



ELSEVIER

Available online at www.sciencedirect.com

SCIENCE @ DIRECT®

Journal of Health Economics 23 (2004) 935–949

JOURNAL OF
HEALTH
ECONOMICS

www.elsevier.com/locate/econbase

To err on humans is not benign Incentives for adoption of medical error-reporting systems

Joshua Graff Zivin^{a,*}, Alexander S.P. Pfaff^b

^a Department of Health Policy and Management, Mailman School of Public Health and International Center for Health Outcomes and Innovation Research, Columbia University, Room 608, 600 West 168th Street, New York, NY 10032, USA

^b Department of Economics, School of International and Public Affairs, Columbia University, New York, NY 10032, USA

Received 31 January 2002; accepted 30 January 2004

Available online 28 July 2004

Abstract

Concerns about frequent and harmful medical errors have led policy makers to advocate the creation of a system for medical error reporting. Health providers, fearing that reported information about errors would be used against them under the current medical malpractice system, have been reluctant to participate in such reporting systems. We propose a re-design of the malpractice system – one in which penalties are a function of the health provider's reporting efforts – to overcome this incentive problem. We also consider some alternatives to this mechanism that address two important ways in which reporting effort may not be observable: hospitals may have interests distinct from individual physicians and may not be able to observe their reporting efforts, and a regulatory agency or a court may not be able to adequately observe reporting efforts by a provider.

© 2004 Elsevier B.V. All rights reserved.

JEL classification: I1; L5

Keywords: Medical errors; Malpractice; Event reporting; Liability

1. Introduction

In the early 1990s, a Harvard Medical Practices Study provided staggering figures concerning the incidence of adverse events within US hospitals (see Brennan et al., 1991; Leape et al., 1991). Perhaps even more shocking than the prevalence of these adverse outcomes was that more than one-quarter of these events were due to negligence. In response, the

* Corresponding author.

E-mail address: jz126@columbia.edu (J.G. Zivin).

Institute of Medicine (IOM) initiated the Quality of Healthcare in America project, which served to underscore the frequency with which preventable injuries occur within the US medical care system (Kohn et al., 1999). Not surprisingly, these findings were widely publicized (Pear, 1999; David and Appleby, 1999). Soon thereafter, as evidenced by the *Medical Error Reduction Act (2000)* and the *Patient Safety and Errors Reduction Act (2000)*, policy makers began to clamor for some sort of solution.

Preventing errors requires the identification of adverse events and their causes so that the appropriate corrective measures can be undertaken. A number of steps could be taken, in the public and private realms, to facilitate such identification. One would be to generate a new knowledge base. The IOM report advocated confidential and voluntary reporting and record keeping systems for injurious events, alongside mandated reporting for more serious events. The Joint Commission on Accreditation of Healthcare Organizations (JCAHO, 2004) has proposed a Sentinel Events Policy to encourage the self-reporting of “an unexpected occurrence involving death or severe physical or psychological injury, or the risk thereof”. Under this proposed policy, health providers would be required to analyze the root causes of detected errors and to create an action plan to address the causes of all ‘sentinel events’. Error prevention activities were deemed so important that failure by a healthcare organization to carry them out within 45 days would jeopardize the organization’s accreditation status.

What could be wrong with a plan to generate such information? *Incentives for self-reporting*. An important feature of error-reporting systems is that the providers are central to the reporting, as is quite natural given their presence when errors occur and their expertise in discerning errors. Yet questions abound concerning health providers’ participation in such reporting systems. When JCAHO announced the Sentinel Events Policy, the American Hospital Association sent notices to its members urging them not to comply, citing concerns about increased liability exposure. The American Society for Healthcare Risk Management was also quite critical of the suggested policy, scorning it as a “boon to the legal profession” (Liang, 2000a,b).

These anecdotes highlight the need to evaluate the incentives faced by a healthcare provider to participate in any system relying upon self-reporting. As our model will demonstrate, given their liability under the current medical malpractice system to ask the healthcare providers to be the reporters of such adverse events is somewhat akin to asking a Tax Evaders Club Of America to provide the IRS with its membership list. If reported information is discoverable – as there is reason to believe (Liang, 1999) – then providers may have strong incentives not to report.¹

The disincentive stems from the widely held belief that self-reporting raises the chance that an error will be detected by a regulator or court, i.e. that there is a ‘detection differential’.² If providers with reporting systems are caught more often, they will face higher expected penalties than those without such a system. This increased chance of incurring liabilities also makes the benefits from reporting lower for a healthcare provider than for society. Any

¹ This is a widely held view. See, for example, Leape et al. (1998): “Currently, the disincentives to disclosing errors seem much stronger than the incentives”. Also see, among others, Studdert and Brennan (2001): “. . . the patient safety reforms spurred by the IOM report are on a collision course with the medical malpractice system”.

² From Liang, 2000b: “Of paramount concern for hospitals and providers, and their primary grievance with regard to the Sentinel Event Policy, is the potential for subsequent discoverability of the information that JCAHO requires”.

modification of the current malpractice system concerning error reporting must address these incentive effects. We suggest a modification, based on the detection differential, to efficiently correct reporting incentives.

1.1. Current landscape of solutions to reporting disincentives

This tension between the social gains from and the private incentives for reporting has received considerable attention in academic and policy communities. The IOM report devotes an entire chapter to the protection of reporting systems from legal discovery, arguing that this would remove a major obstacle to reporting, as providers would no longer need to weigh gains in safety against potential punishment and humiliation.³ If legal discovery of reports were prevented, then reporting would not affect the chance an error is detected. As a result, health providers' incentives would no longer be distorted relative to a societal view.

Preventing legal discovery, however, would make it more difficult for patients to receive compensation when it is warranted. Worse, it could create safe havens for incriminating evidence, substantially weakening firms' incentives to prevent errors.

An alternative policy approach, which does not limit legal discovery, is the implementation of a "no-fault" compensation system (see Bovbjerg and Sloan, 1998; Studdert and Brennan, 2001) funded by premiums unrelated to errors. Here, reporting could not raise expected penalties, as no penalties are levied upon providers. Unfortunately, this alternative may also provide little incentive for firms to prevent errors. "Experience rating", where providers with higher rates of these errors pay higher premiums, could enhance these prevention incentives. Yet this modification would recreate, using premiums instead of penalties, the disincentive to reporting that exists under the current malpractice system.

1.2. Mechanism re-design: efficient incentives for both reporting and prevention

We propose explicitly rewarding providers for generating socially valuable information. Specifically, we maintain the liability in the current malpractice system but add an incentive for reporting by lowering penalties for the same error when reporting efforts are greater. All errors are punished but the magnitude of punishment falls as the information generated rises.⁴

This modified penalty structure counters the disincentive to reporting arising from the detection differential, by adding a subsidy for reporting to the harm-based fines within the current malpractice regime. A provider need not fear reporting if it makes error detection more frequent but also correspondingly less painful. Our proposed modification completely removes reporting disincentives by equating expected penalties for providers with and without reporting systems.

For example, suppose a hospital pharmacy dispatches the wrong dose of the correct drug to a nurse, who fails to check the dose and administers it, causing harm. With a reporting system, a case report and analysis would follow, potentially yielding changes in the care

³ See also, for example, Leape et al. (1998) and Brennan (2000).

⁴ The modeling approach taken here is similar in spirit to work on optimal law enforcement (Kaplow and Shavell, 1994; Innes, 1999), takings (Polasky and Doremus, 1998), and environmental auditing (Pfaff and Sanchirico, 2000).

provided to the current patient and in various processes of care provision. Such changes would not be possible without a well functioning reporting system. But these reports and analyses will make fines and lawsuits more likely than for a non-reporting provider, since the regulators and private parties are more likely to learn about the error and to have materials for prosecution.⁵ To counteract this disincentive, penalties for a detected error should be lower for a reporting provider guilty of a given error, i.e. a modified policy could essentially combine a subsidy to reporting with the tax upon damages. If done correctly, this amended policy could make a firm's expected penalties invariant to the reporting choice.

The remainder of the paper is structured as follows. [Section 2](#) models providers' decisions to report and to correct errors. We construct a hypothetical case with no liability, in which even the correction of detected errors is socially inefficient. This helps to provide intuition for the effects of the current malpractice regime, which leads to efficient correction once a problem is detected but socially inefficient reporting. [Section 3](#) presents our modification to the current malpractice regime – a two-tiered penalty system with penalties conditioned on reporting effort. This penalty structure achieves efficient reporting and correction not only in the case of a unitary provider but also when individual physicians must be induced to report events. [Section 4](#) concludes with a brief discussion of our results, some policy alternatives, and several possible model extensions.

2. Provider reporting and correction decisions

2.1. Base case without malpractice: inefficient reporting and inefficient correction

During the course of the provision of healthcare services, errors occur that result in harm to patients.⁶ Examples include the incorrect administration of a drug, or drug dosage, the wrong device implanted during surgery, or the mismatch of blood type. The harm associated with an error can include increased medical expenses, physical harm, and even death. If its underlying cause is not corrected, any error is likely to recur, causing similar future harm.

Since prevention of harm is the goal driving a focus on reporting, it is worth emphasizing exactly what many experts feel is required to avoid harm. Simply observing that an error has occurred (e.g. a patient received the wrong medicine) is often far from enough to prevent the same thing from happening in the future. [Vincent \(2003\)](#) stresses that error prevention requires understanding how a mistake arose, and thus the reporting and recording of detected errors. 'Root-cause analysis' (in JCAHO's language) or 'systems analysis focused on the possibilities to change provision going forward' (in [Vincent's \(2003\)](#) language) is often conducted by people other than those who detected an error. Reporting and recording are necessary to facilitate any form of enterprise-wide communication that could allow analysis and changes in healthcare provision.

⁵ From [Liang, 2000a](#): "... by reporting a sentinel event and delivering a root cause analysis to the JCAHO, an entity may be compiling, and may need to deliver to the opposition without request, very damaging materials. . . ."

⁶ While our modeling approach could in principle (at least with some modification) be applied to the decisions made by any type of provider of healthcare services, the enterprises we are most directly thinking about here are hospitals.

Following reports such as that by the IOM, there is a common belief about the probability distribution of the magnitudes of errors correctable and preventable with a reporting system, e , specifically the probability density function $f(e)$ and associated cumulative density function $F(e)$. The variable e is an index of the aggregate severity of harm resulting from the present error and its probability-adjusted recurrences that could be avoided through correction and prevention efforts. The decision to invest in reporting can be based on this distribution and in particular on its translation into both private and social harm. Let s denote the per-unit societal damages from an error. This applies to not only current but also future damages associated with an observed error (e.g. a patient death from a correctable process flaw). Thus, knowing the total magnitude of the correctable and preventable societal damages associated with that error requires investigation of the error's underlying causes, which in turn requires the adoption of a reporting system.

Consider a provider deciding whether or not to adopt an error-reporting system at some cost. For simplicity, assume this is a binary decision, i.e. this provider adopts the system or does not. Assume the system functions perfectly to detect, report, record and analyze all observed errors and reveal correctable and preventable error magnitudes e . Denote the cost of adopting the reporting system as k . If a system is adopted and an error is observed, a provider must then decide whether to incur the costs of the steps identified by the system to correct and prevent the associated errors and their societal damages. This package of steps both helps the current patient and prevents a recurrence of similar errors, which, for simplicity, we will generally refer to as error correction. The choice is binary, has cost c and when undertaken eliminates associated e .

Thus, addressing errors requires two steps: first, adopt a reporting system at cost k ; and second, undertake correction and prevention at cost c .⁷ In the absence of malpractice liability, these decisions will depend on the benefits to the providers themselves. Healthcare providers prefer not to harm their patients (at the current time or in expectation), so we assume that for each unit of preventable error e the provider receives x units of disutility, where x is strictly lower than s . Thus, on their own, providers correct errors when $c \leq x$. Since c is assumed to be the same for all e , providers will prevent/correct those errors that cause them the most disutility.

As a result, the provider's decision about reporting can be characterized as choosing

$$\operatorname{argmax} \left[\pi_0 - \int xf(e) de - \int_{c \leq x} cf(e) de - \int_{c > x} xf(e) de \right] \quad (1)$$

where π_0 denotes the healthcare provider's short-run profit (or quasi-rent). Without reporting, i.e. in the term to the left, net returns are simply the quasi-rent minus the disutility from the expected patient harm (in dollar terms). If a provider adopts a reporting system, learning the e associated with the provider's observed errors and identifying how to correct and prevent that harm, then the net return starts with the quasi-rent but now we subtract k , the cost of the reporting system, plus the cost of correcting and preventing errors for which such remedies are worthwhile (i.e., for which $c \leq x$). The residual disutility x is from the

⁷ In principle, some future errors could be corrected without analysis of errors. In the next section, we see that incentives for their correction can be efficient without modification to the current malpractice liability regime.

errors deemed not worthy of these steps. It is analogous to the disutility within the left-hand term but it applies to a smaller set of errors.

That only the integrals in the term to the right within Eq. (1) reflect the choice of whether or not to correct future errors indicates precisely the value of the reporting system. Reporting provides information about e associated with an observed error and thus an option to correct and prevent errors and, in turn, harm. Thus, while reporting has costs, it also provides benefits by allowing providers to correct and prevent errors for which disutility is greater than the cost of taking the needed steps to address the problem. Manipulating Eq. (1), we see that in the absence of malpractice providers will adopt reporting if

$$k \leq \int_{c \leq x} (x - c) f(e) de \quad (2)$$

Recall, however, that society places a per-unit value of s on preventable errors e . This s includes all disutility, the provider's as well as that of current and future patients. Thus, society will want to correct all preventable errors for which $c \leq s$ and will want providers to adopt reporting if

$$k \leq \int_{c \leq s} (s - c) f(e) de \quad (3)$$

As can be seen from Eqs. (2) and (3), a role for regulators and the medical malpractice system arises because providers' disutility from a unit of correctable and preventable error is less than the associated societal damages (i.e., $x < s$). If this were not the case, policies to encourage the correction and prevention of detected errors and the adoption of error-reporting systems would not be needed.

2.2. Current malpractice system: efficient correction given detection, inefficient detection

Given this divergence in decision rules, regulators and the medical malpractice system must levy financial penalties, for suffering, that increase providers' net benefits from error prevention. These penalties should be chosen so that the provider's net benefits from preventing errors will correspond exactly to those of society, in essence implementing a Pigouvian tax on damages so that providers internalize the external benefits of error correction and prevention. In this case the marginal incentives are aligned and private decisions are socially efficient. We assume that not all errors are detected by regulators. Consider first a fixed probability of detection p . In this case, the optimal penalty should be the social-private gap in disutilities from errors, $s - x$, divided by the chance of detection (Becker, 1968). We view this penalty structure – a uniform fine based upon the magnitude of the error – as representative of the current medical malpractice system. The penalty includes all of the financial costs to the provider resulting from the medical liability system, e.g. fines levied by agencies and lawsuits by private parties.

For a known error, we can determine the level of assessed penalty a to align the incentives. The expected penalty for a detected error must equal the extra societal gain from preventing an error, i.e. $pa = s - x$. The cost of correction efforts, c , will then be compared to the expected disutility x plus the expected penalty that can be avoided, i.e. to

$x + pa = x + (s - x) = s$.⁸ The penalty $a = (s - x)/p$ provides efficient incentives for correction. This approach works equally well if we introduce a reputation cost r for providers, distinct from their ‘internal’ disutility from errors, which accrues only for the detected harm. In this case, the penalty a satisfying $pa = s - x - pr$, such that $a = ((s - x)/p) - r$, efficiently bridges the private-social gap in the benefits of correction. Thus, in a framework in which the probability of an error being detected is invariant to firms’ choices, the current medical malpractice system can provide efficient incentives for error correction.

The problem we examine, however, is more complicated. As suggested in Section 1, the main problem with reporting in the eyes of the providers is that the detection probability is not fixed, at p , but rather is higher for providers with reporting systems. This seemingly minor modification to the formal regulatory problem – the existence of a detection differential – in fact lies at the heart of the disincentive to reporting. It must also, then, be featured in the solutions.

The current malpractice liability regime ignores this differential and thus is not able to provide efficient incentives for both correction and reporting. This feature is incorporated into our model by adding a probability of error detection given a reporting system, $P > p$, with p representing a lower probability of detection for an error by a provider without a system. The increased probability of detection due to reporting is assumed to be the same for all types of errors, i.e. the detection differential is independent of the magnitude of e . Efficient correction if a provider detects an error using a reporting system would require that the penalty a satisfies $Pa = s - x - Pr$. Cost c will then be compared to the expected disutility plus the expected penalty that can be avoided, i.e. to $x + P(r + a) = x + (s - x) = s$. As this comparison matches the societal cost-benefit tradeoff (see just above (3)), correction is efficient.

Consider then the provider’s decision to adopt a reporting system. The provider will choose

$$\operatorname{argmax} \left[\begin{array}{l} \pi_0 - \int xf(e) de - \int p(r + a)f(e) de \\ \pi_0 - k - \int_{c \leq x + P(r+a)} cf(e) de - \int_{c > x + P(r+a)} (P(r + a) + x)f(e) de \end{array} \right] \quad (4)$$

Given that the penalty a has been set for errors detected through reporting systems, i.e. that the penalty is set in order to satisfy $P(r + a) = s - x$, the reporting system will be adopted only if

$$k + \int_{c > s} (P - p)(r + a)f(e) de \leq \int_{c \leq s} (p(r + a) + x - c)f(e) de \quad (5)$$

The left-hand-side of this expression describes provider costs of adopting a reporting system. These include not only the direct costs k but also a cost, in penalties and reputational damages, due to uncorrected errors being detected with greater frequency when a reporting system exists. The right-hand-side of the expression describes the net benefits, in

⁸ Note that since the cost of corrections are assumed to be the same for all e , more costly errors (in terms of fines, reputation costs, and disutility) are more likely to be reported and remedied.

terms of avoided penalties, disutility and reputational damage, from error correction possible due to the reporting system. Comparing (5) with (3) reveals that neither the private costs of reporting nor its private benefits match the expression for socially efficient reporting incentives. The difference between the private and social costs of system adoption is $\int_{c>s} (P - p)(r + a) f(e) de > 0$ and the difference between the private and social benefits of adoption is $\int_{c\leq s} (p - P)(r + a) f(e) de < 0$. Given the detection differential arising from reporting, the costs are higher and the benefits are lower from the provider’s perspective. This leads to a private underinvestment in reporting.

In sum, the existing medical malpractice system uses a uniform penalty mechanism that can induce efficient correction conditional on detection. However, this mechanism ignores the differential in the detection probability due to reporting, and thus cannot simultaneously provide efficient incentives for reporting system adoption and efficient correction. Providing efficient incentives to correct those errors for which corrective steps have been identified (through use of a reporting system) will lead to an underinvestment in the process of error identification itself.

3. Aligning reporting incentives through mechanism re-design

Providing efficient incentives to both correct errors and adopt reporting requires a departure from policies with a uniform penalty, which cannot counteract the increase in firms’ expected financial penalties that results from the detection differential created if the provider has adopted reporting. We propose a system in which, for a given error, the financial penalties themselves are lower for providers with reporting systems. This counteracts the increased probability of being penalized, removing the disincentive to adopting a reporting system that is apparent within expression (5).

3.1. Using reporting behavior to create incentives for reporting

Recall the earlier example in which a hospital had administered an incorrect dosage of a drug, causing harm. The adoption of a reporting system by this provider will make fines and lawsuits for this error more likely. To counteract this disincentive, the penalties for a detected error should be lower. We propose distinguishing a , the penalty for providers who have a reporting system, from A , a larger penalty for providers who do not. For expected penalties to be insensitive to the presence of a reporting system, we can simply set $p(r + A) = P(r + a) = s - x$ (where $A > a$, because $p < P$).

Considering the healthcare provider’s decision to adopt a reporting system, now facing this modification to the current medical malpractice system, the care provider will choose

$$\operatorname{argmax} \left[\begin{array}{l} \pi_0 - \int xf(e) de - \int p(r + A) f(e) de \\ \pi_0 - k - \int_{c\leq x+P(r+a)} cf(e) de - \int_{c>x+P(r+a)} (P(r + a) + x) f(e) de \end{array} \right] \quad (6)$$

The only difference between expression (6) and expression (4) is the final part of the upper term within the returns for the provider if not adopting a reporting system. In this case,

providers face expected penalties and reputational damage of $p(r + A)$ rather than $p(r + a)$ for each error. Given a larger penalty for non-reporting providers, the reporting system is adopted if and only if

$$\begin{aligned} k + \int_{c > x + P(r+a)} (P(r+a) - p(r+A)) f(e) de \\ \leq \int_{c \leq x + P(r+a)} (p(r+A) + x - c) f(e) de \end{aligned} \quad (7)$$

Since penalties a and A were chosen such that $p(r + A) = P(r + a) = s - x$, we can re-express the provider's reporting decision as choosing to adopt the error-reporting system if and only if

$$k \leq \int_{c \leq s} (s - c) f(e) de \quad (8)$$

This is identical to the social decision rule in (3). The disincentive to reporting no longer exists.

3.2. Implications of a principal-agent model

Thus far, the healthcare provider has been treated as a single entity responsible for system adoption plus detecting, reporting, recording, analyzing and correcting. Here we distinguish the hospital from the physician, allowing them to have different objectives and roles in reporting systems (as emphasized in proposals for 'enterprise liability' in which enterprises such as hospitals, not individual physicians, are the penalizable parties (see Abraham and Weiler, 1994; Sage et al., 1994). We derive the optimal regulatory solution in light of the need for hospitals to induce reporting by physicians, given distinct objectives. We then compare this solution to that presented just above.

Hospitals function essentially like the unitary provider modeled above, taking adoption decisions concerning reporting systems and taking correction decisions when errors are detected. Now, however, the efficacy of hospitals' corrective actions is a stochastic function of physicians' efforts to detect errors and report them into a system, and physician detection/reporting effort is unobservable. Without useful input from physicians, it is difficult for hospitals to accurately identify an error's cause. Thus, some actions hospitals take to prevent future errors will not succeed. Because efficacy is a stochastic function, a failed correction does not allow the hospital to identify physician-reporting effort and the hospitals will not be able to contract on effort. Instead, contracts are written over outcomes.

Formally, the hospital makes the reporting decision as above, except that now corrective efforts (which still cost c no matter their efficacy) eliminate the potential future errors e only with a probability $q < 1$. This probability depends upon unobservable physician detection and reporting effort \equiv .

Following Eqs. (4) and (6), the hospital's reporting system decision can be characterized as choosing

$$\operatorname{argmax} \left[\begin{array}{l} \pi_0 - \int xf(e) \, de - \int p(r + A) f(e) \, de \\ \pi_0 - \hat{k} - q(\mathbf{B}) \int_{c \leq \bar{q}[x+P(r+a)]} cf(e) \, de \\ - (1 - q(\mathbf{B})) \int_{c \leq \bar{q}[x+P(r+a)]} (c + P(r + a) + x) f(e) \, de \\ - \int_{c > \bar{q}[x+P(r+a)]} (P(r + a) + x) f(e) \, de - T(q) \end{array} \right] \tag{9}$$

where \hat{k} is the fixed cost of adopting the system (distinguished from the earlier costs k for the purposes of comparing models later) and $T(q)$ is a transfer payment from hospital to doctor that depends on the realization of q , i.e. on whether correction decisions successfully eliminate e . Unlike prevention/correction costs, $T(q)$ is not a cost incurred only for those errors corrected. It is a schedule of transfer payments conditional on the realization of q that is determined independently of individual correction decisions. As such, it is best viewed as an additional fixed cost of adopting the reporting system. As before, the hospital undertakes correction efforts when their costs are less than the expected costs faced from not correcting them. The expression for the latter has changed, however, since correction is only sometimes effective. The expected benefits of correction are now equal to the avoided penalties, reputation damages, and disutility, adjusted for the expected effectiveness of correction efforts ($\bar{q} < 1$).⁹ These efforts still cost c and, if they are not successful, with probability $(1 - q)$, the hospital still faces the expected penalties for the errors.

Turning to physicians, in our simplified model physician profits are expressed as follows:

$$\pi_{\text{doc}} = \tilde{\pi} - \gamma(B) + T(q) \tag{10}$$

where $\tilde{\pi}$ are the rents earned by physicians that are independent of reporting efforts, $T(q)$ is the transfer as noted earlier, and γ are costs incurred as a result of error detection and reporting effort B . The γ encapsulates the direct costs of detecting and reporting as well as the possibility that these efforts will lead to reputational damages and medical malpractice expenses for the physician.

Given these expressions characterizing the hospital adoption decision and physician profits, we can address the transfer $T(q)$. The optimal transfer in this standard principal-agent framework is well known. The efficient contract between principal and agent will permit agents to receive the full values of their actions on the margin minus a fixed rent (see Grossman and Hart, 1983). Letting B^* denote optimal physician effort in response to the optimal contract, the optimal transfer payment is $\gamma(B^*)$. In expectation, physicians are completely compensated for their efforts and incentives are aligned (see Appendix A for derivation).

The hospital will adopt the system if and only if

$$\begin{aligned} & \hat{k} + \int [P(r + a) - p(r + A)] f(e) \, de \\ & \leq \int_{c \leq \bar{q}[x+P(r+a)]} (\bar{q}[P(r + a) + x] - c) f(e) \, de - \gamma(B^*) \end{aligned} \tag{11}$$

⁹ If the same stochastic correction function were added to the model with no agency problem, the difference between the two would simply reduce to the fixed costs of inducing physicians to detect/report. The qualitative results obtained here remain unchanged.

Setting fines such that $\bar{q}[P(r + a)] = \bar{q}[p(r + A)] = s$, i.e. implementing our two-penalty policy so that high fines for non-adopters offset a lower chance of being fined, the decision is expressed as

$$\hat{k} + \gamma(B^*) \leq \int_{c \leq s} (s - c) f(e) de \quad (12)$$

Comparing Eq. (12) to the societal decision rule for adopting a reporting system within Eq. (3) reveals one small difference. Instead of the reporting system costs for the unified healthcare provider k on the left-hand side, we have the direct costs to the hospital \hat{k} plus the amount of money γ that must be transferred to doctors to compensate them for their detection and reporting efforts B .

Note that if corrective efforts are always effective (i.e., $\bar{q} = 1$) the fines are identical to those in the earlier model without an agency problem. It also remains the case that, regardless of the expected effectiveness of correction efforts, efficient regulation requires the equating of expected penalties with and without reporting, i.e. $p(r + A) = P(r + a)$. However, each of these fines is now larger, since $p(r + A) = P(r + a) = (s - x)/\bar{q}$. The larger a is required to induce hospitals to exert correction effort, despite its fallibility, and this results in a larger A to retain reporting incentives given the detection differential.

In essence, the agency problem translates into an additional fixed cost of adopting reporting for the hospital. When the fixed or direct cost of the reporting system plus transfer to physicians (i.e., the left-hand side of Eq. (12)) is greater than the direct costs (k) of adoption when the hospital and doctor behave as a unitary actor, then the adoption rule above suggests that system adoption will be less likely in the principal-agent case, *ceteris paribus*. Thus, some hospitals that would prefer to have a reporting system when agents' incentives are aligned may opt not to adopt due to the costs of having to induce reporting. Nonetheless, if society cares about the costs of physician efforts in reporting γ , then our re-designed mechanism as characterized in Eq. (12) remains socially optimal.

4. Discussion

Suffering as a result of medical errors is common and often preventable. Thus, many advocate the use of error-reporting systems to help track and learn from mistakes. Yet other significant players argue strongly against such systems. This might seem surprising, but it is quite easily understood if one considers the disincentives created by the regulatory and legal environments that health providers face. Policies for overcoming these disincentives to reporting have not been adopted in the US, in part as they may come at patients' expense (e.g. making compensation difficult to obtain) or because they may fail to retain incentives for correcting errors. Political and interest group resistance make them even more difficult to enact (see, for example, Sage, 2001).

In this paper, we presented a policy that gives health providers an incentive to both report and correct errors, while still allowing patients to be compensated for harm they have incurred. This policy relies upon financial penalties that, for a given error, depend on

whether the provider exerts reporting effort. Penalties for the providers who report are lower, to counter the fact that a provider who reports is more likely to be caught and penalized. By making expected penalties invariant to a firm's reporting choices, the disincentive to reporting is eliminated. This basic approach, adding a subsidy for reporting to existing damage-based fines in the malpractice system, also functions when hospitals face an agency problem, i.e. when they cannot observe individual physicians' reporting efforts and must induce efficient reporting by the physicians.

4.1. Variations upon this theme

While our mechanism redesign for the regulator achieves efficient correction and reporting, it may be challenging to implement since it relies upon information about reporting efforts that may be difficult to ascertain. Consider this regulation: "if reporting efforts are acceptable, then penalties for errors will be lowered". The existence of a reporting system, which may be easy to document, may not imply "acceptability" or acceptable intensity of reporting effort. A provider could set up a reporting system but operate it in a way that would not detect and correct errors. Thus, a continuous measure of the intensity of reporting effort may be required.¹⁰ When judges, juries, and regulators do not possess such a measure, two variations using proxies may help.

First, the manner in which an error was detected can be an observable proxy for reporting effort. If penalties are lower for errors detected in ways that are more likely if reporting effort has occurred (e.g. via case records or whistleblowers), reporting would again decrease the chance of higher penalties, offsetting the increase in the overall rate of error punishment as a result of reporting. A second variation would use as observable proxies detected errors that were not corrected even though, once root-cause analysis is done, they pass an ex-post cost-benefit test for correction. Put another way, these are errors for which the care provider must not have looked, as if found they would have been corrected. Raising penalties for this subset of the errors provides an incentive to adopt a reporting system so that errors of this type can be detected and prevented. Again, reporting lowers the frequency of higher penalties to offset the higher overall frequency with which errors are punished. In both of these variations, a signal is used to construct a penalty *indirectly* conditioned upon reporting.

4.2. Implementation and extensions

Any implementation of our recommendations hinges on the ability to estimate the difference between the probabilities that the regulatory community will detect an error with and without a reporting system, i.e. how much the private reporting system helps to further the social agenda. It should be noted that judges and juries are often asked to engage in

¹⁰ One possible continuous measure of reporting-system effort, e.g., is that some providers may simply record errors, while other providers might have a system in place to analyze the errors for their underlying causes. Alternatively, detection efforts could be measurably different across systems. Some might adopt, e.g., a new "surgical black box" (analogous to flight-data recorders on planes) that provides the ability to review procedures later (see MIT, 2001).

such exercises. Any number of judgments rest on whether an action “materially helped or hindered” a specific pursuit. Moreover the regulatory community need not play a passive role in this estimation. It could describe what it sees as the important observable elements of reporting systems and then implicitly ascribe levels of helpfulness to those elements through a schedule of penalties associated with those indicators of reporting intensity. Such a schedule is currently employed in an environmental context, where the Environmental Protection Agency offers varying levels of penalty reduction depending on evidence of auditing effort (EPA, 1995).

Our model might be extended in several directions. To start, physicians have quite complicated relationships with many actors, suggesting a multi-dimensional role for providers’ reputations in shaping behavioral incentives. Patients may view assessed penalties as indicating low provider quality, and it is hard to know how these consumers would compare frequent smaller penalties, as advocated to create reporting incentives, to infrequent higher penalties. Such reputational complications could require additional policy interventions to educate patients, not unlike the public discussions of risk adjustment that took place in the wake of physician and hospital report cards (Iezzoni, 1997). Further, providers’ reputations among their colleagues are prized and may be damaged by the reports of errors that would emanate from a reporting system. Similar arguments could be made with respect to third-party payers. On the other hand, having an observably functional error-reporting system could enhance a provider’s reputation with all others by signaling a commitment to the correction of inevitable problems of process and thus to quality of care and patient well-being.

Given stochastic error production and error detection, the model could also be extended to incorporate both patient and provider risk aversion. Lastly, incorporating explicitly the roles of the insurance industry and other healthcare stakeholders into modeling private and social choice may help to identify new error-management strategies. These comprise a future research agenda.

Acknowledgements

For helpful discussions, we would like to thank Hal Kaplan, Kim Putalik, Dahlia Remler, William Sage, and in particular Chris Sanchirico, upon whose work with Pfaff our approach is based. Also, Graff Zivin gratefully acknowledges financial support from the Agency for Health Research and Quality, Grant 1U18HS11905-01.

Appendix A. The optimal contract in the principal-agent framework

For simplicity, assume that corrective efficacy takes on one of two values q_1 or q_2 , where $q_2 > q_1$. Let $\phi(B)$, denoting the probability of the better outcome q_2 , increase with physicians’ reporting effort, i.e. $\phi'(B) > 0$. Also let Θ denote the fixed payment from physician to hospital and let t_i denote the transfer payment for realization q_i . The transfer payments can then be expressed as

$$t_1 = q_1 \int_{c \leq \bar{q}[x+P(r+a)]} cf(e) de + (1 - q_1) \int_{c \leq \bar{q}[x+P(r+a)]} (c + P(r + a) + x) f(e) de - \Theta$$

$$t_2 = q_2 \int_{c \leq \bar{q}[x+P(r+a)]} cf(e) de + (1 - q_2) \int_{c \leq \bar{q}[x+P(r+a)]} (c + P(r + a) + x) f(e) de - \Theta$$

Given this contract, physicians will exert effort $B^*(\Theta)$ (assuming that physicians’ participation constraints are satisfied by the transfers). The optimal contract will be one where the fixed rental payment from physician to hospital makes the physician no worse off than without a contract, i.e. one that just satisfies a physician-participation constraint.¹¹ Thus, the fixed payment will be equal to the expected costs of correction minus the effort costs to the physician:

$$\Theta = \bar{q} \int_{c \leq \bar{q}[x+P(r+a)]} cf(e) de + (1 - \bar{q}) \int_{c \leq \bar{q}[x+P(r+a)]} (c + P(r + a) + x) f(e) de - \gamma(B^*)$$

Substituting the expressions for the optimal contract into the hospital’s reporting decision (9) and simplifying, we obtain an expression for the hospital’s adoption of an error-reporting system within a principal-agent framework. The hospital will adopt the system if and only if

$$\hat{k} + \int [P(r + a) - p(r + A)] f(e) de \leq \int_{c \leq \bar{q}[x+P(r+a)]} (\bar{q}[P(r + a) + x] - c) f(e) de - \gamma(B^*)$$

References

Abraham, K., Weiler, P., 1994. Enterprise medical liability and the evolution of the American healthcare system. *Harvard Law Review* 108, 381.
 Becker, G.S., 1968. Crime and punishment: an economic approach. *Journal of Political Economy* 76, 169–217.
 Bovbjerg, R., Sloan, F., 1998. No fault for medical injury: theory and evidence. *University of Cincinnati Law Review* 67, 53–123.
 Brennan, T., Leape, L., Laird, N., et al., 1991. Incidence of adverse events and negligence in hospitalized patients: results of the Harvard Medical Practice Study I. *New England Journal of Medicine* 324, 370–376.

¹¹ We implicitly assume, as is standard in the agency literature, that the hospital has all of the bargaining power, such that the optimal outcome-contingent contract will be one that keeps the physician precisely at his reservation utility.

- Brennan, T., 2000. The institute of medicine report on medical errors – could it do harm? *New England Journal of Medicine* 342, 1123–1125.
- David, B., Appleby, J., 1999. Medical Mistakes 8th Top Killer. *USA Today*, November 30.
- Environmental Protection Agency (EPA), 1995. Incentives for self-policing: discovery, disclosure, correction and prevention of violations. *Fed. Reg.* 66,706, 22 December.
- Grossman, S., Hart, O., 1983. An analysis of the principal-agent problem. *Econometrica* 51 (1), 7–45.
- Iezonni, L.I., 1997. The risks of risk adjustment. *Journal of American Medical Association* 278, 1600–1607.
- Innes, R., 1999. Self-policing and optimal law enforcement when violator remediation is valuable. *Journal of Political Economy* 7 (6), 1305–1325.
- JCAHO, 2004. Sentinel Events Policy, online at <http://www.jcaho.org>.
- Kaplow, L., Shavell, S., 1994. Optimal law enforcement with self-reporting of behavior. *Journal of Political Economy* 102 (3), 583–606.
- Kohn, L., Corrigan, J., Donaldson, M. (Eds.), 1999. *To Err is Human: Building a Safer Health System*. National Academy Press, Washington, DC.
- Leape, N., Brennan, T., Laird, N., et al., 1991. The nature of adverse events in hospitalized patients: results of the Harvard Medical Practice Study II. *New England Journal of Medicine* 324, 377–384.
- Leape, N., Woods, D., Hatlie, M., et al., 1998. Promoting patient safety by preventing medical error. *Journal of American Medical Association* 280, 1444–1447.
- Liang, B., 1999. Error in medicine: legal impediments to US reform. *Journal of Health Politics, Policy, and Law* 24, 27–58.
- Liang, B., 2000a. Risks of reporting sentinel events. *Health Affairs* 19, 112–120.
- Liang, B., 2000b. Creating problems as part of the ‘solution’: the JCAHO sentinel event policy, legal issues, and patient safety. *Journal of Health Law* 33, 263–276.
- Medical Error Reduction Act, 2000. S2038, 106th Congress, 2nd Session, 2000.
- MIT, 2001. Surgical black box. *Technology Review* 104 (10), 19.
- Patient Safety and Errors Reduction Act, 2000. S2738, 106th Congress, 2nd Session, 2000.
- Pear, R., 1999. Group Asking U.S. for New Vigilance in Patient Safety. *New York Times*, November 30.
- Pfaff, A.S.P., Sanchirico, C.W., 2000. Environmental self-auditing. *Journal of Law, Economics & Organization* 16 (1), 189–208.
- Polasky, S., Doremus, H., 1998. When the truth hurts: endangered species policy on private land with imperfect information. *Journal of Environmental Economics and Management* 35, 22–47.
- Sage, W., Hastings, K., Berenson, R., 1994. Enterprise liability for medical malpractice and healthcare quality improvement. *American Journal of Law and Medicine* 20, 1–28.
- Sage, W., 2001. Principles, pragmatism, and medical injury. *Journal of American Medical Association* 286, 226–228.
- Studdert, D., Brennan, T., 2001. No-fault compensation for medical injuries: the prospect for error prevention. *Journal of American Medical Association* 286, 217–223.
- Vincent, C., 2003. Understanding and responding to adverse events: health policy report on patient safety. *The New England Journal of Medicine* 348 (11), 1051–1056.