Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Biomedical Informatics 36 (2003) 131–143
Journal of Biomedical Informatics
www.elsevier.com/locate/yjbin

Methodological Review

Detecting adverse events for patient safety research: a review of current methodologies

Harvey J. Murff,[a,b,*] Vimla L. Patel,[c] George Hripcsak,[d] and David W. Bates[e,f]

[a] Department of Veterans Affairs, Tennessee Valley Healthcare System, GRECC, 1310 24th Avenue South, Nashville, TN 37212-2637, USA
[b] Division of General Internal Medicine, Department of Medicine, Vanderbilt University, USA
[c] Laboratory for Decision Making and Cognition, Departments of Biomedical Informatics and Psychiatry, Columbia University, New York, USA
[d] Department of Biomedical Informatics, Columbia University, New York, USA
[e] Division of General Internal Medicine, Department of Medicine, Brigham and Women's Hospital, USA
[f] Center for Applied Medical Informatics Systems, Partners Healthcare System, and Harvard University, USA

Received 11 August 2003

Abstract

Promoting patient safety is a national priority. To evaluate interventions for reducing medical errors and adverse event, effective methods for detecting such events are required. This paper reviews the current methodologies for detection of adverse events and discusses their relative advantages and limitations. It also presents a cognitive framework for error monitoring and detection. While manual chart review has been considered the ''gold-standard'' for identifying adverse events in many patient safety studies, this methodology is expensive and imperfect. Investigators have developed or are currently evaluating, several electronic methods that can detect adverse events using coded data, free-text clinical narratives, or a combination of techniques. Advances in these systems will greatly facilitate our ability to monitor adverse events and promote patient safety research. But these systems will perform optimally only if we improve our understanding of the fundamental nature of errors and the ways in which the human mind can naturally, but erroneously, contribute to the problems that we observe.

Keywords: Patient safety; Medical errors; Research methodologies; Adverse events; Quality of care

1. Introduction

Adverse events in health care appear to be responsible for 44,000 to 98,000 accidental deaths and over one million excess injuries each year [1,2]. Estimates of the prevalence of adverse events per hospital admission range from 2.9 to 16.6% [3–6]. While medical errors have not been as extensively studied in the outpatient setting, it is estimated that 3–35% of outpatients experience an adverse drug event [7–10] and that 13% of all adverse events identified during hospitalizations occur during outpatient care [3]. In one study from a tertiary care center, 19% of patient's discharges from the hospital experienced an adverse event during the transition to outpatient care [11]. The pervasive problem of medical errors and adverse events in healthcare has made improving patient safety a national priority [12].

High quality evidence exists supporting many interventions designed to improve patient safety [13], yet much work still remains. Researchers wishing to evaluate the effectiveness of patient safety interventions are faced with a daunting challenge. Identification of adverse events is critical for improving patient safety, yet medical errors and adverse events can be challenging to measure [14]. Efforts to overcome these problems have resulted in the development of several tools that researchers and increasingly clinical teams are using to detect adverse events.

This paper reviews the methodologies that have been utilized for adverse event detection. For the purpose of this review, we have also included methods used to increase the reporting of adverse events within our definition of adverse event detection systems, as well as

* Corresponding author. Fax: 1-615-327-5381.
E-mail address: harvey.murff@med.va.gov (H.J. Murff).

1532-0464/$ - see front matter © 2003 Elsevier Inc. All rights reserved.
doi:10.1016/j.jbi.2003.08.003

more recent approaches to the application of cognitive science principles to understanding and monitoring errors. We will begin with a brief description of both voluntary and involuntary manual methods for adverse event detection, then review combined modalities involving both electronic screening and manual review, and then describe fully automated methods for adverse event detection. We discuss the cognitive perspective in the final section.

No universal definitions for descriptive terminology used within patient safety literature currently exist. This is one of the factors resulting in varying estimates of the prevalence of adverse events and medical errors. For the purposes of this paper, we will use the definitions presented by the Institute of Medicine (IOM) to define adverse events and medical errors. The IOM defined an adverse event as "an injury caused by medical management rather than the underlying condition of the patient" and a medical error as "the failure of a planned action to be completed as intended or the use of a wrong plan to achieve an aim" [1]. A preventable adverse event is an adverse event that results from an error. It is important to note is that most adverse events do not result from an error, an example being a patient who experiences an anaphylactic reaction after their first exposure to a drug to which they had no prior allergic history, and also that most errors do not result in an adverse event, for example a ten-fold overdose order for a drug that gets intercepted and corrected by a pharmacist. Medical errors occur much more frequently than adverse events and medication errors outnumber adverse drug events by 100–1 [15].

2. Manual methods

2.1. Voluntary reporting methods

2.1.1. Incident reporting

A cornerstone of safety initiatives in other industries has been the reporting and analysis of errors [16]. The commercial aviation industry has been a pioneer in this area designing non-punitive, blame-free reporting systems [17,18]. Other industries, such as nuclear power [16], have also follow this lead and established reporting systems, yet despite the successes of voluntary error reporting systems in other industries, voluntary reporting or incident reports in healthcare has been relatively unsuccessful. By incident reporting, we specifically mean the voluntary reporting of a medical event by a health care provider. Incident reports have traditionally been paper-based forms, however some online reporting systems exist [19]. Studies that have compared the rate of adverse event detection through incident reports to those detected by chart review have found that only 1.5% of adverse events [20] and only 6% of adverse drug events [21] are detected through incident reports. Nonetheless, incident reports still remain the mechanism that most institutions use to detect adverse events.

There are several important reasons why incident reporting in its current state may not be the most effective way to detect adverse events [22,23], including interruption in workflow, perception that completing a form will not result any improvement, lack of knowledge that an adverse event had occurred, and fear of exposing oneself to litigation [23–26]. Incident reports remain an attractive source of information because they are generally readily available, it is often possible to obtain detailed descriptions of what went wrong, truly catastrophic errors or events can sometimes be found which can be used to make changes, and they represent the best source of potential adverse events, or near misses [27]. However, problems with underreporting greatly limit their utility for patient safety research.

2.1.2. Prompted spontaneous reporting

Some investigators have attempted to increase voluntary reporting through continuously prompting physicians, nurses, or pharmacists to report errors or adverse events. In one study, medical residents were sent daily electronic mail reminder reminding them to report adverse events [20]. Along with the daily prompts, once weekly all medical residents were again asked to report adverse events. Adverse events were reported through either electronic mail or paper reports. Over the four-month study period the residents reported 89 adverse events while 85 were detected by chart review. Notably, only 41 of the adverse events were identified by both of the methodologies, although all the adverse events appeared real, suggesting that the true underlying rate is higher than that identified by any single methodology.

Welsh et al. [28] studied how the intensity of prompted reporting might influence reporting rates. Medical residents who were prompted on a daily basis tended to report more adverse events per month (55 ± 4) when compared to those prompted weekly (24 ± 15) or those not prompted at all (21 ± 2). Furthermore, the number of housestaff reporting adverse events increased as the intensity of prompting increased (no prompting $= 0.55 \pm 0.3$, low-level prompting $= 1.47 \pm 0.3$, intensive-level prompting $= 4.04 \pm 0.4$, $p < 0.001$). Similar to prior studies, the degree of overlap between events reported by the residents and those detected through incident reporting was small.

Another study created a voluntary reporting system in which residents detected adverse events by interviews with their peers and detected 88 adverse events over a 3-month period [29]. This study, as prior studies of prompted physician report, has several limitations. The intervention was only studied over a short period of time, and it is unclear if because of the high level of effort required maintaining these systems they could be

sustainable. Furthermore, these studies have targeted medical residents only. Not only might prompted physician reporting not improve adverse event detection rates for hospitals without residents, but also these methods might not even be effective for non-medical housestaff, such as surgical or obstetric residents.

Stimulated voluntary reporting has also been used with nurses and pharmacists [15,27]. In these studies, nurses and pharmacists were asked on each shift whether they had noted anything that might represent an adverse drug event or potential adverse drug event. Overall, this approach was much more effective for identifying near misses than adverse events.

2.1.3. Conclusions regarding reporting

Spontaneous reporting has identified a number of clusters of important problems that have resulted in important policy change, including deaths related to concentrated potassium, concentrated lidocaine, and deaths and disability related to confusion between *cis*-platinum and carboplatinum. However, such events are fortunately rare, and utilizing them best requires aggregating them at the regional or national level. At the local level, identification of a single event can sometimes reveal a sufficiently important problem that it makes sense to change a process or system, though this is unusual.

Although voluntary reporting can detect a broad range of adverse events, these systems miss the vast majority of events and cannot provide stable estimates of the true underlying defect rates, which has resulted in the development and evaluation of other detection methods that do not rely on spontaneous reporting. In this next section, we review several manual adverse event detection systems that do not require active participation including chart review, direct observation, and direct patient interviews.

2.2. Involuntary reporting methods

2.2.1. Chart review

Most of the epidemiological evidence describing adverse medical events comes from several large studies that used chart reviews to detect adverse events. The California Medical Insurance Feasibility Study, the Harvard Medical Practice Study, and the Colorado-Utah Study all described similar methodologies to detect adverse events [4,5,30]. Charts underwent a two-phase screening process. Initially a trained reviewer, usually a research nurse, examined charts for 18 screening criteria. (Table 1) Charts with one of these screening criteria then underwent physician review. Physician reviewers made judgments on whether they believed an adverse event had occurred based on the information contained within the chart. Reviewers then rated their confidence that the patients' medical management resulted in the injury.

Table 1
Screening Criteria for Adverse Events in Chart Review[a]

Screening Criteria
Hospitalization within previous year for patients less than 65 years old
Admission to any hospital after this discharge
Previous failure of medical management or unfavorable results
Trauma incurred in hospital
Unfavorable drug reaction in hospital
Transfer from general care to special care unit
Transfer to another acute care hospital
Return to operating room during hospitalization
Treatment for organ damage after invasive procedure
Acute myocardial infarction, cerebrovascular accident, or pulmonary embolus during or after an invasive procedure
Neurologic deficit at discharge
Death
Temperature higher than 38.3 °C on day before or day of discharge
Cardiac or respiratory arrest
Five-minute Apgar score <6, or complication of abortion or labor and delivery
Other undesirable outcome
Indication of litigation in the medical record
Length of hospital stay above 90th percentile for the diagnosis-related group in patient under 70, and 95th percentiles in those 70 or older

[a] Criteria from the Harvard Medical Practice Study [30].

When a physician reviewer was more than 50% confident that medical management resulted in the injury, an adverse event was considered to have occurred. Typically, either all or a fraction of the charts are reviewed independently by two physicians. If the physician reviewers disagree on whether an adverse event occurred, the physicians will try to come to a consensus or involve a third party.

While this two-step chart review process has been used to provide much of the adverse event prevalence data, several problems exist with this methodology. First, the positive predictive value of the initial screening process was low (21%) [31] resulting in the physician reviewers evaluating a high number of false-positive charts. Second, agreement between physicians is generally poor with respect to causality. In one observational study, the number of cases in which there was "extreme disagreement" as to whether an adverse event occurred actually outnumbered the cases in which the physician agreed [32]. While the κ-statistic for the physician agreement concerning an adverse event occurrence in the Harvard Medical Practice Study was moderate ($\kappa = 0.61$) [3], other investigators have found lower rates of agreements ($\kappa = 0.40$, 95% confidence interval [CI] 0.30–0.51) [33] and have questioned this overall methodology [34]. But perhaps the greatest limitation to this methodology to detect adverse events is the overall resources required, which are great when compared to other screening modalities [21]. Some investigators have used fewer screening criteria to help reduce the overall

resource burden of chart reviews [35] but in general, chart review remains an impractical means for routine adverse event detection because of the high costs.

2.2.2. Observers

Another method used to detect adverse events has been the use of trained observers. This approach has advantages over chart review in that is can be performed prospectively. Observers have been used to detect adverse events in the intensive care unit [36] and surgical care floors [37]. In one study, observers were trained in how to recognize adverse events and recorded physician's comments made during patient care activities [37]. These observers were able to detect serious adverse events in 17.7% of all patients studied.

Medication errors, defined as any error occurring during the ordering, transcribing, dispensing, administering, and monitoring of medications, are common occurrences that sometimes result in patient injury [15]. In a study that compared direct observation, chart review, and incident reports to determine the most effective means for detecting medication errors, direct observation detected 300 errors made on 2556 doses compared with only 17 errors detected on the same doses detected by chart review [38]. Several investigators have described direct observation as the most sensitive detection method for identifying medication errors [39]. A limitation of direct observation is that it can be an expensive method for error detection with a mean cost of $4.82 per error detected versus $0.63 per error detected by chart review [38]. Furthermore, even though more errors are detected, the errors tend to be less likely to result in an injury than errors detected through other methodologies [14].

2.2.3. Patient interviews

Much of the data concerning adverse events in the hospital setting has come from chart reviews. However, it has been more challenging to detect adverse events that occur in the outpatient setting using this methodology. Outpatient charts are generally less data intense than inpatient charts with patient visits occurring sporadically. As a result, many researchers have relied on patient interviews to detect adverse events in these settings [9]. Recent survey evidence suggests that patients can be good source for adverse event detection. Two recent patient surveys have indicated that 20–42% of patients have experienced and are aware of an error occurring with their medical care and that the majority of these errors resulted in serious consequences [40,41].

A recent study by Forster et al. [11], utilized patient interviews to help construct case summaries for physician reviewers. Information sources in the preparation of the case summaries included patient interviews, sign-out notes, discharge summaries, and laboratory results. Adverse events were determined based on review of these case-summaries.

Patient interviews offer a novel potential source for adverse event detection yet, similar to using observers, require a substantial resource commitment. Nonetheless, patient observations concerning their care likely uncovers adverse events that might not be detectable by other means and the approach warrants further study.

In summary, manual adverse event and error detection systems offer the ability to detect a wide array of adverse events. All of the above systems are limited by either physician reluctance to use them or the resource requirements to maintain them, but they can be very useful in research, and direct observation can be used in clinical care to detect the medication error rate because the high rate of errors allows rapid assessment of the rate [38,39].

To overcome these limitations, several electronic methods for detecting adverse events have been evaluated. Advances in information technology, the increase in hospitals with integrated information systems, and ambulatory electronic medical records have increasingly made it possible for institutions to automatically detect adverse events [42]. Institutions with computerized event-monitors in place have been able to detect a greater number of adverse drug events when compared to spontaneous reporting [21]. Even more exciting is that some of these methods can detect errors prior to the occurrence of an adverse event or before it becomes severe and thus prevent or ameliorate patient injury [43]. In the following section we review detection methods that combine automated techniques with manual review followed by fully automated systems.

3. Combined modalities

Combined modalities are methods for detecting adverse events that rely on both electronic and manual review process. In general, these systems identify an electronically stored "signal," such as a laboratory abnormality, use of a medication along with development of a symptom, as screening criteria to identify charts for further review. These systems generally require less reviewer time and thus are less expensive to operate than manual systems. However, as a result of low specificity, much of these approaches still require some form of manual review. A few of these screening methods have evolved into stand-alone systems that no longer require any form of manual review. We begin with a review of adverse event detection systems that utilize a single data source, such as administrative data, as a signal, followed by more sophisticated systems that use multiple data sources, such as pharmacy and laboratory databases, and rule-based alert logic to detect adverse events.

3.1. Adverse event detection system

3.1.1. Single data source

As discussed above, chart reviews have relied on a two-step process, which initially involves a screening review followed by physician review. However, many of the criteria used for the initial screening process could be electronically captured. Several investigators have evaluated the test characteristics of generic screens obtainable through administrative or billing data. In one study, five generic screens were used to attempt to screen for adverse events. Overall these five triggers had a positive predictive value of 20% and were able to detect 50% of all AE occurring in the cohort [44]. Removing a poorly performing billing screen resulted in an overall improvement in the positive predictive value (30%) but detected only about half as many AEs as the first method. Other investigators have had similar results with positive predictive values of the generic screens ranging from 12 to 34% [45]. Despite these low positive predictive values, strategies for adverse event detection that incorporate these electronically captured generic screens do cost significantly less per adverse event detected [44] making them an attractive alternative to pure manual review. The Veterans Affairs Hospital has developed an occurrence-screening program that utilizes similar electronically captured generic screens to search for events that might require further peer review [46].

Iezzoni developed a computer algorithm called the Complications Screening Program (CSP), which searched discharge abstracts for potential complications, using ICD-9-CM codes as a signal [47,48]. For surgical event screens, 73% were corroborated on chart review. However, only 32% corroborated for the medical event screens [49]. The major reason for the poor performance of the medical event screens resulted from the tool's inability to distinguish diagnoses that were present on admission from those that occurred during the hospitalization. Of the screens detected for the medical cases, 58% were present on hospital admission as opposed to only 13% in the surgical patients. As a result of these findings the authors did not advise using the tool as a "stand-alone" test for detecting most complications, though it still may be useful with follow-up review. Regardless, ICD-9-CM codes have been utilized to detect medical adverse events in recent publications [50], and are used by many commercial organizations to detect complications [51,52]. ICD-9-CM codes have also been used to detecting adverse events in outpatient care, but identified only a small number of events, in part because adverse event codes were rarely used in this setting [53].

There are several limitations to using ICD-9-CM codes as a search strategy. First ICD-9-CM codes are generated for reimbursement purposes rather than clinical care and may not be appropriate for clinical studies [54]. Second, ICD-9-CM codes frequently are erroneously assigned. In one study designed to determine whether clinical evidence could be found within the chart to support the ICD-9-CM codes found in the discharge abstract, only 70% of medical patient and 81% of surgical patients had objective evidence to support the code within their record [55]. Finally, as mentioned above, ICD-9-CM codes do not discriminate between pre-existing and newly documented diagnoses.

One possibility means for overcoming the limitations of using ICD-9-CM coding to detect possible adverse events is to use clinical narratives as a data source. Clinical narratives contain rich information yet the uncoded nature of this data has made it challenging to develop detection systems. Methods have been developed for both inpatient and outpatient clinical narratives.

Initial approaches to detect adverse events within inpatient clinical narrative have focused on discharge summaries. In one study, simple keyword searches were used to screen for adverse events [56]. Discharge summaries were parsed into individual words and queried for "trigger words" that represented adverse event concepts. An example of a "trigger word" was *agitation*. Of the 424 selected discharge summaries for screening, 3 contained the "trigger word" agitation and all three had an adverse event revealed on chart review. A total of 95 trigger words were included. Overall, the electronic tool had a positive predictive value of 52% (95% CI 46–58%) but had poor specificity (48%, 95% CI 42–54%). Specificity could have been markedly improved by adding a manual discharge summary review step to the screening process. However, the addition of this step would offset some of the time gains made through electronic screening. Addressing the negation terms and lexical variants that resulted from keyword queries could eliminate some false positives, however to make substantial improvements in the tool, more sophisticated searching techniques, such as natural language processing, would be required [57].

3.1.2. Integrated data source

Using an electronic medical record, Honigman et al. [53] developed a tool that searched a patient's administrative data, medication history, and clinical narrative from the visit note to identify adverse drug events. The tool initially parsed the visit note into words and word combinations. These terms were then filtered through a medical dictionary database and joined into master concepts, which could be indicative of an adverse drug event. Drugs and drug classes were then linked to the reported adverse drug events. For example, if the term diarrhea was mentioned within the clinical narrative, this term was joined to the adverse drug event concept of "diarrhea" and linked to any medication the patient might be taking that could cause diarrhea, such as

omeprazole. In a study comparing multiple detection methods (ICD-9-CM screening, allergy screening, event monitor, and text searching) the overall positive predictive value of this process was low (7.2%, 95% CI 6.8–7.5), but this methodology was estimated to have detected 91% of all adverse drug events recorded in the chart [53].

Using clinical narrative to detect adverse events has great potential as a screening tool. Advances in the searching algorithms and the use of natural language processors should greatly improve the test characteristics. Natural language processors extract concepts from free text narrative through using basic grammatical rules and knowledge about the domain [58–60]. A range of natural language processing technologies exists. Some processors cover a broad range of medicine and produce comprehensive coded output [60]. To detect medical events, one applies queries or rules to the output to find cases with relevant evidence. If patient safety terms are missing from the processor's lexicon, it will be necessary to add them. Some natural language processors are designed to detect specific concepts. In this case, it is necessary to modify the processor itself (e.g., modify a Bayesian network [61]). Natural language processors have achieved classification rates similar to that of expected manual reviewers, with receiver operating characteristic curve areas of 0.95–0.96 [58,61,62]. These natural language processors outperformed keyword searches even when negation was explicitly handled [58,61]. There is therefore enthusiasm that natural language processing may replace costly manual review of discharge summaries, operative notes, nursing notes, daily progress notes, and visit notes for detecting adverse events. Unfortunately, as of this writing, natural language processors are not broadly available, and it will take some time for commercial products to reach the performance of research systems.

3.2. Nosocomial infection detection systems

Several groups have successfully developed nosocomial infection event monitors [63]. When compared to manual review, computerized surveillance has superior sensitivity and required less person time [64]. These systems typical work by searching clinical databases for signals and then generating an alert for hospital infection control personnel.

3.2.1. Single data source

Most described nosocomial infection detection systems have utilized multiple data sources and a rule-base to generate alerts. Dessau et al. [65] have described a system that relies solely on laboratory data. This surveillance system detects changes in the incidences of microorganisms identified within the hospital laboratory databases. Isolated microorganisms identified over a weekly period are compared to isolation rates detected during a preceding period. Detection limits are determined based on the distribution of the differences between the expected event count and the observed event count. The authors note several nosocomial infectious outbreaks identified by this system. While this system appears successful at identifying nosocomial outbreaks, it would only be able to capture a small subset of all nosocomial infections.

3.2.2. Integrated data source

More often, nosocomial infection detection systems have relied on data from several sources. One approach has combined laboratory (microbiology) data and administrative data to generate alerts regarding possible nosocomial infections [66,67]. Hirschhorn et al. [66] developed a computer program that captured the duration and timing of antibiotic exposure and the ICD-9-CM coded discharge diagnosis to detect nosocomial infections occurring after non-elective cesarean sections. Alerts generated when an individual had a discharge diagnosis code suggestive of infection and duration of antibiotic exposure greater than expected for prophylaxis purposes had a high positive predictive value for detecting a nosocomial infection (94%) but a poor sensitivity (59%). Major limitations of this detection system resulted from inaccuracies in the diagnostic codes, with 35% of the infections not being identified by ICD-9-CM codes.

Other nosocomial infection surveillance systems have relied on a combination of microbiology data and an extensively developed rule-based expert system [63,68]. These systems typically search for positive culture reports from microbiology reports. Based on information obtained from the culture report, the computers activates specific modules of its knowledge base to extract other information, such as admission dates and dates of recent hospitalizations, from the patient's electronic medical record. This information is used to generate a daily report for hospital infection control personnel. These systems have been developed to be both "data driven," initiating the logic module anytime an alert signal in entered into the system, or "time driven," initiating the logic module at a set time each day. Nosocomial infection surveillance systems coupled with decision support and electronic order entry can decrease inappropriate antibiotic utilization and reduce hospital expenditures [64,69].

3.3. Adverse drug event monitors

Several methods exist for the detection of adverse drug events. Much data has been collected through prompting nurses and pharmacists to report events, manual chart reviews [15,38,70,71] and direct observation [72]. However, because pharmacy and laboratory

Table 2
Examples of signals used for adverse drug event screening

Signal	Example
Non-rule based	
Laboratory data	Bilirubin >1.5 mg/dL
	Serum creatinine rise >30% from initial value
	Increased plasma levels of aminoglycoside
Pharmacy data	Any order for naloxone
	Any order for diphenhydramine
Rule-based	
Pharmacy and Laboratory	Receiving "nephrotoxin" AND blood creatinine has risen >0.3 mg/dL in last 1 day
	On ranitidine and platelet count >100,000 and <250,000/mm^3
Pharmacy and EMR[a] Orders	Receiving diphenhydramine AND no diphenhydramine within last 7 days AND patient no on paclitaxel AND no blood transfusion in last 1 day

[a] EMR, electronic medical record.

data are computerized in most hospitals [73], adverse drug events are particularly suitable for detection through electronic surveillance [74]. In this section, we will describe adverse drug event monitors. These are tools that can either search a single data source, such as pharmacy records for an antidote order, or search a combination of data sources and utilize a rule-base to detect adverse drug events.

3.3.1. Single data source

Computerized adverse drug event monitors have used a variety of signals to detect adverse drug events. Signals include drug orders [75,76], laboratory abnormalities [77–79], and billing codes [80] (Table 2). Generally predefined criteria are specified, such as a platelet count of less than 70,000/μL, and once this threshold has been reached, an alert is generated for a clinical pharmacist to review. The major limitation to these systems has been numerous false positives. False positive rates for some of these single data source detection systems have ranged from 77 to 96% [75,77–80]. To reduce this rate of false positives, event monitors have been designed that incorporate multiple data sources and screening algorithms.

3.3.2. Integrated data source

To increase the overall numbers of adverse drug events detected, many institutions have searched several data sources for signals. Some of these systems have incorporated an alert logic into the searching strategy. An early example of an event monitor, which utilized multiple data sources stems from work done at LDS Hospital in Utah [81]. The event monitor continuously searched patient records looking for laboratory abnormalities, such as eosinophil percent greater than 6 or a digoxin level greater than 2 ng/dL. The program also utilized screening algorithms to detect sudden stop orders, dose reduction orders, or drug antidotes. The event monitor was able to detect 731 adverse drug events over an 18-month period [82]. During this same time period traditional reporting means only detected nine events. The number of false negative reports generated by this system was not described.

False positive reports were a problem and through iterative testing, certain screens that often produced false-positive alerts were dropped as signals [81]. Furthermore, logic was added to the event monitor to help further reduce false-positive reports. Rules were added to the knowledge base that would prevent the event monitor from generating an alert is patient had elevated liver function tests and an admission diagnosis reflective of liver disease.

Jha et al. [21], at Brigham and Women's Hospital in Boston, used similar signals that were used at LDS Hospital for detection of drug events, and further expanded the rule-base. Rules consisted of Boolean combinations of signals detected through medication orders, laboratory results, and medication orders associated with changes in laboratory results. To reduce false positive alerts several rules were conditional on multiple criteria being met. An example would be if pharmacy records indicated a patient was taking a nephrotoxic medication while simultaneously laboratory data indicated the patient's serum creatinine had risen. (Table 2) The event monitor detected 275 adverse drug events compared to 398 detected through manual chart review and 23 detected through incident reports. In the first 8-weeks of the study the overall positive predictive value of the screening tool was 16%. Through removal of poorly performing screens and the addition on more rule combinations, this positive predictive value increased to 23%. Over the entire study 83% of the alerts were considered false positives. Despite this high level of false positive alerts, the computer strategy required only 11 person hours a week to execute as opposed to 55 person hours for manual chart review.

Similar event monitors have been implemented in both community hospitals [83] as well as Veterans Affairs hospitals [84,85]. Raschke et al. [83] developed an event-monitor which targeted 37 specific adverse events. Because this study focused on a pre-defined set of adverse drug events, an extensive alert logic could be developed. Not surprisingly the number of false positives alerts in this study were significantly less that most prior studies, with only 47% of alerts being judged as false positives.

Drug-laboratory rules have also been used to detect adverse drug events occurring in outpatient care. As with several inpatient event monitors, Boolean combinations of new medication orders, laboratory results above or below a certain threshold, and changes in laboratory values over time were used as the alert logic

[10]. Drug-laboratory rules for outpatient care were able to detect more adverse drug events than using ICD-9-CM codes or using patient allergy information but overall had very low positive predictive value of 3%. As with inpatient monitors, with iterative testing it is likely that the performance of drug-laboratory rules will improve.

4. Automated detection systems

The "gold standard" for the determination of an adverse event has traditionally been implicit physician judgment. Implicit judgment is often a requirement because of the difficulty of establishing causality. However, some types of adverse events, such as some adverse drug events, might be able to be determined without a reliance on implicit judgment. An example might be toxic serum levels of digoxin or International Normalized Ratios (INR) greater than 6 in patients on warfarin. Thus, it could be possible to design adverse event monitors that are fully automated and require no manual review component. In this section we will describe adverse event monitors that have the potential of being fully automated.

4.1. Intra-operative monitors

Few automated adverse event monitors have described methods to detect intra-operative adverse events. One such monitor performed a structured query search of drug, laboratory, and vital sign information included within an anesthesia database [86]. Data combinations derived from this database were used to detect five types of intra-operative adverse events: hypotension, hypertension, bradycardia, tachycardia, and hypovolemia. An example of one such rule-combination for bradycardia would be: heart rate under 50 beats per minute for at least 5 minutes *and* administration of a drug to increase the heart rate within a maximum of 15 minutes following the first measurement of the event. Based on a random review of anesthesia records, all events detected through the automated means were subsequently identified by chart review. While the tool was only designed to detect a few types of adverse events, the overall accuracy of the tool could allow for the elimination of manual review.

4.2. Inpatient fall monitors

Hripcsak, Wilcox, and Stetson used natural language processing of the electronic medical record to detect inpatient falls [42]. This tool searched all inpatient radiology reports, excluding those performed during the first two days of an admission, and was able to identify those that had been requested as a result of a fall. The tool was then able to determine in which reports a fracture had been diagnosed. The fall rates detected through this fully automated tool were similar to fall rates reported within the medical literature.

5. Cognitive frameworks for monitoring and detecting adverse events for patient safety research

In light of the growing awareness of the role of human errors in widely publicized incidents such as airline accidents and complications of medical procedures, understanding of the cognitive, social, and information sciences have become crucial in order to make any significant contributions to the study and prevention of human errors. Medical decision-making is constrained by human cognitive limitations, as well as socio-cultural and organizational factors, and errors often arise due to poor and flawed memory, inexperience and poor training, as well as faulty reasoning and ambiguous communication. To improve patient safety, technology should be designed that take into account these limitations. Several industries (e.g., aviation and nuclear power plants) have been very successful in monitoring and preventing human errors through the careful design of user interfaces and use of systems theory. A key issue with many detection systems is that they may not describe errors or events in sufficient detail to make it possible to get to the root of the problem, and make appropriate systems changes, which will often involve interface design. Many organizations are beginning to apply traditional engineering methodologies to the study of patient safety concerns, such as failure mode and effects analysis (FMEA) and root cause analysis (RCA). While these approaches have produce some notable successes [87,88], in general, far too often errors and adverse events are simply accumulated in healthcare, and the data are not effectively used.

A recent report from the Institute of Medicine supports the importance of systems design by suggesting that most medical errors are related to inadequacies in process rather to incompetence of health workers, and that information technology can play a particularly important role in preventing or minimizing the consequences of such errors [12]. Examples include the computer-based verification of dosing information at the time that a drug regimen is ordered, or improved access to (and legibility of) pertinent clinical information that may prevent decision-making errors before they occur. However, there remain major challenges in the implementing and integrating such facilities into clinical environments and even technology that is designed to improve safety could result in unintended errors if improperly designed for the user.

Because errors are fundamentally cognitive, a cognitive framework is necessary to identify and specify the

underlying mechanisms that cause human error. Two ways in which cognitive factors play important roles in medical errors are through the cognitive properties of interactions between human and technology (user interfaces) and the dynamics of information processing in distributed cognitive systems involving groups of people and technology (groupware and distributed cognition). A cognitive foundation for the system is essential for a complete and in-depth understanding of medical errors. Once a cognitive foundation is developed, we can classify medical errors according to their underlying cognitive mechanisms, which we may be able to automate for adverse event monitoring and detection.

5.1. Developing a taxonomy of medical errors

Within this cognitive framework, errors are divided into two major categories: (1) slips or lapses that result from the incorrect execution of a correct action sequence and (2) mistakes that result from the correct execution of an incorrect action sequence [89]. To develop such cognitive theories, we need to start from the very definition of error: an error is the outcome of a system that fails to achieve the intended outcome in a planned sequence of mental or physical actions. In other words, errors are errors of actions. Therefore, a cognitive theory of medical errors is based on a basic theory of human actions. A cognitive theory of human action most appropriate for medical errors is the seven-stage action theory developed by Norman [90,91] and refined by Zhang and his colleagues [92–94].

One critical step toward reducing medical errors in particular and human errors in general is a cognitive taxonomy of errors that can (1) categorize all types of medical errors along cognitive dimensions, (2) associate each type of medical errors to a specific underlying cognitive mechanism, (3) describe how, explain why, and even predict when and where a specific error will occur, and (4) generate intervention strategies for each type of error. Based on Reason's [89] definition of human errors and Norman's cognitive theory of human action [90,91], a preliminary action-based cognitive taxonomy of medical errors has been developed that largely satisfies these four criteria [92]. This taxonomy can categorize all types of errors (slips and mistakes) according to the stages of the action cycle. The authors have identified a set of cognitive mechanisms (substantial but not exhaustive) that underlie each type of slip or mistake. This taxonomy can also explain why and describe how a specific error occurs. With further development and refinement, it may have enough predictive power to help designers and implementers to anticipate more effectively when and where an error might occur. Finally, at a high, conceptual level, the authors have generated guidelines for the development of cognitive interventions and have proposed a framework for the development of medical error reporting systems that over time can monitor and provide solutions or enhance prevention of the kinds of errors that are reported—which may require collection of additional data elements at the time of error or adverse event detection.

5.2. Comprehension of natural language text information

Another major cognitive framework necessary for the study of medical errors is comprehension and natural language processing of text information. A narrative text reflects natural prose, which is typically the way to express ones observations and thoughts. It does not guarantee, however, that the reader would understand this information exactly as the author meant it [95,96]. Natural language text understanding has two central and complementary component processes — production and interpretation. Both of these can be analyzed in terms of natural linguistic structures (e.g., syntax), propositions, concepts and discourse dynamics [97]. Information retrieval from such text has two steps. First, the user has to search through the text, which is guided by its internal structure, to select the section that might contain the relevant information. Second, the user has to read the content of this section to retrieve the information needed [98]. Structured text, in contrast, is a format that has an inherent order to it and is highly organized.

Patel and colleagues [99,100] point out that processing natural language text requires a distinction between the 'text base' and 'situation model'. A 'text base' is a propositional representation of the meaning of the text itself. The reader generates meaning from the text by transforming the written information into some semantic form or conceptual message. In similar context, understanding may be regarded as a process whereby readers attempt to infer the knowledge structure of a writer through the text, based on what they already know. Thus the mental model developed by an individual who is reading a text is not limited to the information contained in the text itself but is extended to incorporate the reader's prior knowledge. In this sense, the reader constructs a 'situation model' of the scenario described in the text. It is, then, from the interaction between the text-base and the situation model that the conceptual representation emerges [100]. This representation varies greatly from reader to reader as prior knowledge and experiences (i.e., expertise) differ.

One example of such application is the paper by Sharda and colleagues [101], who applied the theoretical concepts of variation in format and expertise to a study of discharge summaries. The doctor-patient interaction generated patient chart was viewed mostly as the 'text base' since the chart represents the clinical information gathered from and provided by the patient. The

discharge summary created from the data interpretation of the physician, reflects the representation of biomedical and clinical information originating from the physicians experience, and was equated to the 'situation model'.

In the field of comprehension, propositional analysis is one of the formal cognitive methods for investigating representation of meaning in memory. Details of the use of such methods are given in Patel et al. [100], where the focus is on understanding and the use clinical practice guidelines. Propositional analysis provides the means to identify concepts and relation among these concepts, generating network structures at various levels of granularity. Text-based and situation-based structures to evaluate the level of understanding can be identified through these networks.

The recall and the generation of inferences in reading a text can be coded and is linked to the understandability and the coherence of a text [99]. The amount of prior knowledge (expertise) in reading a text is also related to the process of inference generation and comprehensibility. The detail of the analysis detects and captures the nature of conceptual and procedural errors and problems in understanding this text, such as discharge summaries or data gathering in EMRs [100].

This method of detecting adverse effect through investigating cognitive process in comprehension equally applies to the reasoning by lay people, such as patients. For example, a study reported by Patel and colleagues on errors in interpreting quantities as procedures in the case of pharmaceutical labels addresses the case in point. Specifically, the correct interpretation of these labels demands that the reader translate minimal, quantitative formulas into qualitative, and often complex, procedures [102]. Given that medication errors involving the use of therapeutic drugs are amongst the most frequent [1], an understanding of the way in which the users of such drugs interpret these labels when administering therapy, is a key step in reducing errors

In this investigation of comprehension of instructions on pharmaceutical labels by people of diverse cultural and education backgrounds, participants were asked to read and interpret (using think aloud methods) pharmaceutical labels related to children's medications of varying complexity: (1) oral rehydration therapy (ORT); (2) over-the-counter cough medicine; and (3) over-the-counter fever medicine. Results show that all groups of participants had considerable difficulty in interpreting the instructions, generating errors of overdose or underdose. The errors of comprehension were attributed to three features of the therapeutic situation the labels presupposed: *the uniformity of the application procedure, the complexity of the quantified variables, and the congruency with intuitive models of therapy*. Uniformity of application was violated when there are irregular intervals or varying amounts of medication between doses. Complexity of quantified variables took the form of inherently difficult conversions, such as converting milligrams to milliliters, or too many calculations. Congruency with intuition was violated when procedures or their instruments must be applied in non-standard ways, such as when the recommended frequency of administration exceeded the reader's intuitive representation of the application situation. Instructions to the public could be monitored knowing the basic nature of the problems lay people have in understanding health instructions.

5.3. Usability inspection and usability testing

A third general cognitive method to monitor and capture errors is related to the evaluation of human and machine competencies as they interact. These methods are called usability inspection and usability testing. Kaufman's paper in this issue [103] assesses use of diabetic home monitoring telemedicine system by patients in their homes. The approach incorporates a cognitive walkthrough usability evaluation and field usability testing conducted in patient's homes. The focus was both on dimensions of the interface and on dimension of patient skills and competency. The usability field research involved testing 25 patients in their homes using the system. The analysis included a range of video-analytic methods of varying levels of granularity. The usability evaluation revealed aspects of the interface that was sub optimal and may impede the performance of certain tasks. It also found a range of patient-related factors such as numeracy and psychomotor skills that constitute barriers to productive use. This addresses a serious gap in knowledge regarding the use of technology by an elderly chronic-care patient population and more generally, for understanding how home health initiatives can more effectively use such technology. In another study, Patel and colleagues 99 identify the usability problem in a commercially available EMR, where the temporal sequence of events in patient data is not captured, leading to generating diagnostic errors by less experience physicians.

Zhang and colleagues [104] used usability heuristics methods to evaluate the safety violations of infusion pumps in the hospitals and also attempted to modify the method for more accurate monitoring of adverse events. The modified heuristic evaluation method was successfully applied to medical devices. 192 heuristic violations were categorized for 89 usability problems identified for Pump 1, and 121 heuristic violations were categorized for the 52 usability problems identified for Pump 2. Pump 1 had more usability problems with high severity ratings than Pump 2. In general, Pump 1 was found to have more usability issues that are likely to induce more medical errors. Heuristic evaluation, when modified for medical devices, was found to be a useful, efficient, and

low cost method for evaluating patient safety features of medical devices through the identification of usability problems and their severities.

6. Conclusions

Errors and adverse events can be detected in many ways. Spontaneous reporting—the traditional cornerstone of error and adverse event detection—can provide some useful information, but identifies only the tip of the safety iceberg. While manual review methods have helped to advance our understanding of adverse events, these methods are too impractical for routine use outside of the context of large, well-funded research projects, with the exception of direct observation for identification of medication error rates. Information technology, in contrast, has the potential to make adverse event detection possible in clinical care. Costs and person hours can be significantly reduced. However, not all adverse events can be detected using coded data. Improvements in techniques to extract concepts from clinical narratives should open up novel methods for adverse event detection. With the combination of multiple data sources it might be possible to detect a wide range of adverse events without reliance on manual review. Cognitive frameworks for error-related research will provide insight into the process of error generation, which will not only aid in error detection, but also more importantly inform the development of interventions. Too often, error and adverse event detection systems have not been translated into effective solutions. Increase use of cognitive and systems methods in medical error and adverse event detection systems that could result in major safety benefits. Improvements in the methods of adverse event detection will also allow researchers the means to evaluate the nature of improvement in patient safety associated with various interventions, which has typically been very costly because of the infrequency of adverse events.

References

[1] Kohn LT, Corrigan JM, Donaldson MS, McKay T, Pike KC. To err is human. Washington, DC: National Academy Press; 2000.

[2] Weingart SN, Wilson RM, Gibberd RW, Harrison B. Epidemiology of medical error. Bmj 2000;320(7237):774–7.

[3] Brennan TA, Leape LL, Laird NM, et al. Incidence of adverse events and negligence in hospitalized patients. Results of the Harvard Medical Practice Study I. N Engl J Med 1991;324(6):370–6.

[4] Thomas EJ, Studdert DM, Burstin HR, et al. Incidence and types of adverse events and negligent care in Utah and Colorado. Med Care 2000;38(3):261–71.

[5] Mills DH. Medical insurance feasibility study. A technical summary. West J Med 1978;128(4):360–5.

[6] Wilson RM, Runciman WB, Gibberd RW, Harrison BT, Newby L, Hamilton JD. The Quality in Australian Health Care Study. Med J Aust 1995;163(9):458–71.

[7] Hutchinson TA, Flegel KM, Kramer MS, Leduc DG, Kong HH. Frequency, severity and risk factors for adverse drug reactions in adult out-patients: a prospective study. J Chronic Dis 1986;39(7):533–42.

[8] Hanlon JT, Schmader KE, Koronkowski MJ, et al. Adverse drug events in high risk older outpatients. J Am Geriatr Soc 1997;45(8):945–8.

[9] Gandhi TK, Burstin HR, Cook EF, et al. Drug complications in outpatients. J Gen Intern Med 2000;15(3):149–54.

[10] Honigman B, Lee J, Rothschild J, et al. Using computerized data to identify adverse drug events in outpatients. J Am Med Inf Assoc 2001;8(3):254–66.

[11] Forster AJ, Murff HJ, Peterson JF, Gandhi TK, Bates DW. The incidence and severity of adverse events affecting patients after discharge from the hospital. Ann Intern Med 2003;138(3):161–7.

[12] Corrigan JM, Kohn LT, Donaldson MS, Maguire SK, Pike KC. Crossing the quality chasm: a new health system for the 21st century. Washington, DC: National Academy Press; 2001.

[13] Shojania KG, Duncan BW, McDonald K, Wachterm RM. Making HealthCare Safe: A Critical Analysis of Patient Safety Practices: Evidence Report/Technical Assessment #43; 2001.

[14] Gandhi TK, Seger DL, Bates DW. Identifying drug safety issues: from research to practice. Intern J Qual Health Care 2000;12(1):69–76.

[15] Bates DW, Boyle DL, Vander Vliet MB, Schneider J, Leape L. Relationship between medication errors and adverse drug events. J Gen Intern Med 1995;10(4):199–205.

[16] Barach P, Small SD. Reporting and preventing medical mishaps: lessons from non-medical near miss reporting systems. Bmj 2000;320(7237):759–63.

[17] Helmreich RL. On error management: lessons from aviation. Bmj 2000;320(7237):781–5.

[18] Reason J. Managing the risks of organizational accidents. Aldershot: Ashgate; 1997.

[19] Wu AW, Pronovost P, Morlock L. ICU incident reporting systems. J Crit Care 2002;17(2):86–94.

[20] O'Neil AC, Petersen LA, Cook EF, Bates DW, Lee TH, Brennan TA. Physician reporting compared with medical-record review to identify adverse medical events. Ann Intern Med 1993;119(5):370–6.

[21] Jha AK, Kuperman GJ, Teich JM, et al. Identifying adverse drug events: development of a computer-based monitor and comparison with chart review and stimulated voluntary report. J Am Med Inf Assoc 1998;5(3):305–14.

[22] Pietro DA, Shyavitz LJ, Smith RA, Auerbach BS. Detecting and reporting medical errors: why the dilemma? Bmj 2000;320(7237):794–6.

[23] Leape LL. Reporting of adverse events. N Engl J Med 2002;347(20):1633–8.

[24] Cullen DJ, Bates DW, Small SD, Cooper JB, Nemeskal AR, Leape LL. 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.

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

[26] Gostin L. A public health approach to reducing error: medical malpractice as a barrier. JAMA 2000;283(13):1742–3.

[27] 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.

[28] Welsh CH, Pedot R, Anderson RJ. Use of morning report to enhance adverse event detection. J Gen Intern Med 1996;11(8):454–60.

[29] Weingart SN, Callanan LD, Ship AN, Aronson MD. A physician-based voluntary reporting system for adverse events and medical errors. J Gen Intern Med 2001;16(12):809–14.
[30] Hiatt HH, Barnes BA, Brennan TA, et al. A study of medical injury and medical malpractice. N Engl J Med 1989;321(7):480–4.
[31] Brennan TA, Localio RJ, Laird NL. Reliability and validity of judgments concerning adverse events suffered by hospitalized patients. Med Care 1989;27(12):1148–58.
[32] Localio AR, Weaver SL, Landis JR, et al. Identifying adverse events caused by medical care: degree of physician agreement in a retrospective chart review. Ann Intern Med 1996;125(6):457–64.
[33] Thomas EJ, Lipsitz SR, Studdert DM, Brennan TA. The reliability of medical record review for estimating adverse event rates. Ann Intern Med 2002;136(11):812–6.
[34] Hayward RA, Hofer TP. Estimating hospital deaths due to medical errors: preventability is in the eye of the reviewer. JAMA 2001;286(4):415–20.
[35] Wolff AM. Limited adverse occurrence screening. A medical quality control system for medium sized hospitals. Med J Aust 1992;156(7):449–52.
[36] Donchin Y, Gopher D, Olin M, et al. A look into the nature and causes of human errors in the intensive care unit. Crit Care Med 1995;23(2):294–300.
[37] Andrews LB, Stocking C, Krizek T, et al. An alternative strategy for studying adverse events in medical care. Lancet 1997;349(9048):309–13.
[38] Flynn EA, Barker KN, Pepper GA, Bates DW, Mikeal RL. Comparison of methods for detecting medication errors in 36 hospitals and skilled-nursing facilities. Am J Health Syst Pharm 2002;59(5):436–46.
[39] Allan EL, Barker KN. Fundamentals of medication error research. Am J Hosp Pharm 1990;47(3):555–71.
[40] Blendon RJ, DesRoches CM, Brodie M, et al. Views of practicing physicians and the public on medical errors. N Engl J Med 2002;347(24):1933–40.
[41] Davis KSSC, Collins KS, Tenney K., Hughes DL, Audet AJ. Room for improvement: patients reports on the quality of their health care. The Commonwealth Fund; 2002.
[42] Bates DWE RS, Murff HJ, Stetson PD, Pizziferri L, Hripcsak G. Detecting adverse events using information technology. J Am Med Inf Assoc 2003;10(2).
[43] Bates DW, Cohen M, Leape LL, Overhage JM, Shabot MM, Sheridan T. Reducing the frequency of errors in medicine using information technology. J Am Med Inf Assoc 2001;8(4):299–308.
[44] Bates DW, O'Neil AC, Petersen LA, Lee TH, Brennan TA. Evaluation of screening criteria for adverse events in medical patients. Med Care 1995;33(5):452–62.
[45] Evans RL, Connis RT. Risk screening for adverse outcomes in subacute care. Psychol Rep 1996;78(3 Pt 1):1043–8.
[46] Goldman R. Development of a Veterans Administration occurrence screening program. QRB Qual Rev Bull 1989;15(10):315–9.
[47] Iezzoni LI, Daley J, Heeren T, et al. Identifying complications of care using administrative data. Med Care 1994;32(7):700–15.
[48] Iezzoni LI, Foley SM, Heeren T, et al. A method for screening the quality of hospital care using administrative data: preliminary validation results. QRB Qual Rev Bull 1992;18(11):361–71.
[49] Lawthers AG, McCarthy EP, Davis RB, Peterson LE, Palmer RH, Iezzoni LI. Identification of in-hospital complications from claims data. Is it valid? Med Care 2000;38(8):785–95.
[50] Needleman J, Buerhaus P, Mattke S, Stewart M, Zelevinsky K. Nurse-staffing levels and the quality of care in hospitals. N Engl J Med 2002;346(22):1715–22.
[51] University HealthSystem Consortium. http://www.uhc.org/.
[52] Leatherman S, Peterson E, Heinen L, Quam L. Quality screening and management using claims data in a managed care setting. QRB Qual Rev Bull 1991;17(11):349–59.
[53] Honigman B, Light P, Pulling RM, Bates DW. A computerized method for identifying incidents associated with adverse drug events in outpatients. Intern J Med Inf 2001;61(1):21–32.
[54] Jollis JG, Ancukiewicz M, DeLong ER, Pryor DB, Muhlbaier LH, Mark DB. Discordance of databases designed for claims payment versus clinical information systems. Implications for outcomes research. Ann Intern Med 1993;119(8):844–50.
[55] McCarthy EP, Iezzoni LI, Davis RB, et al. Does clinical evidence support ICD-9-CM diagnosis coding of complications? Med Care 2000;38(8):868–76.
[56] Murff HJ, Peterson FA, Fiskio JF, Heiman JM, Bates HL. Electronically screening discharge summaries for adverse medical events. J Gen Intern Med 2002;17(Supp 1):A205.
[57] Murff H, Forster J, Peterson AJ, Fiskio JF, Heiman JM, Bates HL. Electronically Screening Discharge Summaries for Adverse Medical Events. J Am Med Inf Assoc 2003;10(4):339–50.
[58] Hripcsak G, Friedman C, Alderson PO, DuMouchel W, Johnson SB, Clayton PD. Unlocking clinical data from narrative reports: a study of natural language processing. Ann Intern Med 1995;122(9):681–8.
[59] Friedman C, Hripcsak G. Natural language processing and its future in medicine. Acad Med 1999;74(8):890–5.
[60] Friedman C. Towards a comprehensive medical language processing system: methods and issues. Proc AMIA Annu Fall Symp 1997:595–9.
[61] Fiszman M, Chapman WW, Aronsky D, Evans RS, Haug PJ. Automatic detection of acute bacterial pneumonia from chest X-ray reports. J Am Med Inf Assoc 2000;7(6):593–604.
[62] Hripcsak G, Austin JH, Alderson PO, Friedman C. Use of natural language processing to translate clinical information from a database of 889,921 chest radiographic reports. Radiology 2002;224(1):157–63.
[63] Evans RS, Gardner RM, Bush AR, et al. Development of a computerized infectious disease monitor (CIDM). Comput Biomed Res 1985;18(2):103–13.
[64] Evans RS, Larsen RA, Burke JP, et al. Computer surveillance of hospital-acquired infections and antibiotic use. JAMA 1986;256(8):1007–11.
[65] Dessau RB, Steenberg P. Computerized surveillance in clinical microbiology with time series analysis. J Clin Microbiol 1993;31(4):857–60.
[66] Hirschhorn LR, Currier JS, Platt R. Electronic surveillance of antibiotic exposure and coded discharge diagnoses as indicators of postoperative infection and other quality assurance measures. Infect Control Hosp Epidemiol 1993;14(1):21–8.
[67] Pittet D, Safran E, Harbarth S, et al. Automatic alerts for methicillin-resistant Staphylococcus aureus surveillance and control: role of a hospital information system. Infect Control Hosp Epidemiol 1996;17(8):496–502.
[68] Rocha BH, Christenson JC, Pavia A, Evans RS, Gardner RM. Computerized detection of nosocomial infections in newborns. Proc Annu Symp Comput Appl Med Care 1994:684–8.
[69] Evans RS, Pestotnik SL, Classen DC, et al. A computer-assisted management program for antibiotics and other antiinfective agents. N Engl J Med 1998;338(4):232–8.
[70] Classen DC, Pestotnik SL, Evans RS, Lloyd JF, Burke JP. Adverse drug events in hospitalized patients. Excess length of stay, extra costs, and attributable mortality [see comments]. JAMA 1997;277(4):301–6.
[71] Bates DW, Leape LL, Cullen DJ, et al. Effect of computerized physician order entry and a team intervention on prevention of serious medication errors. JAMA 1998;280(15):1311–6.
[72] Barker KN, Flynn EA, Pepper GA. Observation method of detecting medication errors. Am J Health Syst Pharm 2002;59(23):2314–6.

[73] Grasela TH, Walawander CA, Kennedy DL, Jolson HM. Capability of hospital computer systems in performing drug-use evaluations and adverse drug event monitoring. Am J Hosp Pharm 1993;50(9):1889–95.
[74] Bates DW, O'Neil AC, Boyle D, et al. Potential identifiability and preventability of adverse events using information systems. J Am Med Inf Assoc 1994;1(5):404–11.
[75] Dalton-Bunnow MF, Halvachs FJ. Computer-assisted use of tracer antidote drugs to increase detection of adverse drug reactions: a retrospective and concurrent trial. Hosp Pharm 1993; 28(8):746–9, 752-5.
[76] Tse CS, Madura AJ. An adverse drug reaction reporting program in a community hospital. QRB Qual Rev Bull 1988;14(11):336–40.
[77] Levy M, Azaz-Livshits T, Sadan B, Shalit M, Geisslinger G, Brune K. Computerized surveillance of adverse drug reactions in hospital: implementation. Eur J Clin Pharmacol 1999;54(11): 887–92.
[78] Dormann H, Muth-Selbach U, Krebs S, et al. Incidence and costs of adverse drug reactions during hospitalisation: computerised monitoring versus stimulated spontaneous reporting. Drug Saf 2000;22(2):161–8.
[79] Bagheri H, Michel F, Lapeyre-Mestre M, et al. Detection and incidence of drug-induced liver injuries in hospital: a prospective analysis from laboratory signals. Br J Clin Pharmacol 2000;50(5):479–84.
[80] Whipple JK, Quebbeman EJ, Lewis KS, Gaughan LM, Gallup EL, Ausman RK. Identification of patient-controlled analgesia overdoses in hospitalized patients: a computerized method of monitoring adverse events. Ann Pharmacother 1994;28(5):655–8.
[81] Classen DC, Pestotnik SL, Evans RS, Burke JP. Description of a computerized adverse drug event monitor using a hospital information system. Hosp Pharm 1992;27(9):774, 776-9, 783.
[82] Classen DC, Pestotnik SL, Evans RS, Burke JP. Computerized surveillance of adverse drug events in hospital patients. JAMA 1991;266(20):2847–51.
[83] Raschke RA, Gollihare B, Wunderlich TA, Guidry JR, Leibowitz AI, Peirce JC, et al. A computer alert system to prevent injury from adverse drug events: development and evaluation in a community teaching hospital. JAMA 1998;280(15):1317–20.
[84] Brown S, Black K, Mrochek S, et al. RADARx: Recognizing, Assessing, and Documenting Adverse Rx events. Proc AMIA Symp 2000:101–5.
[85] Payne TH, Savarino J, Marshall R, Hoey CT. Use of a clinical event monitor to prevent and detect medication errors. Proc AMIA Symp 2000:640–4.
[86] Benson M, Junger A, Michel A, et al. Comparison of manual and automated documentation of adverse events with an Anesthesia Information Management System (AIMS). Stud Health Technol Inf 2000;77:925–9.
[87] DeRosier J, Stalhandske E, Bagian JP, Nudell T. Using health care Failure Mode and Effect Analysis: the VA National Center for Patient Safety's prospective risk analysis system. Jt Comm J Qual Improv 2002;28(5):248–67.
[88] Burgmeier J. Failure mode and effect analysis: an application in reducing risk in blood transfusion. Jt Comm J Qual Improv 2002;28(6):331–9.
[89] Reason J. Human error: models and management. Bmj 2000;320(7237):768–70.
[90] Norman D. The design of everyday things. New York: Doubleday; 1988.
[91] Norman DA. Cognitive engineering. Hillsdale, NJ: Lawrence Erlbaum Associates; 1986.
[92] Zhang J, Patel VL, Johnson TR, Shortliffe EH. Toward a cognitive taxonomy of medical errors. Proc AMIA Symp 2002:934–8.
[93] Zhang J, Patel VL, Johnson TR. Medical error: Is the solution medical or cognitive? JAMIA 2002;9(6 Suppl):S75–7.
[94] Patel VL, Shortliffe EH, Zhang J, Freed M. Slips, mistakes and faulty reasoning: The nature of human errors in medicine. Proc AMIA Symp 2000.
[95] Kintsch W. Comprehension: a paradigm for cognition. Cambridge: Cambridge University Press; 1988.
[96] Kintsch W. The role of knowledge in discourse comprehension: A construction-integration model. Psychol Rev 1988;95(2):163–82.
[97] Frederiksen CH. Representing logical and sematic structure of knowledge aquired from discourse. Cognit Psychol 1975;7(371–458).
[98] Tange HJ, Schouten HC, Kester ADM, Hasman A. The granularity of medical narratives and its effect on the speed and completeness of information retrieval. JAMIA 1998;5 (571-582).
[99] Patel VL, Kushniruk AW, Yang S, Yale JF. Impact of a computerized patient record system on medical data collection, organization and reasoning. JAMIA 2000;7(6):569–85.
[100] Patel VL, Arocha JF, Diermeier M, Greenes RA, Shortliffe EH. Methods of cognitive analysis to support the design and evaluation of biomedical systems: the case of clinical practice guidelines. J Biomed Inf 2001;34:52–66.
[101] Sharda P, Das A, Patel VL. Specifying design criteria for electronic medical record interface using cognitive framework. Proc AMIA Symp 2003:In press.
[102] Patel VL, Branch T, Arocha JF. Errors in Interpreting Quantities as Procedures: The Case of Pharmaceutical Labels. Intern J Med Inf 2002;65:193–211.
[103] Kaufman DR, Patel VL, Starren J, Morin PC, Hilliman C, Pevzner J, Weinstock RS, Goland R, Shea S. Usability in the real world: Assessing medical information technologies in patients' homes. J Biomed Inf 2003;36(1–2):In Press.
[104] Zhang J, Johnson TR, Patel VL, Paige DL, Kubose T. Using usability heurisitics to evaluate patient safety of medical devices. J Biomed Inf 2003;36(1–2):In Press.