Final Progress Report

Validation of an Innovative Approach To Error Reduction

Principal Investigator: David Blumenthal, MD, MPP

Team Members:
Carlos A. Camargo, Jr., MD, DrPH (Massachusetts General Hospital)
Paul D. Cleary, PhD (Yale School of Public Health)
James A. Gordon, MD, MPA (Massachusetts General Hospital)
Edward Guadagnoli, PhD (Harvard Medical School)
Rainu Kaushal, MD, MPH (New York Presbyterian Hospital)
David J. Magid, MD, MPH (Colorado Permanente Medical Group)
Sowmya R. Rao, PhD (Massachusetts General Hospital)
Ashley F. Sullivan, MS, MPH (Massachusetts General Hospital)

Project Dates: 09/30/2003 - 09/29/2006

Federal Project Officer: Kerm Henriksen, Program Official

Supported by a grant from the Agency for Healthcare Research and Quality
Grant Number: 5 R01 HS013099

STRUCTURED ABSTRACT

Purpose: To determine if reports by emergency department (ED) personnel about failures in clinical safety processes and the presence of potential systemic contributors to errors are significantly correlated with the actual occurrence of errors in EDs and (2) to examine the characteristics of EDs associated with the occurrence of errors.

Scope: No comprehensive studies of the frequency or types of errors in EDs exist. Previous studies suggest that EDs are a common site of occurrence for adverse events. This study measured errors in a large sample of EDs throughout the U.S.

Methods: Staff perceptions were assessed with a validated survey, and errors were assessed through chart review of three conditions: acute myocardial infarction, acute asthma, and dislocations that use procedural sedation. ED characteristics were collected through a key informant survey.

Results: Data collection for the study has concluded; databases are currently being cleaned, analytic variables are being constructed, and exploratory analyses are underway. The principal results of the study are forthcoming. We are optimistic that our data will provide a landmark contribution to understanding of safety of EDs, approaches to measuring safety in EDs and elsewhere, and the extent of guideline compliance in EDs.

Key Words: emergency department, patient safety, errors

PURPOSE

The objectives of the study are to (1) determine if reports by emergency department (ED) personnel about failures in clinical safety processes and the presence of potential systemic contributors to errors are significantly correlated with the actual occurrence of errors in EDs and (2) examine the characteristics of EDs associated with the occurrence of errors, which will contribute to understanding and reducing the occurrence of errors in EDs.

SCOPE

The significance of the medical error problem is now well known. It is important to reduce the frequency of errors as rapidly and cost effectively as possible; however, the identification of unsafe healthcare processes poses analytic and logistical challenges. Moreover, the development of systems for the detection of preventable adverse events is expensive and difficult. These challenges suggest the value of developing tools for direct and quick identification of unsafe clinical processes and associated systemic factors.

This study responds to the need for error reduction methods, with a focus on the correction of failures in emergency healthcare. Using the ED as the study environment, the overall aim of the study is to determine if reports by personnel about safety processes are significantly correlated with the occurrence of errors. If so, such a reporting system can be used to accurately identify processes for continuous quality improvement. In addition, data collection efforts will provide potentially valuable new information on the frequency and types of errors, and the characteristics of EDs associated with occurrence of errors.

No comprehensive studies of the frequency or types of errors in EDs exist. Previous studies in New York focused on the hospital experience generally; however, their work revealed that 2.9% of adverse events in hospitalized patients occurred in EDs. This made EDs the third most common site of occurrence for adverse events, after operating rooms and patient care rooms. Negligent adverse events (defined as adverse events resulting from "care that fell below community standards") were more common in EDs (70.4% of adverse events) than in any other settings. A similar study in Utah and Colorado found that EDs accounted for 1.7% of adverse events in hospitalized patients and that 52.6% were found to be negligent.

There is reason to believe that these studies underestimated the incidence of errors in EDs. Because adverse events and negligent adverse events were detected through reviewing charts of hospitalized patients, problems that might have occurred among patients seen in the ED and discharged without admission were not counted in the study. Other studies indicate that suboptimal care is reasonably frequent in EDs. In a study of emergency departments at five Harvard hospitals, Burstin et al found that care conformed with process of care guidelines in 59.1% of shortness of breath cases and 65.3% of chest pain episodes. One of the contributions of this research will be to provide the most comprehensive portrait of the number and type of ED safety problems ever developed.

METHODS

We measured the perceptions of staff about potentially unsafe processes and the actual occurrence of errors in EDs. Study implementation first required identification of a sample of 85 EDs. Next, we began actual data collection, which included (1) administration of a survey to personnel in these EDs; (2) data collection on rates of errors through chart review; and (3) collection of other relevant site-specific data. The Institutional Review Board at all participating institutions approved the study.

Data Sources and Site Sample

The 85 EDs recruited for data collection consisted largely of sites affiliated with the Emergency Medicine Network (EMNet), an ED-based research collaboration. We excluded military and VA hospitals as well as hospitals in U.S. territories. Children's hospitals also were excluded, because acute myocardial infarction (AMI) is one of the study conditions of interest.

Because many EMNet sites are affiliated with an emergency medicine residency program (i.e., are academic EDs), we made an extra effort to recruit non-academic, non-metropolitan EDs to increase the study's generalizability. To accomplish this goal, we created the 2001 National ED Inventory. We then selected EDs with annual visit volumes between 28,000 and 45,000. We excluded EDs with <28,000 annual visits, because this is the median annual visit volume of EDs that see at least one patient per hour, and we expected that EDs with <28,000 annual visits would not see the volume of cases needed for the chart review component of the study.

Academic EDs have a median annual visit volume of 48,920, so we used the cutoff of 45,000 visits to capture EDs more likely to be non-academic. We established a criterion of one published article as an indicator of an ED's potential interest in a research collaboration. We performed a Medline search for all hospitals with annual ED visit volumes between 28,000 and 45,000 to determine if any member of the ED published an article from 1996 to April 2004. Because <10 non-metropolitan hospitals had publications, we expanded the Medline search to hospitals in metropolitan statistical areas. Community hospitals in both non-metropolitan and metropolitan areas, with at least one publication, were invited to participate in the study (n=18). Unfortunately, none of these hospitals participated.

Overall, of the 241 sites invited to participate in the study, 102 agreed to participate, 49 declined the invitation, and 90 did not respond to the invitation. Over the course of the study, 32 sites dropped out. Sites withdrew from the study because the site principal investigator or key research personnel left the site, the site's IRB or management prohibited the site from participating in the research, or the site had inadequate administrative or research support to complete the study.

Survey

All participating EDs administered the survey to members of their staff. To develop the survey, the study team revised a previously developed instrument that was based on a human factors framework. Though the survey already had been piloted in two institutions, it had not been subjected to psychometric testing and needed additional work to test its use in the ED setting. We also added questions to assess specific ED process failures that might contribute to errors. Items in each domain probed whether principles of human factors were applied in the ED. For example, within the staffing domain, respondents were asked whether staffing is sufficient to handle the patient care load during busy periods.

To further refine the survey, investigators conducted confidential, in-depth personal interviews with key informants in three EDs. Key informants included ED medical directors, nurse managers, physicians, nurses, and administrators. Interviews followed a structured protocol that covered 1) specific clinical processes that interviewees observed to be associated with medical errors in ED care and 2) systemic factors that may be generally associated with that occurrence of errors in the ED. Interviews were tape recorded and transcribed.

Investigators also conducted focus groups at these three EDs. The focus groups lasted approximately 2 hours and followed the same protocol as the structured interviews. In addition, the survey underwent cognitive testing within the focus groups. Participants included ED physicians, nurses, administrators, and other ED staff to ensure that a variety of perspectives were represented. Focus groups also were recorded and transcribed for review by investigators.

Ten EDs served as sites for psychometric testing. We administered a paper-based version of the survey to all eligible ED staff at these 10 sites. Data from these sites were used to establish the psychometric properties of the survey, with particular attention to clustering of systemic factors of interest.

Based on preliminary factor analyses, the investigators deleted certain questions from the survey and analyzed substantively coherent clusters of items. Decisions about which items to drop also were based on face validity. Items then were organized into pre-specified domains, scales were developed (using the items expected to represent the domains), and reliability statistics were calculated. In some cases, when the domain of a variable was ambiguous, the variable was tested in more than one scale. The process resulted in a revised survey with nine psychometrically coherent domains: physical environment, equipment, triage & monitoring, staffing, nursing, teamwork, culture, information coordination, and inpatient coordination. On review of these data, we were able to decrease the survey length by approximately 20%.

We administered the final survey to a random sample of 80 ED staff at each of the 75 remaining study sites. Sites with fewer that 80 eligible staff administered the survey to all eligible staff. Potential respondents were informed of their right not to fill out the surveys and of the measures taken to ensure their confidentiality. Informed consent was implied by completion of the survey.

The final survey was administered to ED staff who worked at least one shift per week and provided clinical care. Eligible survey respondents were clinical ED staff employed in the ED for at least 3 months, with the exception of residents, who needed to work in the ED for at least 1 month.

The survey asked about working conditions and clinical care in the ED, with a focus on the integrity of certain generic processes that are important to the safety of ED patients regardless of diagnosis. The instrument also asked perceptions of ED personnel about the presence of systemic factors consistent with safety. The survey collected perceptions about all the domains identified in psychometric testing, including equipment, staffing, nursing, information coordination and inpatient coordination. In addition, the survey collected perceptions about the management of AMI, asthma, and dislocations that use procedural sedation (i.e., the three conditions chosen for chart review). Respondents replied to statements using a five-point Likert scale. Respondents were also asked to provide personal background information, including their position and length of employment in that ED.

Although the distributed surveys were a paper-based instrument, staff had the option to complete the revised survey online. Site coordinators distributed surveys, but respondents returned completed surveys directly to the EMNet Coordinating Center at Massachusetts General Hospital. Non-respondents received two additional surveys at 2-week intervals for a total of three surveys over 6 weeks. To improve response rates, the site stipend included funds for a modest honorarium for survey respondents. The choice of honorarium was at the discretion of the site (e.g., cash, gift cards for a local coffee shop, or application of these funds to an ED-related event or project).

Chart Review

The chart review identified the prevalence and characteristics of errors that occurred in our sample of EDs.

The definition of a medical error was a critical issue for this investigation. Following the work of Reason and the Institute of Medicine (IOM), we defined an *error* as the failure of a planned action to be completed as intended or the use of a wrong plan to achieve an aim. We defined an *adverse event* to be an injury resulting from a medical intervention not due to the underlying condition of the patient. We defined a *preventable adverse event* as an adverse event associated with an error.

For the purpose of this study, we used the failure to comply with standards of care and the identification of preventable adverse events as approaches to measure error and actual harm from errors. Failures to comply with standards of care were further classified as serious and non-serious depending on the potential consequences of the particular failure. Serious failures (also called "critical errors") will be grouped with preventable adverse events in some analyses. Accordingly, we conducted explicit chart review both for (1) failure to comply with condition-specific ED guidelines for our three target conditions and (2) the occurrence of adverse events. A physician review panel (described below) made implicit decisions about the preventability of each adverse event identified by chart review.

To increase the generalizability of the study, we also examined charts for the occurrence of failures in generic processes – that is, processes that are important to the care of all ED patients, regardless of diagnosis (e.g., the failure to record vital signs or prolonged waits prior to treatments).

We chose three clinical conditions (AMI, acute asthma, and dislocations that use procedural sedation) on which to concentrate survey questions about failing clinical processes and on which to focus our chart review to validate reports of failing processes.

Using ICD-9 codes from hospital administrative records, sites identified charts with a principal ED or hospital discharge diagnosis of AMI, asthma, or dislocation. We excluded transfer patients and patients who were ineligible due to any of the following criteria:

AMI
Age $\geq$ 90 years
Cardiac arrest prior to or on arrival to the ED
Cardiac enzymes not elevated within 24 hours or ED arrival, or no cardiac enzyme levels documented or drawn

Asthma
Age $\leq$ 13 or $\geq$ 55 years
History of chronic obstructive pulmonary disease or emphysema
No history of asthma before index visit
ED visit not prompted, in large part, by asthma exacerbation

Dislocation
Age ≤ 13 or ≥ 90 years
No dislocated joint of interest
Acromioclavicular shoulder dislocation
No joint relocation procedure*
No intravenous or intramuscular sedative or anesthetic administered*

* A partial chart abstraction was performed for charts with either of these exclusion criteria. We therefore have some data on charts in which no relocation procedure occurred or no sedative or anesthetic was administered.

When a single patient's chart included >1 ED visit for a given condition during the previous 12 months, we included only the first ED visit in the study.

For each condition, onsite chart abstractors reviewed 70 randomly selected ED charts. If the ED patient was admitted and screened positive for a possible adverse event, chart abstractors could expand their review of documentation beyond the ED chart and review the hospital discharge summary for patients who were admitted to the hospital from the ED. The hospital discharge summary allowed chart abstractors to identify adverse events that may have occurred in the inpatient setting as a result of care given in the ED. Sites with <70 charts in the 12-month period reviewed all eligible charts available for the condition. The 10 psychometric survey testing sites reviewed charts for patients who presented to the ED during the 12 months preceding the administration of the staff survey. The remaining sites reviewed charts from calendar year 2004.

We developed computerized abstraction tools so that abstractors had the option to complete the reviews using either computerized tools or paper forms. The protocol included explicit measures of compliance with standards of care developed by authoritative groups. The chart review instruments also included criteria for measuring the occurrence of adverse events in ED settings. Screening criteria for adverse events involved criteria that were adapted from previous studies. Before starting data collection, the abstraction forms were tested on a sample of charts at four EDs. A few ambiguities were identified and corrections made.

Before initiation of the chart reviews, the Project Director and co-investigators also trained abstractors identified by sites. Chart abstractors were required to have some medical training, and >95% were physicians, nurses, residents, and medical students. We conducted 60- to 90-minute training sessions by telephone employing PowerPoint presentations of review protocols. Following the training session, we mailed abstractors six practice charts, which they abstracted and returned. Practice charts were checked against "criterion standard" abstractions provided by the investigators. Abstractors whose accuracy per chart was <80% or who did not identify important adverse events were re-trained before they were permitted to begin reviewing actual charts.

Electrocardiograms for AMI cases were collected and subsequently interpreted by two board-certified cardiology attendings and two emergency medicine attendings. Interpretation was used to help determine eligibility for reperfusion and to help determine eligibility for ß-blockers.

In addition, accuracy of electrocardiogram interpretation is an outcome in and of itself and a potentially important explanatory factor for errors in delivery of AMI care.

For the detection of preventable adverse events and critical errors, we used a multi-level review. The first review detected the occurrence of adverse events or apparent critical errors and was conducted as part of the initial chart review by onsite abstractors using the screening criteria adapted for the detection of adverse events. Charts that met one of the criteria for an adverse event were de-identified and forwarded to the EMNet Coordinating Center for review by a panel of physician reviewers, who judged the type of error, its preventability, and its impact.

The physician review panel consisted of board-certified or board eligible emergency physicians and physician patient safety experts. All panel members completed a 60-minute training session by telephone. A pair of physicians consisting of at least one emergency physician reviewed each chart that screened positive for an adverse event. Each physician reviewed the chart independently and then discussed the case with their paired colleague to reach consensus on how to judge the case. When reviewers could not reach consensus, a third physician served as a tiebreaker and made a final decision on the classification of the event. Charts were randomly distributed to physician reviewers on a rolling basis. Reviewers were not permitted to review charts from their ED. In cases when a critical error was responsible for a preventable adverse event, the data point was entered into the analysis only as a preventable adverse event to prevent double counting.

All chart abstraction data also were reviewed, independent of whether an adverse event or critical error was identified by the onsite reviewer. This enabled the identification of guideline violations that did not involve critical errors or critical errors that might not have prompted a referral for physician review.

Other Relevant Site-Specific Data

Attributes of EDs may influence either perceptions of errors or actual error rates. To explore these possible relationships, we distributed a key informant survey to collect ED attributes, including volume of ED visits in the past year; number of FTE ED staff during the past year; average numbers of hours of ED divert per month over past year; average patient waiting times over the past year; and the proportion of patients arriving by ambulance. Sixty-nine sites completed the key informant survey.

Measures

We are using bivariate and multivariate analytic techniques to examine the relationship between ED personnel perceptions of process failures (as determined on the survey) and rates of occurrence of medical errors (as determined by chart abstraction). In testing all hypotheses and study questions, we are interested in institution-level inferences and are using hierarchical models as appropriate to derive those inferences.

We are calculating the means, medians, and distributions of key variables. Correlation coefficients will be calculated to relate chart review measures and ED personnel perceptions of process failures across sites.

We will conduct multivariate analyses examining the relationship between dependent variables of interest (e.g., overall compliance rates, condition specific compliance rates, process-specific compliance rates, generic error rates, preventable adverse events, preventable adverse events combined with critical errors) and independent variables (e.g., staff reports of average percent of cases with condition that meet standards, staff perceptions of different systemic factors). Covariates will include attributes of the staff (average time of service in ED, turnover), ED workload (volume of visits/year; visits per FTE), and patient acuity (proportion of patients arriving by ambulance).

We will assess relationships among the institution-level variables in an attempt to identify any collinear groups of variables in constructing models. Statistical assumptions also will be assessed (e.g., normality, heteroscedasticity), and necessary adjustments will be made to make these assumptions more plausible. All analysis will be weighted for the sampling design, non-response and respondent type (e.g., physicians, nurses). We will examine differences in the covariate distributions among the respondents and non-respondents and attempt to reduce any non-response bias.

Limitations

Our sample of EDs consisted predominantly of EDs affiliated with an emergency medicine residency program. However, a sample that represents the nation's teaching hospitals is highly relevant from a policy standpoint, because these institutions train the vast majority of emergency physicians. All the participating EDs were in metropolitan statistical areas, so rural EDs are not represented in this study; however, the majority (72%) of US EDs are in an urban setting.

The chart review at most sites yielded <70 dislocation cases with procedural sedation. Although this possibility was discussed during study planning, we selected this condition because a possible decrease in charts might be offset by a higher error rate. To further mitigate risk, we asked sites to abstract data about dislocation cases without procedural sedation, such as the patient's level of pain and analgesics administered. Quality of care issues will be examined using data gathered on dislocation patients who did not receive procedural sedation.

Some sites did not have 70 eligible AMI and/or asthma cases. Although this limits the available data for the study, we felt it was more important to include EDs that saw a lower volume of visits (e.g., 50 cases) rather than exclude them and decrease the generalizability of the study and statistical power.

Our study may underestimate the incidence of errors in EDs. For example, because adverse events are detected through the review of ED records and hospital discharge summaries, problems that might have resulted in adverse events after discharge from the ED would not have been counted in NEDSS. We tried to address this by including a return to the ED within 48 hours as a potential adverse event that warranted physician review.

RESULTS

Data collection for the study has concluded; databases are currently being cleaned, analytic variables are being constructed, and exploratory analyses are underway. The principal results of the study are not yet available. However, we are optimistic that our data will provide a landmark contribution to understanding of safety of EDs, approaches to measuring safety in EDs and elsewhere, and the extent of guideline compliance in EDs.

Our database consists of:

- Surveys from 3,684 eligible emergency clinicians from 70 sites.
- Chart review data for 10,205 charts (AMI: 3,813 from 60 sites; asthma: 4,049 from 62 sites; dislocations with procedural sedation: 2,213 from 59 sites).
- Dual physician review data for 1,564 charts screening positive for a potential adverse event.
- Key informant surveys from 69 sites.

The following papers/abstracts have been submitted and accepted for publication:

Sullivan AF, Richman IB, Ahn CJ, et al. A profile of U.S. emergency departments in 2001. *Ann Emerg Med* 2006; 48: 694-701.

Sullivan AF, Camargo CA Jr, Cleary PD, et al. Do emergency physicians and nurses differently perceive factors that affect safety? The National ED Safety Study. [abstract] *Acad Emerg Med* 2007; in press.

The following papers are in draft form:

Sullivan AF, Camargo CA Jr, Cleary PD, et al. The National Emergency Department Safety Study (NEDSS): Study rationale and design.

Magid DJ, Rao SR, Sullivan AF, et al. Perceptions of safety among emergency department clinicians.

Kansagra SM, Rao SR, Sullivan AF, et al. Perception of workplace safety and incidents of violence and weapons in the ED.

The following papers are planned at the current time:

1. Do Emergency Physicians and Nurses Differently Perceive Factors That Affect Safety? The National ED Safety Study.

2. The Determinants of Errors in Emergency Departments.

LIST OF PUBLICATIONS and PRODUCTS

Sullivan AF, Richman IB, Ahn CJ, et al. A profile of U.S. emergency departments in 2001. *Ann Emerg Med* 2006; 48: 694-701.

Sullivan AF, Camargo CA Jr, Cleary PD, et al. Do emergency physicians and nurses differently perceive factors that affect safety? The National ED Safety Study. [abstract] *Acad Emerg Med* 2007; in press.