Malcolm K. Sparrow

Fraud in the U.S. Health-Care System: Exposing the Vulnerabilities of Automated Payments Systems

IN 1993, ATTORNEY GENERAL JANET RENO DECLARED HEALTH-CARE fraud the "number two crime problem in America" after violent crime—a remarkable status for a category of white-collar crime. In 1995, FBI Director Louis J. Freeh testified that cocaine-traffickers in Florida and California were switching from drug dealing to health-care fraud. The traffickers had discovered that health-care fraud was safer, easier, and more lucrative than the drug trade, and carried a smaller risk of detection (Freeh, 1995: 2). In 1997 the *New York Times* reported that mafia families in New York City and New Jersey were abandoning their traditional lines of business (extortion and bid-rigging rackets) in favor of new criminal enterprises, including health insurance (Raab, 1997: A1, B4). In 2003, Columbia HCA, America's largest hospital chain, finalized a $1.7 billion settlement with the U.S. Department of Justice, the largest in history, following 10 years of investigation into an array of whistleblower allegations (Department of Justice, 2003). In July 2008, Abner and Mabel Diaz, a couple in Miami Lakes, Florida, pleaded guilty to fraud, admitting they had submitted to Medicare $420 million in false claims for medical equipment (Weaver, 2008: 1).

All sorts, apparently, find attractive opportunities in health-care fraud. But why steal from the health-care system? Perhaps because, at least in the United States, that's where the money is! No other nation on earth spends as much on health care as the United States, where health care expenditures for 2006 (the last year for which reliable figures are currently available) reached $2.1 trillion (CMS, 2006: 1). Projections for calendar year 2008 put total costs at $2.4 trillion, equivalent to $7,868 per person or 16.6 percent of GDP (CMS, 2007, table 1). The future of American health care looks even more expensive, with costs projected to outpace economic growth by an average of 1.9 percent per year, so that by 2019 health care will account for 19.5 percent of GDP (CMS, 2007: 1). Current spending levels for the United States are roughly double the average for other Organization for Economic Cooperation and Development (OECD) countries, and several countries (for example, the United Kingdom, Holland, Denmark, Japan) enjoy significantly better medical outcomes spending less than half as much.

Health-care economists, in their attempts to explain how America spends so much compared with others yet fares worse in medical terms and leaves roughly 16 percent of the population without health insurance coverage, pay little attention to the possibility that fraud contributes substantially to these costs. Scandals abound in which a person or business is discovered to have stolen millions of dollars from health insurers without supplying any legitimate medical care at all. Nevertheless, reliable data regarding the underlying extent of the problem does not exist. Each scandal can be interpreted as evidence of "a few bad apples, thankfully detected, amidst an otherwise sound system," or as "the tip of an invisible iceberg." Each stakeholder group can choose whichever interpretation it prefers, and the majority prefer *not* to consider the possibility that the integrity of major public programs—such as Medicare and Medicaid, each of which now consume more than $400 billion in public funds each year—has been severely undermined by criminal enterprise.

STRUCTURAL FEATURES OF THE U.S. HEALTH SYSTEM

The financial and operational structure of any given health-care system profoundly affects the types of fraud liable to emerge within it.

Transparency International's *Global Corruption Report* for 2006, focusing on corruption in health care, presents a wonderfully broad survey of health-system structures worldwide, and the distinctive patterns of corruption that emerge within them (Transparency International, 2006).

The following structural features of the American system help to account for the distinctive nature of the major fraud types that appear here:

- Fee-for-Service structure: Reimbursement for medical providers is mostly on a fee-for-service basis. Bills are presented to insurers by health-care providers, their staff, or billing agents; and the veracity of these claims is generally assumed, in the absence of any obvious indication to the contrary.
- Private-Sector Involvement: Private-sector entities provide the majority of health-care services. The insurers can be for profit, not for profit, or public. Purchasers of health-care insurance can be individuals, corporations, unions, associations, or public entities. For the majority of working Americans, the purchaser of their health-care insurance (that is, their employer), their insurance company, and their health-service providers are all nongovernmental entities.
- Highly Automated Claims-Processing Systems: The majority of health-care claims are now submitted electronically and processed automatically by computerized, rule-based systems. If the claims satisfy the criteria encapsulated within the *edits and audits* built into the system, then automatic payment follows, generally without any human involvement. Most claims paid, therefore, are not subject to any human scrutiny.
- Processing Accuracy Emphasized Over Verification: Claims-processing systems, designed with honest but possibly overworked and error-prone physicians in mind, do little or nothing to check that services billed were actually provided, or necessary, or that patients' diagnoses are genuine. Controls serve to ensure that claims are presented correctly and processed accurately: rule-based software checks that the prices charged are within appropriate

limits, that the treatments lie within the bounds of policy coverage, and that the combinations of diagnoses and procedure codes represent orthodox medical practice. The implications for fraud perpetrators, who may choose to submit claims that are totally unwarranted or fictitious, is that they must take great care to submit their bogus claims *correctly*. Provided they learn to make the claims appear normal, then they remain free to lie. They can fabricate entire medical episodes and submit the resulting bills without the patient's knowledge.

- Postpayment Audits Focus on Medical Appropriateness, Not Truthfulness: A small proportion of claims paid may be selected later by insurers for *postpayment utilization review* (PPUR). Fraud perpetrators can generally beat such audits by taking the simple precaution of fabricating medical records to match their fictitious claims. Prevailing audit practices at the PPUR stage involve mailing requests for copies of the relevant medical record to providers. A medical record, once received, is then reviewed and compared with the claim or claims it is supposed to justify. Providers have plenty of time (typically 90 days) to prepare and provide such medical documentation, and the subsequent "desk-audit" accepts the documents provided as true, and uses them primarily to test the orthodoxy and appropriateness of the provider's treatment patterns. Fraud perpetrators subject to such reviews may therefore be required to lie twice, and consistently, in order to pass these audits. All but the least sophisticated perpetrators routinely generate matching medical records at the same time they produce their fraudulent claims, just in case anyone ever asks to see them.

The predominant forms of fraud, given this combination of factors, consist of overprovision of services based on false or exaggerated diagnoses, and billing for services that were not actually provided. Claims involving some material misstatement or deception are broadly termed *false claims*. The deception may relate to the diagnosis for the patient, or to the treatments provided, or both. Diagnoses and proce-

dures may be exaggerated (which is called "upcoding"), or totally fictitious. The false claims problem remains the most blatant, extensive, and poorly controlled fraud issue within the health system.

The false claims problem is not the only fraud problem, of course. In the last five years significant attention has been paid to the behavior of pharmaceutical companies—in particular, to aggressive and deceptive advertising practices, off-label promotion of drugs (promoting uses not approved by the Food and Drug Administration), price manipulation, and improper "detailing" techniques (offering illegal inducements for physicians to prescribe specific drugs).

Alternate financial structures have also been introduced within the health industry that fundamentally alter the prevailing incentives, and thus alter the types of fraud liable to appear. Managed care programs that involve *capitated payments* reimburse providers a fixed amount per patient per month, regardless of the level of service the patient consumes. Managed care organizations, receiving capitation payments, therefore acquire an incentive to deny services, or underprovide, rather than to overuse or overbill. When managed care gained a substantial foothold in the industry, many officials believed it would "structurally eliminate fraud." In fact, what happened—given that as the structure changed many of the same bad actors remained in place—was that those inclined to cheat and steal quickly adapted to the new incentives. Fraud, under managed care, involves denial of services, substandard care, and construction of a daunting array of logistical and administrative obstacles for patients to navigate in order to be served. The resulting patterns of abuse involve diversion of resources away from frontline health-care delivery, and bring serious consequences for patients' health, sometimes death. Within the managed-care arena, as fraud became more dangerous to patient health, it also became harder to detect and to prosecute. (Sparrow, 2000: 98-113)

The advance of managed care has slowed somewhat over the last five years, and even slipped into reverse in parts of the country. In 2008, fee-for-service payments still account for the majority of health-

care spending within the United States, and the false claims problem remains the most pressing uncontrolled fraud issue.

The majority of cases brought against major corporations involve false claims of one kind or another, and most are revealed through the actions of whistleblowers rather than the operation of routine detection systems. Most whistleblowers are employees of the offending company, and thus well placed to report the business policies and practices at issue. Over the last decade, the *qui tam* provisions of the federal False Claims Act have emerge as a principal tool in the government's efforts to protect public programs from fraud. In 1986 the federal False Claims Act amended the original civil war version of the act to extend its reach beyond defense procurement fraud and into the health care arena. Health-care cases now routinely dominate the caseload of *qui tam* (or whistleblower) suits filed with the Department of Justice under the False Claims Act. There is apparently no other area of federal spending so vulnerable to fraud, and so deeply infected.

UNDERLYING CHARACTERISTICS OF THE HEALTH CARE FRAUD PROBLEM

Putting aside the particularities of the American setting for a moment, health-care fraud more generally exhibits several properties, some generic to white-collar crime and some particular to the health-care setting, which complicate the task of controlling the risk. The combination of these several properties makes health-care fraud an extraordinarily intractable problem, and may help account for its persistence and scale.

Invisible by Nature

Well-designed fraud schemes remain invisible in perpetuity, and hence the underlying scope of the problem remains unknown. The class of invisible risks is familiar within the field of criminology, and includes a range of problems that generally pass undetected, or unreported, or seriously underreported. White-collar crime and corruption generally fall in this category, as do consensual crimes such as drug-dealing,

bribery, illegal gambling, and prostitution; also underreported crimes such as domestic violence, date rape, and child abuse (Sparrow, 2008: 181-198).

The task of controlling such invisible problems is complicated by the underlying uncertainty about the pervasiveness of problem, and about the ways in which it might or might not be concentrated. Authorities' knowledge of the issue derives from the small proportion of cases detected or reported, and investments in control are pegged to the magnitude of this visible sliver. Low levels of investment result in continuing low levels of detection, which provide false assurance through a paucity of cases. Underinvestment in the control enterprise becomes a circular trap, perpetuating itself. For invisible problems, significant underinvestment remains the norm. Targeting may also suffer from circularity, as authorities pay more attention to areas in which they have found cases before. In their efforts to understand the underlying problem, control strategies rely too much on the few cases that come to light, not realizing that these may represent a small and biased subset of the underlying issue, largely influenced by *where* and *how* they have looked for it in the past.

Available metrics—such as the number of cases detected, or volume of claims denied, or total value of settlements obtained—are all ambiguous too. If the level of detected fraud doubles, this could mean detection has improved or that the fraud situation had deteriorated dramatically. In the absence of measures unambiguously reflecting the underlying level of the problem, changes in the readily available metrics remain open to diverse interpretations.

In relation to insurance fraud, reliable metrics can be obtained. Standard measurement techniques demand rigorous audit of a random or representative sample of claims, followed by extrapolation of the sample's overpayment rates to the universe of transactions. Some forms of overpayment are easier to discover than others. Errors—either committed by the submitter or by the processing operation—are generally easier to detect than fraud. To be useful as a fraud-measurement tool, audit protocols used on such a sample must be rigorous enough to

uncover all the known types of fraud, and preferably thorough enough to reveal novel forms as well.

Few attempts have been made within the health-care field to generate any reliable estimates of fraud-loss rates. Several studies have been conducted to measure "error rates" or "overpayment rates" within various programs; but the audit protocols involved are generally too weak to uncover most types of fraud—even the familiar types.

Best known, perhaps, have been the "Medicare Overpayment Rate" studies, conducted by the Office of Inspector General (OIG) (for the Department of Health and Human Services) from fiscal year (FY) 1996 through FY 2002. These involved statistical valid samples of recently paid claims in the Medicare program (a federally operated program covering beneficiaries who are either over 65 years old, chronically disabled, or suffering from end-state renal disease—that is, dialysis patients). The OIG studies employed an audit protocol resembling a typical postpayment utilization review. These desk-based audits did not involve any face-to-face contact with providers, no contact at all with the majority of patients, and medical records mailed in by providers were assumed to be truthful. Thus the overpayments detected by the studies would not have included the majority of fraud losses, except for those cases where a fraudulent provider refused to mail in supporting (and suitably fabricated) medical records. Nevertheless, the first of these OIG studies, reported in 1997, showed an overpayment rate of 14 percent, equivalent to $23 billion in annual losses from the Medicare program. These findings shocked Congress, and the nation. In subsequent years, the measured overpayment rates came down (see table 1), providing some comfort for alarmed taxpayers.

These figures provided the basis for the Clinton administration's claim (which left office at the end of 2000) that it had correctly identified health-care fraud as a problem, and had cut the problem in half during its time in office. But the weakness of the audit protocols employed in these studies make available a range of other plausible explanations for the observed decline:

Table 1. Medicare Overpayment Rates (by Fiscal Year)

Financial Year	Point Estimate	Extrapolated Loss Rate
1996	14%	$23.2 billion
1997	11%	$20.3 billion
1998	7.1%	$12.6 billion
1999	7.97%	$13.5 billion
2000	6.8%	$11.9 billion
2001	6.3%	$12.1 billion
2002	6.3%	$13.3 billion

Note: Figures reported annually by the Office of the Inspector General, Department of Health and Human Services.

a) the overpayments captured consisted mostly of processing and documentation errors, and increased automation of the claims process naturally reduced these categories of errors over time.

b) fraud perpetrators who happened to be caught in these samples learned, over time, that if they did *lie twice*—by supplying a fabricated medical record to match the fabricated claims—then the authorities would make no further inquiry into the matter and deem the claim payment "correct." So their initial reluctance to send in fabricated medical records (reflected in low response rates at the outset) diminished over time as they became more familiar with the limited extent of the audit.

Early in 2000, the General Accounting Office (GAO) was asked by the congressional House Budget Committee to examine the methodology the OIG had been using to estimate Medicare overpayment rates. In a letter to the committee chairman, Representative John R. Kasich, the GAO reported:

> Overall, our work shows that because the methodology was not intended to detect all fraudulent schemes such as kickbacks, and false claims for services not provided, the estimated improper payments of $12.6 billion would have been greater. How much greater, no one knows. . . .

> It was not designed to identify or measure the full extent of levels of fraud and abuse in the Medicare program. The HHS OIG testified [in July, 1997] that the estimate of improper payments did not take into consideration numerous kinds of outright fraud such as "phony records" or kickback schemes. The methodology assumes that all medical records received for review represent actual services provided (GAO, 2000).

Despite the clear admission that these studies did not capture most forms of fraud, and in particular would not capture false claims, which are the most obvious and central form of fraud, the OIG continued to use the same audit protocols in subsequent years. The OIG argued that it had to employ the same methodology year after year in order to make the results comparable, and for any trends observed to be meaningful.

In January 2003 the OIG discontinued the Medicare overpayment measurement program, asking the Medicare agency itself to run an equivalent annual study. The Centers for Medicare and Medicaid Services (CMS) continues to use weak audit methodology in its Claims Error Rate Testing (CERT) program, and hence nobody has any reliable indication of overall fraud loss rates for the Medicare program.

Medicaid programs (which serve the poor and are administered by the states rather than by the federal government) display greater variability in policies and procedures than the centrally administered Medicare program. Several states have designed and conducted Medicaid overpayment measurement studies in recent years, similar in character to the Medicare studies. In general, these tend to use valid sampling techniques, but fairly weak audit protocols. The federal government, through its Payment Accuracy Measurement project (CMS, 2004), has sought to encourage broader use of loss-measurement by state Medicaid agencies, but does not push the states to use the kind of rigor necessary to capture fraud.

One might imagine that private sector insurers, driven by their bottom-line and fiduciary responsibility to shareholders, would do a much better job of exposing and dealing with fraud than their public and not-for-profit counterparts. In fact, private insurers almost *never* conduct valid loss-rate studies. They defend this particular omission with a set of familiar arguments as to why measurement is either impossible (and therefore should not be attempted), or undesirable. The most frequent justifications given for this failure are as follows:

- Rigorous audits on a random basis are fundamentally unfair, and not an appropriate way to treat medical providers.
- We do not have the time or the money to waste conducting "academic" research.
- All available audit and investigative resources are consumed following leads, and it would be irresponsible to impede the progress of investigations by diverting resources.
- We get better return on investment focusing all of our audits on known high-risk areas and high-risk players. It is wasteful to apply such techniques on a random basis.
- No audit protocol could possibly capture all the possible types of fraud, and therefore it is impossible in any case to measure fraud in any reliable way.
- Fraud involves a state of mind and requires criminal intent. No study could ever determine that.

The antidote to all such arguments, of course, lies in the potential value of the information that rigorous measurement studies might produce. Reliable information about loss rates would give authorities the chance to resolve the otherwise persistent ambiguity about the scope and nature of the problem. Such information could lead in turn to the possibility that investments in control might be pegged in some more sensible way to scientifically or statistically valid estimates of fraud losses.

For any *invisible* problem, effective control begins with valid measurement. For health-care fraud, control breaks down at this very

first hurdle. No one knows quite how bad the situation has become, and industry practices seems to reflect a broad reluctance to find out. Exposing the scale of the problem, after all, might involve a dose of very bad news; and news of major breaches in the integrity of health programs tends to alarm shareholders, drive down stock prices, reveal past failures, and alarm the public.

Conscious Opponents

Fraud also belongs within the class of risks that involve conscious opposition: risks that have a *brain* behind them. Many classes of risk, such as occupational and transportation hazards, as well as most environmental threats, do not have a brain, as such, behind them; and thus these risks do not exhibit adaptive behavior designed to circumnavigate control initiatives or enhancements. Eliminate a specific occupational hazard, for instance, and it does not go searching for another way to kill you.

In this regard, fraud perpetrators belong more naturally with drug smugglers, terrorists, computer hackers, and thieves. Such groups constantly study the relevant defenses, adapt quickly to changes in those defenses, and thrive on novelty and surprise.

The presence of adaptive opposition complicates the challenge of control. The controllers must engage in a game of intelligence and counterintelligence. They must take pains to learn what the opposition is thinking, or what they *might* be thinking. They must respond quickly to the opponents' initiatives, and hassle them out of the fray by forcing them to adapt often. They must seek out and exploit specific vulnerabilities of the opponents' strategies, using such points of vulnerability as resource-efficient opportunities to sabotage their enterprise. They must retain an air of mystery and unpredictability, and vary their detection methods so the opposition can never be sure where, or how, they are looking (Sparrow, 2008: 199-216).

One does not often hear health-care authorities speaking this language, or thinking in these ways. Claims-payment systems, by design, are utterly predictable and transparent. If a claim for payment

is denied, helpful computer-generated explanatory notices explain the reasons for the denial so that the claim submitter can get it right next time. Everything is geared toward the honest physician, possibly error-prone, but basically well intentioned. The result, from the perspective of fraud perpetrators, is a target that exhibits all of their favorite qualities: it pays fast, because it is required to by law. It is perfectly predictable (so if it pays one claim without a hiccup, then it will reliably pay 10,000 similar claims for other patients exactly the same way). If a fraud perpetrator bills "incorrectly" and receives a denial, then the system explains the mistake and teaches how to fix it. And even when the system denies a lot of claims from one provider, it does not become suspicious. Provided the claims submitted are fashioned to reflect medical orthodoxy, then there is very little risk of encountering a human being at all, let alone a criminal investigator.

Health industry practices tend to miss or underestimate the significance of the fact that they confront opponents, sometimes quite sophisticated ones. Insurers place too much trust in the latest and most comprehensive rule-based software packages, imagining that once they have put these in place, their system is properly protected. They underestimate the extent to which the opposition immediately begins testing and trying the new controls, and just how quickly they will determine its parameters and locate its vulnerabilities.

Postpayment claims review operations similarly undervalue any broad or exploratory casting-about by which they might discover emergent problems never seen before. Instead, like fishermen of habit, they fish in the same waters day after day and month after month, because that is where they have caught fish before. One known high-risk area can dominate their thinking and consume their time to such an extent that authorities can remain completely oblivious of entirely new patterns of fraud, which can therefore grow to significant proportions within other industry segments completely out of sight.

The health industry, in addressing fraud, confronts conscious and adaptive opposition, with the following implications:

- Historical experience provides unreliable guidance in identifying risk areas for the present and for the future. Those responsible for fraud control should expect and anticipate novelty from the opposition, and they must design analytic, audit, and investigative strategies with that in mind.
- Fraud controls should include routine use of intelligence-gathering techniques such as surveillance, undercover shopping for medical services, development of informants within corrupt networks, and making deals with convicted perpetrators in exchange for information and intelligence about fraudulent practices.
- Insurers and investigators should incorporate a counterintelligence mindset in their control operations, concealing parts of their detection capabilities, altering thresholds and focus areas constantly, and incorporating degrees of randomness and unpredictability so their methods cannot be reverse-engineered by the opposition.
- Those operating automated claims-processing systems should place less faith in state-of-the-art, but static, rule-based systems and software packages. Instead, they should invest in analytic versatility, and stress nimbleness and rapid response to emerging patterns. Rather than technology-driven control systems, they should develop human-driven, but technically sophisticated, *intelligence* operations.

It takes committed fraud perpetrators at most a few weeks to fathom the nature of new controls, and to redesign their scams accordingly. It takes at most a few months for these newly adapted fraud methods to spread across the country. But it can take authorities *years* to make the legislative, policy, or system changes necessary to suppress specific fraud threats. Health-care payment systems, as targets for fraud, are not only fat and rich, but tend to be *very* slow moving indeed.

Risk Control in a Hostile Setting

Risk control, and crime control in particular, is easier to do when the control function lies within an agency set up to do precisely that. But

when risk control functions appear as ancillary or peripheral to an organization's core enterprise, then those responsible for control may find the general culture and assumptions of the organization somewhat at odds with, or even hostile to, their purposes and methods. Classic examples of such cultural discomfort include the task of providing security in an academic environment, or controlling embezzlement within a charitable organization, or dealing with the risk of child or sexual abuse within a religious community. The prevailing organizational assumptions of trust, and the preference for guidance as the primary method for influencing behaviors, often seem at odds with the less charitable assumptions and harsher methods required for effective crime control.

The core task of health-care systems is to deliver health care, not to carry out fraud control. The crime control imperative comes along later as an uninvited guest, and the rest of the system would rather not hear about it, or hear from it, at all. The awkwardness of the fraud control setting is particularly acute within the health-care industry, for a variety of reasons.

First, insurance companies and government health programs do not generally engender much sympathy as victims of fraud. Their own conduct (for example, in relation to the payment of legitimate claims), often criticized as substandard or unethical, makes defrauding them seem more socially acceptable. Segments of the public view stealing from insurers as a natural form of revenge, either against ruthless and heartless businesses, or against wasteful, inefficient, or incompetent government agencies. Of course, the view that fraud actually hurts the insurers misses the point that—assuming the fraud remains invisible, and the insurers can therefore pass on the cost to those who pay premiums—the real victims of fraud turn out to be the patients, subscribers, and taxpayers.

Second, society holds medical practitioners in high esteem, recognizing the rigor and intensity of their training. Professional judgments made by physicians cannot generally be critiqued, except by another qualified physician. Medical associations fight vigorously

to prevent their members' judgments from being assessed or second-guessed by anyone else on administrative or financial grounds. And if by chance fraud investigators should come sniffing around, medical professionals tend to adopt a haughty position, pointing out that these investigators have no medical training and are therefore not qualified to understand, or render judgments about, diagnostic or treatment decisions. Investigators frequently encounter medical practitioners as arrogant and condescending, counting on their professional status to afford them protection or immunity. Even when investigators persist and make their case, prosecutors may be reluctant to pursue cases that rely in any material way on questions of medical appropriateness or necessity.

Third, medical professionals display an extraordinary reluctance to condemn the most egregious acts of their peers. Even when a physician or other provider is convicted of outright criminal fraud, and even when their actions have had profound adverse consequences for patients' health, their professional associations scarcely ever speak out against their conduct. One has to wonder why it would not be in the interests of the profession, and professional associations, to step forward and explain to the public, quite deliberately, that this person was genuinely one bad apple, and that the rest of this profession abhors what they did. But this almost never happens.

One plausible explanation for this failure relates to the range of possible malfeasance. The spectrum of misbehaviors available to medical providers is rather long, continuous, and not easily divided. At one end lie minor forms of code manipulation designed to compensate for unfairly low reimbursement levels; or a little diagnosis-substitution for the sake of the patient, so that treatments required can be covered by the insurance policy. Such actions contravene the rules but seem to have some plausible social justification. At the other end of the spectrum lie unambiguous, even rapacious, fraud scams that may leave in their wake a trail of victims. The difficulty, if anyone in the profession wants to condemn anything at all, lies in drawing satisfactory dividing lines between what the *criminals* did, and what *they* do, or might do someday.

The fuzziness of the lines between fraud, abuse, waste, overutilization, helping patients circumnavigate unfair policy restrictions, and differences of opinion about medical orthodoxy make it dangerous for the medical profession to condemn anything. Who can tell where such condemnation, once mobilized, might end? Keeping quiet, or emphasizing the extraordinary difficulties under which medical professionals labor, is a much more comfortable course. As a result, those engaged in the fraud control task end up convinced that the entire industry opposes them, despises them, and has no interest in fraud control.

Fourth, societal trust in physicians extends, by association, to a broad range of ancillary provider groups not subject to the same rigor in training and not bound by stringent codes of professional ethics. Investigators see medical equipment suppliers, home health agencies, medical transportation companies, behavioral health clinics and billing agencies as businesses, run by businessmen, for profit. They regret society's assumptions—based purely on the fact that these groups operate within the health industry—that such businesses could or should be trusted to subvert their own private economic incentives to any higher-level professional or ethical obligations. Nevertheless, major payment systems within the industry treat such groups in basically the same way as physicians, accepting the claims they submit as true, and paying them on trust without any routine validation that the services billed were necessary or were actually provided.

Fifth, highly automated claims-processing environments emphasize efficiency and timeliness, not caution and risk control. The responsibilities for processing efficiency and for fraud control lie with different officials, and within different organizational departments. Culturally, the two purposes seem at odds. Process management focuses on the administrative cost of processing a massive volume of claims, and doing so in a timely manner. Fraud control is more interested in finding and examining exceptions, and holding payments up where necessary to reduce the organization's exposure.

One might imagine that the simple concept of *return on investment*, applied to investments in caution and scrutiny, would adequately

instruct health insurers how best to integrate these two competing imperatives. Typically, every dollar spent on protecting the integrity of the system pays off handsomely, saving $10 or more in terms of funds paid out. So why do investments in control not escalate naturally to the optimal level (at which the marginal dollar spent on control returns just one additional dollar in savings or recoveries)? The answer is both legal and organizational. Seldom is one official in a position to consider the return-on-investment equation. Officials are responsible for one thing or the other, and each official—with his or her own metrics and motivations—gets in the way of the other. But processing efficiency always wins, because of the massive volumes and visible embarrassment to the organization if the system does not keep up. Moreover, savings from gains in processing efficiency are visible, concrete, calculable, and certain. By contrast, savings from fraud detection or fraud reductions, given the invisible nature of the problem, are uncertain, highly ambiguous, and cannot be guaranteed, even though they could potentially be much larger.

In the case of Medicare, the funds being paid out are actually legally distinct from the administrative costs of paying them. The payments themselves come from the Medicare Trust Fund, whereas the processing costs are drawn from general tax revenues. This legal separation makes it virtually impossible to set control investments at an appropriate level. Fraud control costs, rather than being weighed against reductions in fraud losses, form part of a zero-sum game with other administrative functions (for example, handling beneficiary enrollment, queries, complaints), all of which are completely inescapable, and all of which draw on the same general pool of administrative costs.

Even where there is no *legal* separation between claims expense and processing expense, organizational divisions of labor seem to produce the same dysfunction. Fraud control functions lose out in terms of budget, and in terms of influence over operational policies. Culturally, in a highly-automated and massive-volume environment, fraud control is just a nuisance.

All of these factors exacerbate the cultural hostility to the fraud control function. Fraud investigators and analysts often express the frustration that even their own bosses behave as if they would rather not hear from them. Senior executives prefer no mention of fraud, because they do not know much about it, they have not been trained how to think about controlling it, and any fraud issue or case that does pop up gets in the way of an otherwise smoothly functioning business model and embarrasses the enterprise. Fraud control becomes a miserable task, unappreciated, stressful, and loaded with organizational tension. Those responsible for fraud control soon learn that, when it comes to fraud at least, no news is good news.

CRITICAL FAILURES OF CONTROL: THE MACHINERY

An examination of the machinery trusted by the health-care industry to control fraud shows it to be profoundly inadequate for the task (Sparrow, 2000: 162-182). Claims-processing systems incorporate extensive suites of rule-based checks (edits and audits) to make sure services have been billed correctly, priced reasonably, and fall within the bounds of medical orthodoxy and policy coverage; but these systems do nothing to verify truthfulness. Prepayment medical review, conducted by nurses or claim specialists, provides an opportunity for examination of selected claims in much greater detail, and by a person—but the claims for review are those picked out of the processing stream by the computerized edits and audits. So, if a fraud perpetrator learns to bill correctly and thereby beats the edits and audits, then their claims effectively bypass any chance of human inspection, and will be paid.

Postpayment utilization review provides an opportunity, later, for the aggregate billing patterns for any particular provider to be compared with their peers. Aggregate billing patterns that deviate from statistical norms for any one specialty, once observed, may trigger a broader audit of that provider's practice and billing behavior. The auditor will draw a sample of the selected provider's recently paid claims, and ask the provider—by sending them a request in the mail—to provide medical records and other relevant documentation (for exam-

ple, test results) to support the claims. Postpayment utilization review does sometimes uncover fraud, but seldom. The PPUR function is more focused on medical orthodoxy and appropriateness; and the audit methods are quite trusting. Providers are typically given up to 90 days to supply the necessary documents, and what they supply is assumed to be genuine. If the documents match the claims, the provider will most likely pass the audit. If the provider fails to provide documentation, or provides inadequate documentation, then those particular claims may be reversed, and the payments adjusted. If the provider shows a pattern of poor documentation, most often they will be "educated" about the need for proper documentation in the future.

Perpetrators of outright criminal fraud do not much fear PPUR for a number of reasons. First, they know that PPUR only detects fraud where fraud produces anomalous billing profiles. If fraud perpetrators fashion fake billing schemes to mirror legitimate billing patterns, then PPUR will never find them. Second, PPUR units are very small, and can only pay attention to a few industry segments at a time. They look mostly where they have looked before, or where the last scandal was. Novel scams are liable to remain completely outside of PPUR's sights, and for a good long time. Third, when PPUR does examine a particular industry segment, it will select only the extreme outliers for audit. Fraud perpetrators can fashion their schemes to avoid these statistical tails, and so stay out of sight. Fourth, even when PPUR does find an anomalous billing pattern, it tends to employ soft and friendly methods, providing guidance and instruction to providers on how to correct their billing behaviors for the future. Fifth, PPUR works long after the fact, from 6 to 18 months after claims have been paid. Fraud schemes can net millions of dollars within such a window, and the operators can shut down and shift to alternate provider numbers as soon as anyone starts asking questions.

The remaining piece of a health insurer's fraud-control apparatus is the Special Investigative Unit (SIU). Most insurance markets require the existence of such units, but do not require any specific performance from them. SIUs employ former police and other investigators, and

are therefore more fraud aware than the rest of the organization. SIUs, however, are tiny; and most of them sit passively on the end of fraud-referral systems. The referral systems from which they get their work (consisting of the other parts of the organization) are not focused on fraud, and therefore the levels of fraud detected and referred to SIUs remain extremely low. SIUs may apply professional investigative skills in a case-disposition mode, but generally do not engage in intelligence work or use investigative field craft to monitor for emerging fraud patterns and to diagnose fraud concentrations or patterns. The performance metrics for SIUs include cases opened and closed, and dollars recovered or settlements obtained as a result of specific investigations. They do not generally include anything relating to *fraud problems identified and suppressed* and their contribution to effective fraud risk-control is diminished by their reactive and case-based stance.

The health-care industry generally relies on these four standard pieces of apparatus—the edits and audits, prepayment medical review, postpayment utilization review, and special investigative units—to provide protection against fraud. What this set of functions manages to accomplish, given the typical resource levels and configuration, is to provide reasonably good protection against *seeing* fraud. Fraud perpetrators with any degree of sophistication at all can easily remain out of sight.

Critical Failures of Control: The Mindset

Even while fraud control *machinery* remains inadequate, one might hold out hope for better control in the future if the fraud control *mindset* were in good shape. If authorities understood what was needed, and knew how to make the case for it, then surely the situation would improve over time. Sadly, there is plenty of evidence that even those officials and organizations most critically placed to address health-care fraud still fail to grasp the nature of the beast, and hence fail to wrestle with it effectively.

The last 10 years has seen an extraordinary series of reports produced by the Office of Inspector General for the Department of

Health and Human Services. The OIG is responsible for overseeing all of the federally funded programs that DHHS operates. These include the two largest public health care programs: Medicare (for the elderly) and Medicaid (for the poor). Medicare, being federally administered, receives the most scrutiny at the national level. Medicaid is funded through a combination of federal and state expenditures, and is administered by state agencies. The OIG is the primary agency responsible for overseeing the integrity of the Medicare program, and shares oversight of Medicaid programs with other state-level authorities. But both of these programs now cost more than $400 billion per year, and so there is a great deal of public concern about the need to protect the funds flowing through these programs.

According to OIG reports, several different categories of patients, none of whom should be getting treatment under these programs, have been showing up in significant numbers within paid Medicare and Medicaid claims. The most obvious embarrassment involves treatments apparently rendered to patients who were already dead on the date of treatment. In March 2000, the OIG published its investigation into provision of medical services to Medicare beneficiaries after their dates of death. The OIG audit methodology was straightforward enough: obtain up-to-date records of death from the Social Security Administration, and search the paid Medicare claims files for services delivered after death. They quickly found $20.6 million in such claims, paid in 1997 (OIG, 2000: 1). For some of these claims there was a plausible "error" story: the rental for a wheelchair or a series of monthly capitation payments had not been stopped when the patient died, and hence payments after death continued when they should have been shut off. But these cases represented a small minority, and can easily be filtered out of the analysis. A significant volume of the claims showed *new* treatments beginning for a patient, more than a month after they had died.

Dead patients also showed up in Medicaid claims around the country. An OIG report in 2006 summarized findings from 10 different states, revealing $27.3 million in Medicaid payments for services after death (OIG, 2006: 3).

Other patient groups that also should not show up in paid claims, but apparently do so remarkably often, include patients who have previously been deported, and which US Citizenship and Immigration Services (CIS) records show had been banished from the country prior to the reported treatment dates, and prohibited from returning. How did these patients manage to receive their treatments here within the United States, and at public expense? In March 2002, the OIG reported finding 43 deported Medicare beneficiaries for whom fee-for-service claims had been received and paid after the recorded date of deportation (OIG, 2002(a): 1-2).

Similarly, patients who are incarcerated generally ought not to show up in Medicare and Medicaid paid claims. Health care for prisoners is provided through prison systems, not by Medicaid or Medicare. There are a few specific exceptions to this general rule, relating to hospital and other treatments delivered outside the prisons. The OIG has conducted investigations into both Medicare payments (OIG, 2002(c)) and Medicaid payments (OIG, 2002(b)) apparently made "in error" for patients in prison.

All of these reports from the Office of Inspector General basically follow the same logic. They point out that the requisite data about deaths, deportations, and incarcerations is available somewhere within government; therefore the Medicare and Medicaid programs can and should do a better job of obtaining it from the relevant agency in a timely fashion, and incorporate it into the claims processing edits and audits, so that such claims could be rejected up front by the payment system.

This approach typifies the prevailing government view that overpayments in health care systems represent processing errors. The cure, once an overpayment problem comes to light, is to fix the process. The OIG seems to understand that such claims—for which there can be no legitimate explanation—should not be paid; but they do not seem to understand that such claims ought never to be generated and submitted in the first place. The obvious question, for any astute observer, would surely be "How on earth did these claims get generated? What type of business practice produces such nonsense?" The most striking feature

of the OIG reports on each of these categories of implausible claims is that they pay no serious attention to these questions. They focus on claims *payment,* not on claims *production.* They assume the problem, and therefore the solution, lies within government's technology, policies, and processes. None of these reports treat seriously the possibility that these claims result from fraudulent billing practices.

In July 2008, another group came to light, adding to Medicare's public embarrassment. The Senate Permanent Subcommittee on Investigations revealed the presence of *dead doctors* within Medicare's paid claims. The subcommittee's investigation revealed that from 2000 and 2007 between $60 million and $92 million was paid for medical services or equipment that had been ordered or prescribed by dead doctors. In many cases, the doctors had been dead for more than 10 years on the date they supposedly ordered or authorized treatments (US Senate, 2008: 1-5).

In testimony before the Senate subcommittee, the OIG presented its analysis and recommendations on the dead doctors problem. The recommendations followed the same formulaic approach they developed for dead patients, deportees, and prisoners. Medicare should fix the processing system, they propose, so that up-to-date information about the status of each Unique Physician Identification Number (UPIN) is properly available to the claims-processing system, and Medicare's processing contractors can bounce back any claims that do not have a valid UPIN in the authorization field (Vito, 2008: 5-13).

While the OIG focuses on process improvement, the scandals all around the country are about fraud. The media provide a steady stream of stories about one petty crook, or group, who—without ever seeing a patient or providing any valid medical services at all—managed to bill Medicare or Medicaid, or some other health insurer, millions of dollars. We know from these cases that fake billing scams exist since they sometimes come to light. When claims are submitted, and they involve dead doctors or dead patients or some other feature that renders them obviously false, the most obvious explanation (if only someone would ask how they could have been generated) is that these claims arise as a by-product of fake billing scams. To understand why the authorities'

response to these billing issues is inadequate, even dangerous, one has to briefly contemplate what life looks like on the other side of the fence.

Let us imagine that these claims have actually been produced by Billy, the crook. Like so many others queuing up to attack the health-care industry's massive payment systems, Billy's goal is to steal as much as he can, as fast as possible. Billy pays a nominal fee to sign up as a Medicare provider himself, or infiltrates a billing service that submits claims on behalf of others. In order to bill Medicare, Billy does not need to see any patients. He only needs a computer, some billing software to help match diagnoses to procedures, and some lists. He buys on the blackmarket lists of Medicare or Medicaid patient IDs. If he wants to bill for services that require a prescription or authorization, he will also need to buy, steal, or otherwise obtain lists of physician numbers (UPINs) to enter on the electronic claims forms.

Billy is actually vulnerable because his lists are not entirely "clean." They contain just a few cases, probably no more than one in a hundred, of doctors or patients who are dead, deported, or incarcerated. The older the lists, the less clean they will be, as more of the patients will have had a chance to die or get deported or imprisoned, and more of the doctors will have retired, moved away, or died themselves. The impurities on Billy's lists are a problem for him, because they provide the authorities some chance to detect his false-claims scheme. The obvious implausibility of these claims, apparent if only the government had the right data in hand, provides an opportunity for the authorities to detect Billy's scam. Hence, unsure about the "cleanness" of his lists, Billy would pay a lot to know which patients' and doctors' numbers not to use, so as to avoid detection.

Now consider the standard government response to these various billing anomalies. In particular, what do the OIG's proposals mean for a fraud perpetrator like Billy? If the Medicare and Medicaid programs perfect their prepayment edits, and operate them as recommended, then Billy will receive computer-generated auto-rejection notices for the very small fraction of his claims that are obviously implausible. If he happened to use the identity of a dead patient, the

computer-generated notice he gets back from Medicare will politely inform him: "Medicare rejected this claim because, according to government records, this patient died prior to the date of service." The other 99 percent of Billy's claims, not involving any such detectable aberrances, will all be paid. From Billy's viewpoint, life is good. Government programs, with their emphasis on process-management, help him "scrub" his lists, making his fake billing scam more robust and less detectable over time. At the same time, the government pays all of his other claims without blinking an eye, and does not become the least bit suspicious.

In relation to the dead doctors problem, the OIG also recommends that the Medicare program, through its contractors, "educate providers" about the importance of using valid physician numbers on their claims. Dedicated fraud perpetrators like Billy will be diligent and grateful students. They are quite eager to perfect their billing practices, and—unlike the legitimate providers who are busy with their patients—they have plenty of time available to incorporate the government's feedback to their advantage.

Even the briefest of glances over this fence puts all of these categories of implausible claims in quite a different light. Rather than *processing errors to be corrected* these claims represent *detection opportunities for massive fake billing scams*. Once you see them in this light, an important question follows: Just how large might these billing scams be? For that, there is no empirical evidence. But one might imagine that the average list of Medicare providers (or patients), available to fraud perpetrators, would typically contain only a few instances of people who were in fact dead, retired, deported, or incarcerated. Suppose these accounted for 1 percent of the list, and that the fake billing scheme used the numbers on the lists evenly. Then one might surmise that the billing scams would likely be 100 times the size of the dead doctor or otherwise implausible claims that these scams would typically generate.

So, while congressional and public concern focuses on the several millions of dollars in obviously implausible claims that are apparently processed and paid in error, the real problem may well be *billions of dollars* in fake billing schemes. The obvious fictions represent impor-

tant detection opportunities; but they themselves are not the problem, but visible symptoms of it.

Insurers should by all means improve their capacity to detect such obviously implausible claims. Better interagency data exchange can facilitate this. But once such claims become visible, auto-rejection of the obviously bad claims is a feeble response. All assumptions of trust should be dropped immediately. A proper fraud response would do whatever was necessary to rip open and expose the business practices that produce such fictions. Relevant methods include surveillance, arrest, or dawn raids. Computers should be seized, and business practices examined. All other claims from the same source should be put on hold. Whenever a provider submits claims for treatment *of* the dead, or treatment *by* the dead, there is almost no chance that any of their other claims—submitted in the names of the living—are any more valid.

It seems extraordinary, given the long history of health care fraud in the United States, that even the Office of Inspector General, centrally placed to oversee the fight against fraud, displays such an obvious lack of comprehension when it comes to false claims and fake billings. Medicare officials and their overseers fail, like so many others across this industry, to properly distinguish between the imperatives of *process management* and the imperatives of *crime control.* By focusing so heavily on the first, they make life easier and safer for fraud perpetrators. One fundamental truth of the fraud-control business is this: *fraud works best when claims processing works perfectly.*

The health-care industry still acts as if it imagines that process accuracy is the cornerstone of effective fraud control. In fact, process accuracy (with the transparency and predictability it produces) is a large part of what makes health care payment systems such attractive targets for fraud.

ASSESSING FRAUD RISKS: TWO DIAGNOSTIC QUESTIONS

In order to assess the seriousness of different fraud threats—in terms of their potential to undermine the integrity of major public programs—two diagnostic questions turn out to be useful.

First, is the fraud *invisible by its nature*? Many frauds are not invisible. Credit card frauds, of the type where perpetrators usurp the existing accounts of others, are *visible*. Cardholders will generally notice unauthorized activity on their accounts, because they are being asked to pay, and so have an incentive to check. Most such frauds will be reported, and thus those responsible for controlling the problem at least know how much of it occurs. Of course, they may learn about credit card fraud too late to find the offenders or to prevent the loss, but the system overall sees the problem. As a matter of course, visible problems tend to get controlled, eventually. But *invisible* types of fraud can grow to a significant scale without anyone knowing how much damage is being done. Sophisticated fraud schemes are not only invisible at the time of commission, but remain invisible in perpetuity. Nobody ever knows they happened. Hence the underlying scale of the problem remains unknown.

Second, is there a *business opportunity* in the fraud? This question could be asked another way: Can a small number of dishonest players do a disproportionate amount of damage? Perhaps the patients can cheat too, to some extent. They might overstate their out-of-pocket expenses, or fabricate their own medical episodes while abroad on vacation. But any patient that begins to look too expensive, from the insurer's point of view, will draw scrutiny and be pulled back into line. So any one patient can only cheat so much on his or her own account, and hence the overall economic cost of patient fraud will be constrained by the proportion of dishonest patients.

Medical providers, routinely submitting bills in the names of hundreds or thousands of patients, can certainly ratchet up the volume. Other intermediaries, such as billing services, can spread their fraudulent activities across hundreds of provider accounts as well. Hence a few bad actors, suitably placed, can steal hundreds of millions of dollars. Judging by the nature of the cases that come to light, they often do.

The most dangerous fraud risks are the ones that combine these two qualities: they are both *invisible by nature*, and there is a business opportunity in the fraud itself. The health care industry in the United

States has constructed payment systems with a perfectly valid set of customer-service values in mind, assuming that the providers it is dealing with are delivering legitimate and necessary medical services, and can be trusted to tell the truth. The systems the industry has constructed, regrettably, turn out to be perfect targets for fraud, and criminal assault against them has run rampant. Unless authorities recognize the true nature of the fraud threat, and substantially increase their effectiveness in exposing and controlling it, there is a very real danger that fraud may end up destroying the integrity and viability of some vitally important public programs.

REFERENCES

Centers for Medicare and Medicaid Services (CMS). Department of Health and Human Services. "National Health Expenditure Data." Washington D.C., 2006 <http://www.cms.hhs.gov/NationalHealthExpendData/downloads/highlights.pdf>.

———. "National Health Care Expenditure Projections." Washington D.C., 2007. http://www.cms.hhs.gov/NationalHealthExpendData/Downloads/proj2007.pdf

Centers for Medicare and Medicaid Services (CMS). Center for Medicaid and State Operations. "Payment Accuracy Measurement Project: Year 2 Final Report." Washington D.C., April 2004.

Department of Justice. "Largest Health Care Fraud Case in U.S. History Settled: HCA Investigation Nets record Total of $1.7 Billion." News Release, Washington D.C., June 26, 2003 <http://www.usdoj.gov/opa/pr/2003/June/03_civ_386.htm>.

Freeh, Louis J., Statement of FBI Director before the Special Committee on Aging. U.S. Senate, Washington D.C., March 21, 1995.

Government Accountability Office. "Efforts to Measure Medicare Fraud." Letter to House Budget Committee Chairman, Rep. John R. Kasich. GAO/AIMD-00-69R. Washington, D.C., February 4, 2000.

Office of Inspector General (OIG), Department of Health and Human Services. "Medicare Payments for Services after Date of Death." Report OEI-03-99-00200. Washington D.C., March 2000.

———. "Review of Medicare Payments Made on Behalf of Deported Beneficiaries." Report A-04-01-05004. Washington D.C., March, 2002(a).

———. "Review of Medicaid Payments for Outpatient Services and Prescription Drugs Provided to Incarcerated Recipients in the State of Florida." Report A-04-01-05011. Washington D.C., October 2002(b).

———. "Review of Medicare Payments for Services Provided to Incarcerated Beneficiaries." Report A-07-02-03008. Washington D.C., October 2002(c).

———. "Audit of Selected States' Medicaid Payments for Services Claimed to Have Been Provided to Deceased Beneficiaries." Report A-05-05-00030. Washington D.C., September 2006.

Raab, Selwyn. "Officials Say Mob Is Shifting Crimes to New Industries." *New York Times,* February 1, 1997.

Sparrow, Malcolm K. *License to Steal: How Fraud Bleeds America's Health Care System.* Denver, Colo.: Westview Press, 2000.

———. *The Character of Harms: Operational Challenges in Control.* Cambridge: Cambridge University Press, 2008.

Transparency International. *Global Corruption Report 2006: Special Focus—Corruption and Health.* Berlin, 2006.

U.S. Senate, Permanent Subcommittee on Investigations, Committee on Homeland Security and Governmental Affairs. "Medicare Vulnerabilities: Payments for Claims Tied to Deceased Doctors." Committee Staff Report, Washington D.C., July 9, 2008.

Vito, Robert. "Medicare Payments for Claims with Identification Numbers of Dead Doctors." Testimony before the Permanent Subcommittee on Investigations, Committee on Homeland Security & Governmental Affairs, U.S. Senate. Washington D.C., July 9, 2008.

Weaver, Jay. "Couple Pleads Guilty to Medicare Fraud," *Miami Herald*, July 3, 2008.