Looking for Trouble in All the Right Places: The Legal Implications Associated with "Electronic Signatures" and High-risk Clinical Situations

Gabriel J. Escobar, Bruce F. Folck, Marla N. Gardner, June Ma, Larry I. Palmer, Bryan Liang, Linda K. Nozick

Abstract

Background: Voluntary reporting systems identify only a fraction of medical errors. Electronic identification mechanisms, which are more efficient, have been defined for adverse drug events. However, similar systems are lacking for other types of errors. **Objective:** The investigators sought to define probabilistic strategies that could support quality improvement and medical error detection by decreasing the need for unselected manual chart review. **Design:** Combinations of administrative data and laboratory test results ("electronic signatures") were employed to identify discrete, high-risk clinical situations among health plan members of a large managed care organization. The design used was a retrospective cohort study linking hospitalization records, outpatient records, and laboratory results that were formatted using approaches developed for physiologic severity scoring. The original outcomes of interest for the study were clinical situations (e.g., birth injuries or delayed diagnosis of myocardial infarction) that have a strong association with human error. **Results:** When presented with preliminary results, senior leaders in the investigators' parent organizations raised a number of objections to any public presentation or publication of the results. Because of these objections, the quantitative results presented in this report focus on rapid detection of one outcome—prolonged neonatal assisted ventilation—that has a weak association with human error. Using recursive partitioning, the investigators were able to define subsets of newborns for whom the frequency of the outcome of interest was substantially higher than in the general population (1 percent). For example, an electronic signature identified a subset of infants (comprising 4 percent of the birth cohort) in which the outcome of interest occurred in 22 percent of the newborns. **Conclusions:** Use of probabilistic electronic strategies could yield significant benefits in medical error research as well as major operational improvements in medical error detection and reporting, quality assurance, and quality improvement. However, three barriers are likely to limit the use of such "electronic signatures"—fear of malpractice litigation, fear of lawsuits invoking "enterprise liability," and high development costs. Entities most likely to benefit from these approaches are those with a critical mass of experienced personnel, a circumstance that can spread the development costs over a large number of hospitals and/or clinics.

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
2005	**N/A**	**-**

4. TITLE AND SUBTITLE
Looking for Trouble in All the Right Places: The Legal Implications Associated with "Electronic Signatures" and High-Risk Clinical Situations

5a. CONTRACT NUMBER
5b. GRANT NUMBER
5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

5d. PROJECT NUMBER
5e. TASK NUMBER
5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Agency for Healthcare Research and Quality 540 Gaither Road, Suite 2000 Rockville, MD 20850

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSOR/MONITOR'S ACRONYM(S)
11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release, distribution unlimited

13. SUPPLEMENTARY NOTES
Published in Advances in Patient Safety: From Research to Implementation. Volumes 1-4, AHRQ Publication Nos. 050021 (1-4). February 2005. Agency for Healthcare Research and Quality, Rockville, MD. http://www.ahrq.gov/qual/advances/

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			
unclassified	**unclassified**	**unclassified**	**UU**	**18**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

Introduction

Considerable public attention has been devoted to whether reporting of human error in medicine should be mandatory or voluntary.[1–3] Less attention has been given to the problems of implementing error-detection and error-reporting systems in health care organizations. In this report, we describe the problems encountered when we developed and attempted to report a probabilistic electronic scanning mechanism in a large managed care organization, the Northern California Region of the Kaiser Permanente Medical Care Program (KPMCP), which consists of three corporations representing the program's doctors (The Permanente Medical Group, Inc.), hospitals and clinics (Kaiser Foundation Hospitals, Inc.), and insurance plan (Kaiser Foundation Health Plan, Inc.) operating under a mutual exclusivity arrangement.

Most hospitals rely on three mechanisms for error detection—manual chart review, which often involves a clinical department examining all eligible patient charts; voluntary reporting (often referred to as "incident reporting"); and chart reviews combined with clinician interviews following inquiries from the legal system. These mechanisms have been shown to be inferior to structured approaches,[4, 5] with a major problem being the reluctance of individuals to report incidents.[6] Current error-reporting mechanisms also have another major limitation—timeliness, with notification time lags as long as several years from when an event occurred.

One of the best examples that demonstrates the potential of electronic scanning is a study that electronically captured adverse drug events (ADEs) among 36,653 hospitalized patients over an 18-month period. Bates et al.[7] employed a probabilistic approach that relied on the fact that ADEs are often associated with distortions in the usual pattern of electronic information (i.e., "signals") generated by the process of care. These researchers developed computer algorithms to scan their hospital's computer systems for specific patterns—which we call "electronic signatures"—that were associated with ADEs (e.g., reports of elevated drug levels or of unusual drug or test combinations). Using several different methods, they verified 731 ADEs out of a total of 557,860 drug exposures in their cohort. Of these events, only nine (1.2 percent) were identified using traditional incident reports. In contrast, electronic scanning identified 631 (86 percent) of the ADEs, while the remaining cases were identified by voluntary reporting following electronic prompting.

The purpose of our study was to expand on the work of Bates et al. in three ways. First, we sought to use electronic scanning for the identification of discrete clinical events other than ADEs. Medication errors are important, but they are not the only kind of error in medicine. High capture rates of other types of errors can be achieved using rigorous chart review protocols,[4, 5] but it would be useful to define strategies that minimize manual chart review. Second, we employed electronic scanning based on two commonly available information systems: an admission-discharge-transfer (ADT) hospitalization database and a laboratory database, both of which are in common use. This differs somewhat from the

approach used by Bates et al.[7] in that their algorithms took advantage of a hospital with unusually sophisticated information systems. Finally, unlike Bates et al.'s single-center study, our electronic scanning algorithms were designed for use across an entire hospital-and-clinic system consisting of 16 hospitals and 34 clinics that employed the same information systems.

Design and results

Our original goal was to define "electronic signatures" for clinical events that could have a strong association with medical error, such as birth injuries or delayed diagnosis of myocardial infarction. However, when we presented our preliminary results to senior KPMCP leaders, they raised a number of significant institutional concerns that are the primary focus of this report. The first concern was that any report involving public presentation or publication in the peer-reviewed literature should be restricted to outcomes having only a weak association with human error. Because of this concern, the quantitative component of this report focuses on the occurrence of neonatal length of assisted ventilation (LOAV) ≥ 5 days, an outcome which is rarely of medical-legal concern. This is because many newborns, particularly premature infants or term infants treated for sepsis, pulmonary hypertension, or congenital anomalies, experience prolonged assisted ventilation without any error having occurred. The data for LOAV ≥ 5 days is provided as background for a description of the interactions we had with our organizational leadership. The problems we encountered, and the solutions we adopted, have significant implications for other investigators seeking to improve error detection, error reduction, and quality improvement processes.

Our conclusions regarding the implications of our approach are based on extensive conversations and correspondence with several individuals. These involved senior leaders, including legal counsel, of the KPMCP, and program staff at the National Patient Safety Foundation, one of the entities funding this research.

Development of electronic signatures

The electronic signatures described in this report for identifying babies with LOAV ≥ 5 days were selected on the basis of potential usefulness to a hypothetical obstetrics or neonatology department that uses traditional screening mechanisms (i.e., those involving review of 100% of obstetric or neonatal charts). We are reporting on signatures that achieved the maximum degree of "enrichment." For example, if a department ordinarily reviewed 1,000 charts to identify 12 events, it would be useful to identify a reduced group of charts, which was "enriched" by a factor of 10 (e.g., a subset of 75 records, 9 of which had the outcome of interest).

We developed electronic signatures using a retrospective cohort study design, using linked electronic records from five KPMCP hospitals. The study was

approved by the KPMCP Institutional Review Board for the Protection of Human Subjects. Eligible study subjects were (1) all women who delivered a live-born infant during a 3-year period and (2) all babies born alive to these women.

Inferences regarding application of these methods to other situations are based on consideration of the following themes in the medical literature: common, well-described physiologic processes (e.g., the known rise of certain enzymes following myocardial infarction),[8] probabilistic approaches based on the relationship between errors and adverse patient outcomes,[9–12] specific descriptions of error detection and reduction efforts,[13–15] and the superiority of process (as opposed to outcomes) measures for detecting problems in the quality of care.[16]

Using methods described elsewhere,[17–21] we identified and linked all eligible maternal hospitalization records, maternal outpatient visit records, neonatal hospitalization records, and specific components ($PaCO_2$, PaO_2, FiO_2, and base deficit) of neonatal arterial blood gas measurements obtained during the first 24 hours of age. We defined 22 candidate predictor variables, which were of two types: raw (i.e., laboratory data that were not transformed in any way) and derived (i.e., laboratory results that were transformed into scores or indices). The approach used to create derived predictor variables was based on the development of the Score for Neonatal Acute Physiology (SNAP), Versions I and II.[22–24] Table 1 provides examples of the candidate predictor variables. Interested readers can obtain a complete list of predictors from Dr. Escobar.

Because all of the predictor variables that we employed were highly correlated with positive or negative correlation coefficients (absolute value exceeding 0.5), we elected to develop our models using a nonparametric technique known as recursive partitioning. Recursive partitioning analyses were performed using Classification and Regression Trees (CART) software.[25, 26] The CART software we employed uses one-fourth to one-third of a dataset as a "learning" dataset and then generates a validation outcome tree with terminal nodes. These nodes are clusters of records that meet specific criteria, and thus are enriched with the outcome of interest. For simplicity, we are reporting our results as extrapolated to the entire dataset, not just to the validation dataset. Also, in-hospital deaths of neonates have been grouped with those infants with LOAV ≥ 5 days, making the assumption that newborns who died in the neonatal period would have required prolonged assisted ventilation had they survived.

Findings

During the study period, a total of 41,439 deliveries occurred at the study sites, with a total of 42,110 live births. Table 2 summarizes data from this population. Figure 1 shows the results of recursive partitioning applied to the subset of 40,490 babies born at 36 weeks or more of gestation, a group in which babies who experience prolonged assisted ventilation are very rare (only

Table 1. Examples of candidate predictor variables employed to predict for prolonged neonatal assisted ventilation

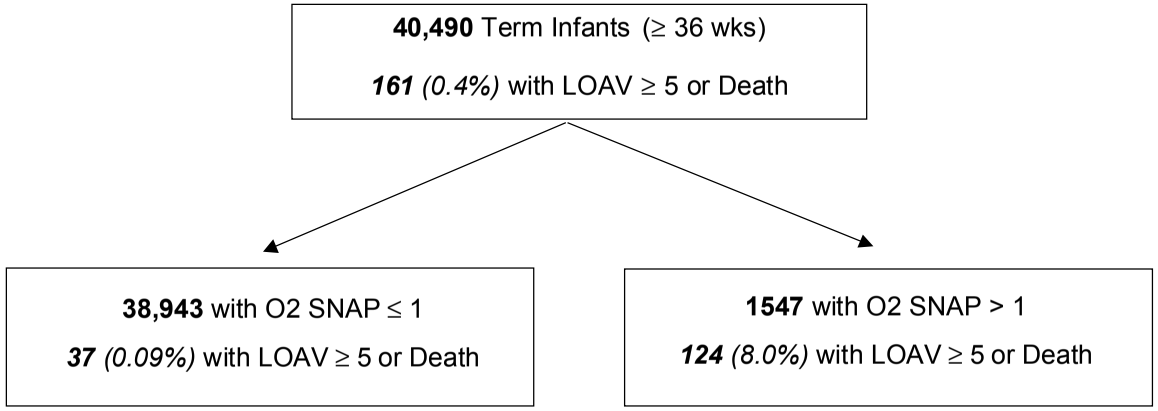
Variable(s)	Definition	Rationale for definition of variable
O2SNAP1 O2SNAP2	Oxygen component of the Score for Neonatal Acute Physiology (SNAP), versions I and II. Ratio of PaO_2 to FiO_2, with point scores assigned for specific ranges (e.g., O2SNAP2 = 28 if ratio is < 0.3).	High oxygen requirement is known to be associated with neonatal mortality and morbidity in multiple outcomes studies conducted by investigators developing two versions of the SNAP.
ABSCORE	Acid burden score. Cumulative score, with points assigned for each arterial pH obtained during first 24 hours of age. Point scheme is based on SNAP-II and is relatively independent of testing frequency. Details available from Dr. Escobar.	Persistent acidosis would be expected to be associated with a requirement for assisted ventilation.
FTOX	Failure to oxygenate score. Cumulative score, with points assigned for each abnormal arterial value of PaO_2/ FiO_2. Point scheme is based on SNAP-II and is relatively independent of testing frequency. Details available from Dr. Escobar.	Persistent oxygen requirement would be expected to be associated with a requirement for assisted ventilation.
DELTAPHA DELTAPHB	Difference between *first* non-cord (arterial, venous, or capillary) pH measurement and the *second* such measurement (DELTAPHA); difference between *first* and *worst* pH's (DELTAPHB). Variables set to equal zero if a) no blood gas measurements were obtained or b) only a single blood gas measurement was obtained.	Sick infants would be expected to have multiple blood gases, often with an initial, highly unphysiologic value (e.g., 7.12) followed by a second value which might represent either correction to normal range (7.30–7.45) or an overshoot into the alkaline range (> 7.45, which often occurs during management of pulmonary hypertension). Healthier infants might have only a single blood gas.

Table 2. Patient population employed to derive electronic signatures for prolonged neonatal assisted ventilation

Gestational Age	Number of births (%)	Number ever ventilated (%)	Number with LOAV ≥ 5 days (%)	Number of in-hospital deaths (%)	Number with either death or LOAV ≥ 5 days (%)
< 33 weeks	868 (2.06)	551 (63.48)	334 (38.48)	151 (17.40)	454 (52.30)
33–36 weeks	2695 (6.40)	323 (11.99)	62 (2.30)	17 (0.63)	78 (2.89)
≥ 37 weeks	38547 (91.54)	369 (0.96)	90 (0.23)	26 (0.07)	112 (0.29)
All gestations	42110	1243 (2.95)	486 (1.15)	194 (0.46)	644 (1.53)

LOAV = length of assisted ventilation

Figure 1. Results of recursive partitioning among infants born ≥ 36 weeks gestation



Node with twentyfold enrichment (8.0% versus 0.4%) is on lower right. LOAV = length of assisted ventilation; SNAP = Score for Neonatal Acute Physiology (see text for citations).

0.4 percent of babies). An electronic signature based on one of the oxygenation components of the SNAP,[22–24] and on the failure to oxygenate (FTOX) variable described in Table 1, achieved a twentyfold enrichment in prolonged assisted ventilation (i.e., the proportion with the outcome of interest rose from 0.4 to 8.0 percent). Applied to the entire cohort of 42,110 births, the signature shown in Figure 1 identified a subset of 1,724 infants (4 percent) in whom the rate of LOAV ≥ 5 days was 22 percent (in contrast to 1 percent for the entire cohort).

Figure 2 shows the results of a different electronic signature, based on the relationship between maternal triage in the labor and delivery service and the occurrence of LOAV ≥ 5 days among deliveries occurring at less than 36 and at least 36 weeks of gestation. It shows that women who deliver at 36 or more weeks of gestation and who were "triaged" home (i.e., who have to return for labor and delivery) are more likely to deliver a newborn with LOAV ≥ 5 days. In contrast, no such relationship was evident among deliveries occurring at less than 36 weeks.

Other possible uses of this approach

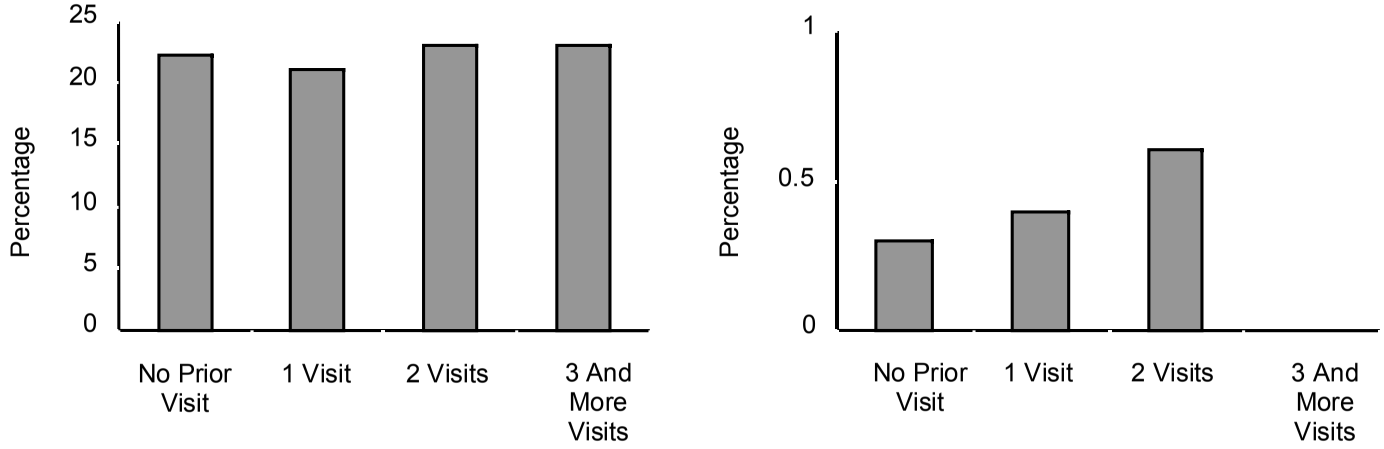
Table 3 summarizes five possible electronic signatures for important clinical situations. Each of these signatures is based upon commonly available information systems and plausible clinical scenarios. For example, situation #3 involves patients who undergo routine surgical procedures or who deliver an infant. Such patients ordinarily do not require arterial blood gas measurement and, in these cases, after removing patients with known pre-existing comorbidities, the simple fact that an arterial blood gas measurement was obtained—rather than any particular value of the pH or PaO_2—would be indicative of a greater chance of error.

Interactions with our parent organizations

The reactions we encountered from our parent organizations were mixed. Some quality and risk managers were extremely enthusiastic about transplanting these techniques from the research environment into the operational arena. Others were uncomfortable with that possibility because, given limited resources and competing priorities, it would be difficult to incorporate these techniques into existing operations.

While some senior leaders were enthusiastic about conducting research on electronic signatures for purely internal purposes (i.e., within the protected environment of quality management bodies), considerable discomfort was expressed with the notion of conducting research and publishing it in the public domain. Our initial reaction to this was one of dismay. At first, it seemed unlikely that presenting or publishing data involving the electronic signatures shown in Table 3 could present a risk to any organization or individual. None of these signatures, by themselves, constitute proof that an error leading to an adverse

Figure 2. Relationship between maternal triage in labor and delivery and the occurrence of prolonged assisted ventilation



Frequency in percent of lengthy assisted ventilation (duration ≥ 5 days) among preterm (<36 weeks, left panel) and term (≥ 36 weeks, right panel) deliveries. Note that there were no women who delivered at term who made more than two visits to labor and delivery prior to giving birth. GA = gestational age.

Table 3. Potential electronic signatures based on commonly available information systems

High risk clinical situation	Electronic signature	Rationale
Delayed or incorrect diagnosis in patient with myocardial infarction	Outpatient diagnosis of "abdominal pain" at time T_0 followed by elevated creatine phosphokinase and/or troponin levels at time $T_0 + x$	Patients with myocardial infarction may have an atypical presentation or a typical presentation may be erroneously categorized. In a subsequent visit, elevations of cardiac enzymes might be detected and used to make correct diagnosis.
Kernicterus (or near miss for kernicterus)	Elevated serum bilirubin for age at time T_0 with no subsequent serum bilirubin measurement obtained by time $T_0 + x$	Newborns with significantly elevated serum bilirubin should have prompt followup testing.
Postpartum hemorrhage	"Delta hematocrit"–difference between admission hematocrit and hematocrit obtained after delivery	Hematocrit is routinely measured in many hospitals. If mother experiences a serious deterioration, postpartum hematocrits are likely to be measured.
Surgical or anesthesia complication; major postpartum physiologic derangement	Operating room registration for simple procedure (e.g., appendectomy, routine hernia repair), followed by presence of any arterial blood gas measurement	Arterial blood gas measurement is not indicated in most surgical procedures or following normal delivery. Presence of such measurements is a strong indication that a patient deterioration occurred.
Birth injury	Patterns similar to those described in text; cord blood gas measurements; liver function tests; renal function tests	Newborns with birth injury often develop multiorgan failure in addition to respiratory failure and pulmonary hypertension.

outcome or a "near miss" had occurred. Plausible clinical scenarios exist that could be associated with the presence of any of these patterns.

However, upon further reflection and consultation, we realized that fears about publishing such data in the public domain have a rational basis. We suspect that these fears and concerns are shared by other organizations—even those with lower profiles in their communities than the KCMCP—and are acting as strong disincentives to conduct human error research throughout medicine. The major concerns raised were the following:

1. Being the first to report rates of errors—or even proxies for errors—would be damaging from a public relations standpoint in those situations for which there are no published comparison data, so that even a low rate would look bad to the public.
2. Results of an electronic scanning project might be subpoenaed by a plaintiff's attorney and lead to identification of other patients who became critically ill, subjecting the organization to further litigation. It should be noted that, in the United States, the statute of limitations for some specialties, like obstetrics, can be considered, for practical purposes, virtually limitless due to a wide variety of legal theories.[27, 28]
3. Results of electronic scanning might be subpoenaed and used in court as evidence of a pattern of inappropriate care within an institution, an assertion that could be employed in a number of situations.
4. Publication of a paper describing results of electronic scanning might lead plaintiffs' attorneys to seek out other families or patients who experienced an adverse outcome during the time period and at the locations reported in a scientific publication.
5. Some clinicians were also very worried about being "judged" on the basis of purely electronically collected data (i.e., that electronic signatures could become de facto measurement standards). Clinicians want an impartial and scientific form of peer review to assess whether an avoidable error did in fact occur.
6. Because electronic scanning is more efficient than manual chart review, the number of cases brought for review by quality management bodies could increase substantially. Given limited resources, these bodies might not be able to handle the increased load.

After extensive discussions and consultations, a compromise strategy to overcome these objections was agreed upon. For internal purposes, we employed the methods described here to support KPMCP projects whose measurement components reside within protected domains, such as facility or regional peer review bodies. For purposes of communicating with the scientific community, we selected the outcome reported here—LOAV ≥ 5 days—as an example of a serious clinical situation that, although significant in terms of a patient's illness severity, has a weak association with human error.

Discussion

We have shown that it is possible to employ arterial blood-gas results and triage patterns to define electronic signatures associated with the occurrence of a discrete clinical event (i.e., occurrence of prolonged, assisted ventilation lasting at least 5 days during the initial neonatal period). Use of this approach could be extended to situations that are more likely to have a high association with the occurrence of human error, such as those shown on Table 3. In busy hospitals using older methods, the effects of switching to probabilistic approaches could be substantial. For example, a neonatology director employing manual chart review to identify babies with long ventilator courses in a hospital with 3,000 deliveries per year could reframe his or her search strategies. One option is to use traditional approaches, reviewing 250 charts per month to find all eligible charts. Another option is to use electronic scanning, reviewing 10 charts a month to find two problem charts very rapidly (the signature we developed is based on data elements available within 24 hours after birth, which makes them more timely than data generated at the time of discharge, such as billing data). In the usual course of scientific inquiry, the next logical steps would be to test such signatures by scanning data from an eligible population, defining their sensitivity and specificity, and calculating the area under their receiver-operator characteristic curves, and then to present these findings to the scientific community for further discussion, validation, and/or refutation. We have, in fact, refined our work on the prediction of assisted ventilation and have submitted this for publication elsewhere.[29]

In the case of adverse drug events, some peer-reviewed publications have been fairly specific with respect to delineating times and locations, sometimes to the point that a reader could pinpoint the specific hospital ward involved.[7, 30] Nonetheless, just a few years later, and despite a significant increase in both public attention to human error[31] and funding for patient safety research,[32] large health care institutions are still uncomfortable with the publication of research findings in this area. What has changed?

We suspect, to quote Davidoff,[33] there is another "elephant in the room"—enterprise liability[34–37]—and that its impact may differ from what Kohn anticipated in the Institute of Medicine report.[38] The term "enterprise liability" itself needs clarification in the context of present efforts to reduce medical error and assuage provider concerns about malpractice claims.[39] For our purposes, "enterprise liability" refers to various circumstances when an organization, for instance a hospital, health maintenance organization, or health plan, is potentially liable to injured patients. In one situation, the hospital might be liable for failure to fulfill its obligation to patients, for instance, by not providing appropriate staff in the emergency room. In another situation, a hospital might be held legally responsible for the injuries caused by a nonemployee, such an independent contractor surgeon. Fear of litigation aimed at establishing some form of organizational liability is likely to exceed fear of malpractice litigation aimed at physicians and to deter quantitative research in the area of human error.

Fear of enterprise liability could make performing research on systematic human error and reporting the findings potentially risky to health care organizations. Consider this hypothetical scenario: A physician-researcher employed by a university desires to perform research using data and information from affiliated hospitals and their feeder clinics. The researcher proposes to employ institutional databases to assess a new quality marker (e.g., ratio of time to diagnosis of a particular disease state on the weekend to time to diagnosis on a weekday). The corporation feels that the project is of theoretical interest and might conceivably provide some benefits, and so gives permission for the researcher to submit a proposal for a Federal research grant. After the researcher obtains funding and approval by the organizational Institutional Review Board, the appropriate corporate officer grants access to the databases required to conduct the study.

The study is performed, and its results indicate that significant variation exists across the feeder clinics for the new quality marker. One study conclusion is that, due to delays in diagnosis, not all patients with the relevant diagnosis appear to have an equal chance of being referred to entity hospitals with appropriate clinical facilities for treatment. Although the publication does not make strong inferences about the possible clinical consequences, a reasonably careful reader might find some significant ramifications. The results are reported; the sites of the study are not named, but the state and affiliation of the researcher are. The researcher notes that although the quality indicators obtained do not suggest poor care, no inferences are possible due to the newness of the indicator and the lack of comparative data in the literature.

Some time later, an attorney for an individual whose diagnosis of the above disease state was delayed hears of the researcher's study through a presentation at a scientific conference. The attorney makes the correct inference that the sites described in the researcher's presentation include the one where his client received care. The attorney sues the hospital corporation for two reasons: First, it is a "deeper pocket" than an individual provider, and thus amenable to greater damage judgments. Second, by suing the corporate entity, the attorney aims to employ the legal discovery process to obtain all information regarding the researcher's data as well as corporate documents.

These documents would be desirable to the attorney because they could provide persuasive evidence to a jury that (1) the corporate entity approved the study, as evidenced by its Institutional Review Board approval and its officer granting access to organizational databases; (2) the hospital corporation knew—or should have known—about the high level of variation between facilities it owned; and (3) the corporate entity failed to address the problems which were implied by the researcher's study. Importantly, neither the corporate documents nor the researcher's data would be protected by any legal shields such as peer review/ quality assurance privilege or attorney-client privilege.[40–42] In fact, even root cause analyses and corrective action plans have been deemed discoverable.[43, 44]

Such a sequence of events, plausible under the U.S. legal system and which could also include the possibility of plaintiff bias,[44] would tend to chill

institutional promotion of human error research, or, at the very least, publication of its findings. Several implications are likely. First, since it is well known that existing mechanisms underestimate the incidence of errors,[45] an organizational entity that considers making well-structured efforts to detect errors will necessarily need to take a broader view of the costs of the error-detection process. The organization will need to consider not just the direct costs of error detection, but also the costs of implementing corrective action, the potential costs of legal exposure, and the effects on public relations should litigation occur. Second, corporate entities are likely to require the use of de-identification mechanisms such as those strongly recommended by the Institute of Medicine's Report, *To Err Is Human.*[38] A similar approach, using pooled datasets, has become routine in aviation,[46] and some de-identification has been used in medical research.[47–49] However, de-identification may not be a viable alternative for individual entities (as opposed to consortia that pool data) interested in promoting patient safety research. Consequently, both cost and confidentiality concerns would tend to make research into medical error more difficult and expensive.[43, 50]

Indeed, some authors have already noted that attempts to gather more data as part of the efforts to enhance patient safety are "on a collision course with the medical malpractice system."[51] Their preferred solution—no-fault compensation for medical injuries—is not likely to be considered by policymakers until the notion of enterprise liability is addressed as part of a system of public accountability for error prevention.[39] Furthermore, the logic of a systems approach to patient safety is to look at organizational, as opposed to individual, practices as a means of error prevention or reduction. The proponents of a no-fault system of compensation recognize the changing face of health care financing and delivery. They point to organizations or enterprises, not individual physicians or their insurers, being the source of compensation.[51] This is consistent with some calls for reformulation of ethical precepts suitable for system accountability.[52]

At least one author has suggested that the term "vicarious liability" should be used instead of "enterprise liability" to remind us of the potential that liability rules have to act as incentives for organizations to improve the quality of care. Such improvements could take place through new kinds of contractual arrangements with physicians and other professionals.[53] Organizations, as opposed to individual professionals, are in the position to use newer approaches to error reduction, such as the ones suggested by our study. Organizations must understand the risks, not only to themselves as corporate entities, but also to the quality of care. Furthermore, they must understand the possible liability risks of *not* undertaking new studies to reduce medical error.[54] Perhaps a day will come when corporations will increase support for human error research and human error reduction, because having corrective processes in place may not only be better business practice, but also provide protection from enterprise liability.[55]

While it is desirable to have health care organizations begin to use probabilistic approaches such as described in this report, two other practical considerations must be kept in mind by policymakers wishing to increase patient safety research and societal attention to patient safety.

The first consideration is that probabilistic approaches to error detection are most likely to be employed by large corporate entities. The costs of developing and validating electronic error detection mechanisms are high and, with the exception of a few stand-alone hospitals that have experience in attracting research funding, it is unlikely that smaller institutions will be able to afford the high development costs. This is particularly the case in the context of smaller institutions' historical financial weakness and need to address other legal mandates.[43]

The second consideration is that entities wishing to employ these techniques for internal quality assurance and improvement must consider not only enterprise/organizational liability, but also the specific provisions of the laws relating to the legal discovery process in the jurisdictions in which they operate. In the United States, there has been a general trend towards decreasing legal protections of the internal peer review process.[56] Limited protection is extant in some circumstances, but the specific legal rules vary by jurisdiction, and researchers must pay meticulous attention to the law as it applies to them at the location where they are working.[38, 56] For example, in some jurisdictions, even casual sharing of data from a peer review committee to others internally in the "normal course of business" could lead to loss of legal protection of that information and thus disclosure of that sensitive information.[56]

Conclusion

Use of electronic scanning of medical records for detection of human error is a viable approach for both research and quality improvement. However, implementation of this approach requires that individuals and organizations consider a number of issues that have received relatively little attention in the literature. One issue is the need to separate activities into two domains: those conducted under peer review protection and those conducted for publication or dissemination. A second issue is that a broader cost- and risk-benefit assessment needs to be performed prior to incorporating this approach into operations. Finally, significant advances in the use of electronic scanning are unlikely to occur unless researchers and practitioners take due consideration of enterprise liability and of the need to employ de-identification mechanisms.

Acknowledgments

This research was funded by the National Patient Safety Foundation, the Innovation Program of The Permanente Medical Group, Inc., and Kaiser Foundation Health Plan, Inc.

We wish to thank the following individuals for their advice and encouragement during the course of the study and for reviewing the manuscript: Philip Madvig, Joseph V. Selby, Andy L. Avins, Suzanne Graham, Douglas Bonacum, and Dean Sittig. We thank Kimberley Harris for her help with manuscript formatting and preparation.

This research was presented on April 23, 2002, at a platform session at the National Patient Safety Foundation's Annenberg Conference on Patient Safety, in Indianapolis, Indiana.

Author affiliations

Kaiser Permanente Division of Research, Oakland, CA (GJE, BFF, MNG). Cornell University School of Civil and Environmental Engineering, Ithaca, NY (LN, JM). University of Louisville School of Medicine, Louisville, KY (LIP). Institute of Health Law Studies, California Western School of Law and University of California-San Diego School of Medicine, San Diego, CA (BAL).

Address correspondence to: Gabriel J. Escobar, Kaiser Permanente Division of Research, Systems Research Initiative, 2000 Broadway, 2nd Floor, Oakland, CA 94611; phone: (510) 891-3502; fax: (510) 891-3508; e-mail: Gabriel.Escobar@kp.org.

References

1. Wu AW. Handling hospital errors: is disclosure the best defense? [editorial]. Ann Intern Med 1999;131(12):970–2.
2. Kraman SS, Hamm G. Risk management: extreme honesty may be the best policy. Ann Intern Med 1999;131(12):963–7.
3. Cohen MR. Why error reporting systems should be voluntary [editorial]. BMJ 2000;320(7237):728–9.
4. Stanhope N, Crowley-Murphy M, Vincent C, et al. An evaluation of adverse incident reporting. J Eval Clin Pract 1999;5(1):5–12.
5. Vincent C, Stanhope N, Crowley-Murphy M. Reasons for not reporting adverse incidents: an empirical study. J Eval Clin Pract 1999;5(1):13–21.
6. Lawton R, Parker D. Barriers to incident reporting in a healthcare system. Qual Saf Health Care 2002;11(1): 15–8.
7. Bates DW, Cullen DJ, Laird N, et al. Incidence of adverse drug events and potential adverse drug events. Implications for prevention. ADE Prevention Study Group. JAMA 1995;274(1):29–34.
8. Alpert JS, Thygesen K, Antman E, et al. Myocardial infarction redefined—a consensus document of The Joint European Society of Cardiology/American College of Cardiology Committee for the Redefinition of Myocardial Infarction. J Am Coll Cardiol 2000;36(3):959–69.
9. Brennan TA, Leape LL, Laird NM, et al. Incidence of adverse events and negligent care in hospitalized patients. Trans Assoc Am Physicians 1990;103:137–44.
10. Leape LL, Brennan TA, Laird N, et al. The nature of adverse events in hospitalized patients. Results of the Harvard Medical Practice Study II. N Engl J Med 1991;324(6):377–84.
11. Localio AR, Lawthers AG, Bengtson JM, et al. Relationship between malpractice claims and cesarean delivery. JAMA 1993;269(3):366–73.
12. Cooper JB, Newbower RS, Kitz RJ. An analysis of major errors and equipment failures in anesthesia management: considerations for prevention and detection. Anesthesiology 1984;60(1):34–42.
13. Bates DW, O'Neil AC, Boyle D, et al. Potential identifiability and preventability of adverse events using information systems. J Am Med Inform Assoc 1994;1(5):404–11.
14. O'Neil AC, Petersen LA, Cook EF, et al. Physician reporting compared with medical-record review to identify adverse medical events. Ann Intern Med 1993;119(5):370–6.
15. Lagasse RS, Steinberg ES, Katz RI, et al. Defining quality of perioperative care by statistical process control of adverse outcomes. Anesthesiology 1995;82(5):1181–8.
16. Mant J, Hicks N. Detecting differences in quality of care: the sensitivity of measures of process and outcome in treating acute myocardial infarction. BMJ 1995;311(7008):793–6.
17. Escobar GJ, Li DK, Armstrong MA, et al. Neonatal sepsis workups in babies ≥2000 grams at birth: A population-based study. Pediatrics 2000;106(2):256–63.
18. Selby JV. Linking automated databases for research in managed care settings. Ann Intern Med 1997;127(8 Pt 2):719–24.
19. Escobar GJ, Gardner MN, Chellino M, et al. Identification of neonatal deaths in a large managed care organization. Paediatr Perinat Epidemiol 1997;11(1):93–104.

20. Escobar GJ, Yanover M, Gardner MN, et al. Structured review of neonatal deaths in a managed care organisation. Paediatr Perinat Epidemiol 1998;12(4):422–36.
21. Escobar GJ, Fischer A, Kremers R, et al. Rapid retrieval of neonatal outcomes data: the Kaiser Permanente Neonatal Minimum Data Set. Qual Manag Health Care 1997;5(4):19–33.
22. Richardson DK, Gray JE, McCormick MC, et al. Score for Neonatal Acute Physiology: a physiologic severity index for neonatal intensive care. Pediatrics 1993;91(3):617–23.
23. Richardson DK, Phibbs CS, Gray JE, et al. Birth weight and illness severity: independent predictors of neonatal mortality. Pediatrics 1993;91(5):969–75.
24. Richardson DK, Corcoran JD, Escobar GJ, et al. SNAP-II and SNAPPE-II: simplified newborn illness severity and mortality risk scores. J Pediatr 2001;138 (1):92–100.
25. An Introduction to CART Methodology. San Diego, CA: Salford Systems; 1985.
26. Steinberg D, Colla P. CART: Tree-structured non-parametric data analysis. San Diego, CA: Salford Systems; 1995.
27. Rheingold PD. Solving statutes of limitation problems. American Jurisprudence Trials 2001; 4(supp):1–45.
28. Runer WB, Swiss DR, M.L. M. Obstetrical malpractice. American Jurisprudence Trials 2001; 40(supp):1–25.
29. Escobar GJ, Shaheen SM, Breed EM, et al. Richardson Score Predicts Short-Term Adverse Respiratory Outcomes in Newborns ≥ 34 Weeks Gestation. J Pediatr 2004;145:754-760
30. Leape LL, Bates DW, Cullen DJ, et al. Systems analysis of adverse drug events. ADE Prevention Study Group. JAMA 1995;274(1):35–43.
31. Millenson ML. Pushing the profession: how the news media turned patient safety into a priority. Qual Saf Health Care 2002;11(1):57–63.
32. Meyer GS, Eisenberg JM. The end of the beginning: the strategic approach to patient safety research. Qual Saf Health Care 2002;11(1):3–4.
33. Davidoff F. Shame: the elephant in the room. Qual Saf Health Care 2002;11(1):2–3.
34. Abraham KE, Weiler PC. Enterprise medical liability and the evolution of the American health care system. Harvard Law Rev 1994;108:381–436.
35. Sage WM, Hastings KE, Berenson RA. Enterprise liability for medical malpractice and health care quality improvement. Am J Law Med 1994;20(1-2): 1–28.
36. Liang BA. Managed care. In: Health law & policy. Woburn, MA: Butterworth-Heinemann; 2000;115–31.
37. Prosser WL, Keeton RE, Dobbs DB, et al. Respondeat superior. In: Prosser and Keeton on the Law of Torts. 5 ed. St. Paul, MN: West Publishing; 1984.
38. Kohn LT, Corrigan JM, Donaldson MS, editors. To err is human: building a safer health system. Washington, DC: National Academy Press; 2000.
39. Sage WM. Principles, pragmatism, and medical injury. JAMA 2001;286(2):226–8.
40. Liang BA. Risks of reporting sentinel events. Health Aff (Millwood) 2000;19(5):112–20.
41. Liang BA, Storti K. Creating problems as part of the "solution": the JCAHO Sentinel Event Policy, legal issues, and patient safety. J Health Law 2000;33(2): 263–85.
42. Liang BA. Promoting patient safety through reducing medical error: a paradigm of cooperation between patient, physician, and attorney. S III Univ Law J 2000;24:541–68.
43. Liang BA. The adverse event of unaddressed medical error: identifying and filling the holes in the health care and legal systems. J Law Med Ethics 2001;29(3-4):346–68.
44. Ba L. Layperson and physician perceptions of the malpractice system: implications for patient safety. Soc Sci Med;57:147–53.
45. Cullen DJ, Bates DW, Small SD, et al. The incident reporting system does not detect adverse drug events: a problem for quality improvement. Jt Comm J Qual Improv 1995;21(10):541–8.
46. Aviation Safety Reporting System. Aviation Safety Reporting System: Program overview. http://asrs.arc .nasa.gov/overview_nf.htm, accessed 2 February 2005.
47. Boruch R, Cecil J. Assuring the confidentiality of social research data. Philadelphia: University of Pennsylvania Press; 1979.
48. Astin A, Boruch R. "Link" system for assuring confidentiality of research data in longitudinal studies. ACE Research Reports, Vol. 5. Washington, DC: Office of Research, American Council on Education; 1970.
49. Avins AL, Woods WJ, Lo B, et al. A novel use of the link-file system for longitudinal studies of HIV infection: a practical solution to an ethical dilemma. AIDS 1993;7(1):109–13.
50. Liang BA, Small SS. Communicating about care: addressing federal-state issues. In: Peer review and mediation to promote patient safety. Houston J. Health L Policy. 2003;3:219-264.

51. Studdert DM, Brennan TA. No-fault compensation for medical injuries: the prospect for error prevention. JAMA 2001;286(2):217–23.

52. Liang BA. A policy of system safety: Shifting the medical and legal paradigms to effectively address error in medicine. Harvard Health Policy Rev 2004;5(1):6-13.

53. Havighurst CC. Vicarious liability: relocating responsibility for the quality of medical care. Am J Law Med 2000;26(1):7–29.

54. Arlen J, McCleod WB. Malpractice liability for physicians and managed care organizations. New York University Law Review. 2003;78:103-174.

55. Palmer LI. Patient safety, risk reduction, and the law. Houston Law Rev 1999;36(5):1609–61.

56. Liang B. Error in medicine: legal impediments to U.S. reform. J Health Politics Policy Law 1999;24(1):27–58.