Improved incident reporting following the implementation of a standardized emergency department peer review process

MARTIN A. REZNEK AND BRUCE A. BARTON

Department of Emergency Medicine (MAR) and Department of Quantitative Health Sciences (BAB), University of Massachusetts Medical School, 55 Lake Avenue, Worcester, MA 01655, USA

Address reprint requests to: Martin A Reznek, Department of Emergency Medicine, UMassMemorial Medical Center, 55 Lake Avenue North, Worcester, MA 01655, USA. Tel: +1-508-421-1400; Fax: +1-508-421-1490; E-mail: mreznek@hotmail.com

Accepted for publication 14 March 2014

Abstract

Objective. Incident reporting is an important component of health care quality improvement. The objective of this investigation was to evaluate the effectiveness of an emergency department (ED) peer review process in promoting incident reporting.

Design. An observational, interrupted time-series analysis of health care provider (HCP) incident reporting to the ED during a 30-month study period prior to and following the peer review process implementation and a survey-based assessment of physician perceptions of the peer review process’ educational value and its effectiveness in identifying errors.

Setting. Large, urban, academic ED.

Participants and Interventions. HCPs were invited to participate in a standardized, non-punitive, non-anonymous peer review process that involved analysis and structured discussion of incident reports submitted to ED physician leadership.

Main Outcome Measures. Monthly frequency of incident reporting by HCPs and physician perceptions of the peer review process.

Results. HCPs submitted 314 incident reports to the ED over the study period. Following the intervention, frequency of reporting by HCPs within the hospital increased over time. The frequencies of self-reporting, reporting by other ED practitioners and reporting by non-ED practitioners within the hospital increased compared with a control group of outside HCPs (P = 0.0019, P = 0.0025 and P < 0.0001). Physicians perceived the peer review process to be educational and highly effective in identifying errors.

Conclusions. The implementation of a non-punitive peer review process that provides timely feedback and is perceived as being valuable for error identification and education can lead to increased incident reporting by HCPs.

Keywords: incident reporting, peer review, quality improvement, emergency medicine

Introduction

In 1999, the Institute of Medicine reported that up to 98,000 patients die annually from medical errors in the USA alone spurring health care organizations around the world to focus on quality improvement [1]. Error detection is essential for error reduction because it facilitates investigation and development of preventative measures. Health care provider (HCP) incident reporting can be a valuable source for error detection; however, reporting is often suboptimal [2].

Increasing incident reporting stands to improve patient safety, but how best to achieve increased reporting remains unclear. A limited number of original, experimental studies have shown increases in incident reporting after the introduction of electronic reporting systems [3–5], anonymous reporting systems [6, 7], HCP education [3, 8, 9], inter-disciplinary safety teams [10] and structuring reporting into established clinical work flows [4, 11, 12]. However, the quantity of original research in this area has been limited and the quality of several studies has been questioned as well. A Cochrane Collaboration review in 2012 drew no conclusions about the evidence, having identified only 21 original, experimental studies of interventions to improve incident reporting with only four meeting strict methodology criteria. The authors encouraged future investigators to consider ‘time series’ analyses [13].

In the two studies meeting inclusion criteria in the Cochrane report that showed positive results, the interventions were multi-faceted. The designs differed from each other, but two themes...
were consistent; the interventions were non-punitive and included improved feedback for reporters [14, 15]. While the importance of these components in improving incident reporting has yet to be established definitively via experimental research, it is supported by survey-based research and expert opinion [16–22].

Based on available research and expert opinion, it seems likely that a system can increase HCP reporting if those making reports perceive that the system is safe and likely to result in quality improvements [16–22]. Accordingly, we designed and implemented a standardized emergency department (ED) peer review process in order to increase HCP reporting. The goal of this investigation was to evaluate the effectiveness of this intervention in promoting HCP incident reporting. A secondary goal was to describe the process so that others might consider it as a template for similar efforts at their institutions.

Methods

Study design and setting

A two-part study design was employed. Part A was an observational, interrupted time-series (ITS) analysis of incident reports submitted by HCPs prior to and following the process implementation. Part B was a survey-based assessment of the perceptions of physicians participating in the peer review process. The data collection and study methods were approved by the institution’s Committee for the Protection of Human Subjects.

Data were collected in a teaching hospital ED with an annual census of ~66 000 adult patients between March 2010 and August 2012. Board-certified emergency medicine (EM) physicians cared directly for the majority of patients as the sole provider or through supervision of EM and rotating residents. A limited number of patients (~5%) were seen primarily by mid-level practitioners.

Interventions

Prior to the peer review process intervention, all ED incident reports were investigated via a single-reviewer format. The reviewer had >10 years of experience in quality improvement; however, the single-reviewer format allowed for subjectivity and anecdotal led to inconclusive investigations if the reviewer and involved practitioner(s) disagreed over the findings. The reviews considered both practitioner-based and systems errors, but the process was not standardized leading to concerns that systems errors especially may have been under-identified [23]. Feedback was provided to involved practitioners; however reporter, departmental and hospital-wide feedback were not standardized.

At the beginning of the 8th month of the study, a new, standardized peer review process was implemented with the intention of fostering a non-punitive environment and providing timely and more robust feedback to reporters and staff in general. The process was based on a format previously implemented at Detroit Receiving Hospital (Detroit, MI, USA) with modifications to the system error analysis and feedback processes [24]. Figure 1 outlines the implemented process.

All incident reports were submitted electronically via the hospital incident reporting system or by direct electronic, written or verbal communication with ED leadership and were investigated via the new process. For each report, a screening review was performed by the ED Clinical Director. A single-reviewer format for screening was chosen for practical considerations, but to limit potential bias, cases were closed only if the review determined with absolute certainty that no errors or near misses had occurred. If they could not be excluded with absolute certainty, the case progressed to full Peer Review Committee (PRC) evaluation. Prior to PRC evaluation, ED practitioners involved in the care were required to comment in writing about the care that was provided. This feedback allowed for first-hand accounts of pertinent information that might not be available in the medical record. In addition, involved ED practitioner(s) were encouraged to attend the PRC meeting.

The PRC met monthly and was open to all ED attending physicians, residents and mid-level practitioners; however, the PRC included a core group of eight board-certified, attending physicians who were present regularly to ensure accuracy and consistency. Mean PRC attendance was 10 participants, and 86% of the attendants were board-certified/board-eligible ED attending physicians. During PRC proceedings, all participants reviewed the medical record and practitioner responses. Records were de-identified as were practitioner responses unless the practitioner(s) chose to be present for the committee review. The PRC Chair facilitated an open discussion focused on potential errors. Specifically, the committee reviewed cases for 11 types of systems and practitioner-based errors (Table 1). A majority vote by all participants determined the presence or absence of errors. Providers involved in a case in question were required to leave the room during voting on practitioner-based errors to reduce potential influence.

Following the PRC meeting, the findings were recorded in a spreadsheet database (Microsoft® Excel, Microsoft, Redmond, WA, USA) and feedback was provided as described in Fig. 1. Data tracking was transparent to all ED practitioners. They were regularly reminded that peer review was undertaken only to guide quality improvement and would not be used for punitive purposes. The entire process, including all records and feedback, complied with peer review protection laws meaning the data were shielded from use by outside parties as a part of litigation.

Methods and measurements

Part A

Every incident report received by the ED between March 2010 and August 2012 was recorded in a database with the date of its preliminary review and its source. Reports submitted by HCPs were categorized into four source types: (i) self-reported (Self)—an ED practitioner directly involved in the care of the patient when the perceived incident occurred, (ii) ED—an ED practitioner not directly involved in the care of the patient when the perceived incident occurred, (iii) hospital practitioner (HP)—a non-ED practitioner within the hospital or a hospital committee comprised primarily of non-ED
Figure 1 Flow diagram of the peer review process.
practitioners and (iv) outside practitioner (OP)—a practitioner from outside the hospital. For the purposes of this study, HCPs included physicians, nurses, allied health care workers and post-graduate trainees.

The preliminary review date served as a surrogate for the submission of the incident report because submission dates were not recorded. With one exception, cases were consistently reviewed within a week of receipt, so the preliminary review was a consistent surrogate data point. In a 3-week period, other administrative requirements precluded performing preliminary reviews. This resulted in an artificially low number of reports for that month and an artificially high number for the following month. Consideration for statistical correction to most accurately account for this anomalous period is described in the analysis section.

**Part B**

PRC participants had the opportunity to complete voluntary, anonymous, 'activity assessments' following each PRC meeting.
Offering this survey was required for the meetings to be eligible for continuing medical education (CME) credit (The PRC Chair was not eligible to complete the survey as the director of the activity). Five of the questions assessed participant perceptions of the PRC process itself and were therefore pertinent to this investigation (Table 3). They included one Likert-type question, three yes/no questions and one multiple-choice question.

Analysis

Part A

The primary study outcome was the change in monthly frequency of incident reporting by HCPs. Prospectively collected data from the peer review database were reviewed and descriptive statistics were calculated. The data were imported into SAS (SAS Institute, Cary, NC, USA) to perform an ITS analysis of monthly reporting frequencies for HCP source groups before and after the peer review process intervention. Frequencies for the Self, ED and HP groups were compared with the OP group. The OP group served as a control because while it also contained HCPs, OP providers did not work in the hospital and had minimal exposure to the intervention.

Within the ITS analysis structure, we used SAS Proc Autoreg to test for autocorrelation (with a lag of 2 months) among the count data. For each source group, the Durbin–Watson statistic was close to 2.0 with non-significant P-values, indicating no significant autocorrelation, and the estimated autoregressive parameters (for both 1 and 2 months) were also uniformly non-significant [25]. As a final test of the effect of autocorrelation, we fit the same models using generalized estimating equations (SAS Proc Genmod), which produced similar results, also indicating no significant autocorrelation.

Based on this assessment of autocorrelation, we used segmented general linear models (ordinary least squares) [25] to fit the ITS model for the four groups. The models fit included count of incident reports as the outcome and as predictors: month of study (to obtain the pre-intervention slope), intervention (0/1 dummy variable with a 1 for the months at or after the implementation to obtain the immediate post-intervention effect), group (the OP group was used as the control/reference since they were not exposed to the intervention) and month post-intervention (to obtain the slope of the change in incident report counts after the intervention). In addition to these main effects, interactions of group with the intervention effect and with the post-intervention slope were tested to determine whether the groups differed in either predictor. We used a standard t-test to test each regression parameter for a significant difference from zero. Separate models were fit comparing the groups: Self vs. OP, ED vs. OP and HP vs. OP. An intercept term was included but it was difficult to interpret directly since it was a combination of the OP group at the start of the study period as well as the pre-intervention slopes for both groups. It was included for completeness of the results without further comment.

For reasons explained in the Methods and measurements section, incident reports were artificially low in September 2011 followed by a compensatory spike in October 2011. Given that the results were definite departures from the months before and after and were byproducts of factors outside of the study, a modification to these data points was considered. We performed two sensitivity analyses: (i) using a loess smoother for that time period with a three-period lag (and using the loess-predicted levels of reports in the subsequent ITS models) and (ii) using a term in the model to represent the values in that time period. For the ED group, the loess smoother produced predicted values (5.05 for September and 5.03 for October) that were slightly lower than the mean of the 2 months (6.0). Substituting these into the model yielded almost identical results to substituting the mean ($P = 0.0024$ with the loess-predicted values and $P = 0.0025$ with the mean substituted). In addition, a term for those 2 months was generated and included in the model but was not significant ($P = 0.30$). Finally, we used the original data in the model, with almost the same parameter estimate, but a slightly larger standard error, which reduced the $P$-value of the slope from 0.0025 to 0.0143 (analyses for the Self and HP groups produced similar results). Based on these sensitivity analyses and because we felt that the mean of the 2 months presented a better approximation of the real number of reports, we used the modified data in the analysis.

Part B

PRC participant perceptions of the peer review process’ effectiveness in identifying errors and its educational value were secondary outcomes. Survey results were retrospectively reviewed and entered in a Microsoft® Excel spreadsheet. Descriptive statistics were calculated.

Results

Part A

HCPs submitted 314 incident reports over 30 months (24 prior to the intervention and 290 following). Of note, no self-reports occurred prior to the intervention, but they accounted for 58 after implementation (20%). Table 2 summarizes the number of incident reports by source group prior to and following the intervention.

Figure 2a–c shows the observed monthly incident reporting rates and the predicted, fitted ITS model curves for the Self, ED and HP groups compared with the OP group. The ITS model results for the Self, ED and HP groups compared with the OP group are presented in Tables 3, 4 and 5. Using

<table>
<thead>
<tr>
<th>Source type</th>
<th>Pre-intervention (7 months) (%)</th>
<th>Post-intervention (23 months) (%)</th>
<th>Total (30 months) (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Self</td>
<td>0 (0.0)</td>
<td>58 (20.0)</td>
<td>58 (18.5)</td>
</tr>
<tr>
<td>ED</td>
<td>7 (29.2)</td>
<td>91 (31.4)</td>
<td>98 (31.2)</td>
</tr>
<tr>
<td>HP</td>
<td>14 (58.3)</td>
<td>125 (43.1)</td>
<td>139 (44.3)</td>
</tr>
<tr>
<td>OP</td>
<td>3 (12.5)</td>
<td>16 (5.5)</td>
<td>19 (6.1)</td>
</tr>
</tbody>
</table>
Table 3 as an example, interpretation of the results is as follows: the month pre-intervention slope showed a decrease of 0.13 reports per month in the pre-intervention period that was not significant ($P = 0.5234$); the Intervention term showed an increase of 1.23 reports in the month immediately after the intervention that was not significant ($P = 0.2082$); the Group effect term showed that the ED group had an average of 0.12 reports per month less than the OP group throughout that was not significant ($P = 0.8267$) and the Month Post-Intervention Slopes showed that after the intervention there was an average increase of 0.27 reports per month in the ED group and 0.10 per month in the OP group, with a resulting difference in the slopes of 0.17 that was statistically significant ($P = 0.0019$). This final $P$-value for the difference in slopes is also reported in Fig. 2.

The comparison analyses revealed that prior to the intervention, the frequency of reporting by the Self, ED and HP groups did not differ from the OP group ($P > 0.1$ for each). The frequency of reporting by the Self, ED and HP groups appeared to increase immediately following the intervention; however, the immediate effects did not achieve statistical significance ($P > 0.10$, $P > 0.10$ and $P = 0.0902$). Long-term effects following the intervention, however, were observed in all three groups; the Self, ED and HP groups showed a greater...
rate of change in reporting compared with the OP group ($P = 0.0019, P = 0.0025$ and $P < 0.0001$).

**Part B**

In 23 months following the process implementation, 22 PRC meetings occurred with a total attendance of 221. The mean attendance was 10.0 (range: 5–17, SD: 2.9). Attending physicians accounted for 86.0% of the attendance. Subjects completed 163 surveys (response rate: 82.3%) (The PRC Chair’s attendance was not included in the response rate calculation because he was not eligible to complete the survey). The results are summarized in Table 6.

**Discussion**

Our results show that implementing a peer review process such as the one in this study can result in increased HCP incident reporting over time. These findings are particularly meaningful because studies showing objective improvements in HCP incident reporting due to any process implementation are rare [12, 13]. Furthermore, to our knowledge, no such studies have been reported related to a peer review process implementation.

Experts assert that HCPs are more likely to report incidents if they perceive value in doing so and if they trust reporting to be non-threatening [17, 18]. Accordingly, our peer review processes were designed to foster such an environment. While it can be concluded that the process as a whole led to increased reporting, it remains unclear to what degree the components of the process contributed to the results. The physician survey suggests that physicians perceived value in the process. On balance, respondents reported that the process identified errors and provided information that could improve practice. They perceived the process to be highly educational, reporting that it improved understanding of systems barriers to quality and would lead to a change in their practice.

The survey tool was originally designed to fulfill CME requirements and initially was not intended for scientific investigation. While it did provide useful data on physician perceptions of the value of the peer review process, it provided no data on other components that may have increased reporting such as perceptions of the process being non-punitive. Future surveys of this type should address all potential factors that may affect HCP reporting.

While we did not measure perceptions of the potential for reporting to be punitive, our results and anecdotal feedback suggest that the intervention was perceived as non-punitive. The significant increase in self-reported incidents suggests that practitioners trusted the process. Anecdotal feedback from ED HCPs suggested that increased transparency as well as the opportunity for them to be directly involved in the review process did foster trust. This feedback is consistent with prior research, which found physicians to be more inclined to report incidents in departments that had investigational processes.

---

**Table 5** ITS model results of HP vs. OP incident reports

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>SE (estimate)</th>
<th>$T$-value</th>
<th>$P$-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>1.09</td>
<td>0.91</td>
<td>1.20</td>
<td>0.2337</td>
</tr>
<tr>
<td>Month</td>
<td>-0.21</td>
<td>0.19</td>
<td>-1.10</td>
<td>0.2743</td>
</tr>
<tr>
<td>Pre-intervention slope</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intervention—post</td>
<td>1.66</td>
<td>0.96</td>
<td>1.73</td>
<td>0.0902</td>
</tr>
<tr>
<td>Group Effect—HP</td>
<td>1.94</td>
<td>0.57</td>
<td>3.43</td>
<td>0.0012</td>
</tr>
<tr>
<td>Month Post-intervention slopes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OP</td>
<td>0.19</td>
<td>0.20</td>
<td>0.95</td>
<td>0.3479</td>
</tr>
<tr>
<td>HP</td>
<td>0.43</td>
<td>0.20</td>
<td>2.18</td>
<td>0.0339</td>
</tr>
<tr>
<td>Diff: HP–OP</td>
<td>0.24</td>
<td>0.05</td>
<td>4.84</td>
<td>$&lt;0.0001$</td>
</tr>
</tbody>
</table>

**Table 6** PRC assessment survey questions and subject responses

<table>
<thead>
<tr>
<th>Survey questions</th>
<th>Subject responses $(n = 163)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. How would you rate this educational activity overall? Please circle 5 4 3 2 1</td>
<td>Mean: 4.97</td>
</tr>
<tr>
<td>(5 = excellent, 1 = poor)</td>
<td></td>
</tr>
<tr>
<td>2. Do you plan on making any changes in your practice as a result of this activity?</td>
<td>Yes: 71.1%</td>
</tr>
<tr>
<td>3a. Case discussions led to improved understanding of systems barriers to quality</td>
<td>Yes: 100%</td>
</tr>
<tr>
<td>3b. Case discussions identify errors from which I and the department as a whole can improve our practice</td>
<td>Frequency of option being selected by respondents</td>
</tr>
<tr>
<td>4. Which of the following competency areas do you feel have been improved as result of this activity? Select all that apply: Patient care</td>
<td>87.1%</td>
</tr>
<tr>
<td>Professionalism</td>
<td>25.2%</td>
</tr>
<tr>
<td>Practice based learning</td>
<td>42.3%</td>
</tr>
<tr>
<td>Medical knowledge</td>
<td>57.1%</td>
</tr>
<tr>
<td>Communication skills</td>
<td>39.9%</td>
</tr>
<tr>
<td>Systems based practice</td>
<td>60.1%</td>
</tr>
</tbody>
</table>
that were administered by their peers [19]. The increase in reporting in a non-anonymous system also suggests a reduced fear of retribution. Previous studies suggest that anonymous reporting systems can increase reporting [6, 7]. Our results suggest that anonymity is not essential to increase incident reporting. If an incident reporting and review process is perceived to be safe, anonymity may not be necessary.

Standardized feedback may also have contributed to increases in reporting. Previous publications have reported timeliness and quality of feedback as important factors in promoting incident reporting [14, 15, 20–22]. While the value of standardized feedback per se was not evaluated in our study, HCPs within the hospital and ED expressed appreciation for the timeliness and robustness of feedback following the new process implementation. Several unsolicited comments highlighted a perceived improvement in the ED’s overall commitment to quality.

Based on the study results, our experience and previously reported research, we believe that three components of the peer review process (the perceived value of the process, the perceived non-punitive environment and the improved feedback) contributed to the increase in HCP reporting. Because the relative importance of these components remains unclear, future studies are warranted to identify opportunities to further improve the process.

The study had several additional limitations. First, the OP group was chosen as a convenience control group presumed to have minimal exposure to the intervention. It is possible that OP members were exposed indirectly to the intervention through unknown mechanisms. Second, we cannot exclude the possibility that other influences within the hospital caused the changes we observed, but we are unaware of other specifically relevant quality interventions implemented within the hospital or ED during the study period. General messages about improving quality and safety seem unlikely to have produced the time-series pattern we observed. Nevertheless, we cannot exclude a general change in safety awareness contributing to increased reporting. Finally, the peer review process was implemented in a single department at one institution. We cannot be sure that other institutions or departments with different cultures would respond to this intervention in the same way.

In conclusion, improving HCP incident reporting is an important component in optimizing health care quality and safety. The implementation of a non-punitive peer review process that provides timely feedback appeared to increase incident reporting by HCPs and was perceived as a valuable by physicians.

References


