problematic.

We know, however, that there are major problems with today's medical care system. We can estimate—within broad ranges—their cost to patient, insurer, and taxpayer. And we have reasonably good ideas about how to address some of these problems. For instance, there is no question that certain costly high-tech procedures are used too often and that a variety of reforms would reduce the incentives behind this overuse.

51. Id.
Secondly, we know that the consolidation of purchasing power in medical care would provide greater countervailing pressure on relentlessly rising medical fees. Moreover, the extension of capitation methods of payment would reduce pressures to increase the volume of services provided.

Administrative costs, as we have shown, are substantially greater in the United States than among our economic competitors. There is no mystery as to how we can reduce these costs significantly.

Outright fraud should be high on the agenda for reform. This is true even though we have no precise financial measure of the extent of current fraudulent practices.

Applying effective managed care techniques to a larger proportion of the population has the potential to cut costs. The problem is how to pursue these savings sensibly, while holding in check excessive expectations for their implementation.

New research on the effectiveness of care holds promise for improving the quality of our medical system as well as for restraining costs. Even more importantly, such research will challenge medical professionals in beneficial ways. It is worth pursuing even given the difficulty of quickly translating research into improved care.