

Reducing errors made by emergency physicians in interpreting radiographs: longitudinal study

James A Espinosa, Thomas W Nolan

Abstract

Objectives To reduce errors made in the interpretation of radiographs in an emergency department.

Design Longitudinal study.

Setting Hospital emergency department.

Interventions All staff reviewed all clinically significant discrepancies at monthly meetings. A file of clinically significant errors was created; the file was used for teaching. Later a team redesigned the process. A system was developed for interpreting radiographs that would be followed regardless of the day of the week or time of day. All standard radiographs were brought directly to the emergency physician for immediate interpretation. Radiologists reviewed the films within 12 hours as a quality control measure, and if a significant misinterpretation was found patients were asked to return.

Main outcome measures Reduction in number of clinically significant errors (such as missed fractures or foreign bodies) on radiographs read in the emergency department. Data on the error rate for radiologists and the effect of the recall procedure were not available so reliability modelling was used to assess the effect of these on overall safety.

Results After the initial improvements the rate of false negative errors fell from 3% (95% confidence interval 2.8% to 3.2%) to 1.2% (1.03% to 1.37%). After the processes were redesigned it fell further to 0.3% (0.26% to 0.34%). Reliability modelling showed that the number of potential adverse effects per 1000 cases fell from 19 before the improvements to 3 afterwards and unmitigated adverse effects fell from 2.2/1000 before to 0.16/1000 afterwards, assuming 95% success in calling patients back.

Conclusion Systems of radiograph interpretation that optimise the skills of all clinicians involved and contain reliable processes for mitigating errors can reduce error rates substantially.

Introduction

To better serve the needs of patients and to reduce costs, new ways of caring for patients must be designed. Many of these improvements will optimise the work of a team of clinicians, with team members working at their highest capability. Our aim is to illustrate a collaborative approach used by radiologists and emer-

gency physicians to achieve a significant reduction in the errors made by emergency physicians in the interpretation of radiographs.

Rates of disagreement between emergency physicians and radiologists in the interpretation of radiographs range from 8-11%.¹⁻⁵ A change of treatment was required for 1-3% of these patients. These errors in interpreting radiographs in the emergency department can also have significant clinical and legal consequences.⁶ As early as 1984, Trautlein published a study of 200 consecutive malpractice claims stemming from treatment in emergency departments. Radiographs were found to have been incorrectly interpreted in 38 of the cases⁷; these misinterpretations included missed fractures in 14 patients and missed foreign bodies in eight. Between 1974 and 1985 the liability programme of the American College of Emergency Physicians identified the most frequent cause of malpractice actions as the failure to diagnose fractures of the extremities and other fractures. The second most frequent cause was the failure to identify a foreign body in a wound. Errors related to the misdiagnosis of fractures accounted for the third largest amount of dollars paid out to settle malpractice claims.⁸

Case study

Baseline analysis

In the third quarter of 1993, JAE initiated an analysis of the errors made by emergency physicians in interpreting radiographs at his hospital. Clinically significant errors in interpretation were defined, according to the method of Fleisher, as a false negative interpretation that would have resulted in a change in the patient's care.¹ The rate of clinically significant errors was tracked each month and was based on an analysis of daily events. From July 1993 to December 1994 the average was 3% (figure). This was consistent with previously reported rates of clinically significant misinterpretations, although it was at the high end of the 1-3% range. Films that were not interpreted by emergency physicians, such as specialised scans, ultrasound scans, and intravenous pyelogram studies, were not included in the data.

At the time of the analysis, four separate radiology processes were in place for patients seen in the emergency department. One process operated from 8 00 am to 5 00 pm on weekdays; during this time radiologists interpreted the radiographs and transmitted the results to the emergency physician. From 5 00 pm until

Department of
Emergency
Medicine, Overlook
Hospital, Atlantic
Health System, 99
Beauvoir Avenue,
Summit, NJ 07902,
USA

James A Espinosa
medical director

Associates in
Process
Improvement, 1110
Bonifant Street,
Silver Spring, MD
20910, USA

Thomas W Nolan
statistician

Correspondence to:
J A Espinosa
jim010@aol.com

BMJ 2000;320:737-40

website
extra

Additional
information about
reliability modelling
appears on the
BMJ's website

www.bmj.com

10 00 pm radiology residents interpreted the radiographs and transmitted the results to the emergency physician. From 10 00 pm until 8 00 am the emergency physician interpreted the radiograph and a subsequent interpretation (an "over read") was done by a radiologist within 12 hours as a quality control measure. On weekends, a fourth system consisted of a mixture of the other systems. However, exceptions to the first process, in which primary responsibility for interpretation was with the radiologist, were frequent and resulted in the use of the third process, in which primary responsibility fell to the emergency physician even between 8 00 am and 5 00 pm on weekdays. The data available for the initial analysis were from cases in which the emergency physician was the primary interpreter of the radiograph and the radiologist provided the subsequent over read.

This variety of processes resulted in a standard of care that differed, at least implicitly, depending on the day of the week and the time of the day. This situation conformed to O'Leary et al's assessment of practice patterns in the interpretation of radiographs in emergency departments: "The interhospital variability of responsibilities within the course of a 168-hour week suggest that the evolution of these arrangements may have been driven by staffing constraints, scheduling practices and convenience, rather than by efforts to provide a consistent high standard of care."⁹

Initial improvement efforts

The initial aim was to reduce the number of interpretation errors made by emergency physicians. This initial work preceded a more extensive redesign of the system that would address the variation in the responsibility between emergency physicians and radiologists.

A quality improvement effort began in the first quarter of 1995, using an approach similar to that described by Berwick.¹⁰ Several improvements were implemented. All staff reviewed all clinically significant discrepancies at monthly meetings in a blame free environment. A common error discussed, for example, was the finding of small chip fractures on an ankle film series that had been interpreted by the emergency physician as negative. A file of clinically significant errors was created. This has been maintained since 1994 and is used for training. Mandatory study of the entire file has become part of the orientation of all new staff. Overall departmental patterns of error were identified from this file, and a focused review of these patterns took place at staff meetings. Patterns of errors made by each emergency physician were identified and were discussed as part of routine, ongoing communication and during performance reviews.

During the baseline period from July 1993 to December 1994 false negative errors had been made by the emergency physicians in 3% (95% confidence interval 2.8% to 3.2%) of the 28 161 cases reviewed. From August 1995 to August 1996, the rate of false negatives fell to 1.2%. (1.03% to 1.37%) for the 20 236 cases studied.

Fundamental redesign

Background

The interpretation of radiographs by emergency physicians had been made more reliable but the four different processes for having the radiographs interpreted were still in place. Long delays in taking and

processing the films were common. Patients, emergency physicians, attending physicians, and nurses were unhappy with the process, as documented by data from both external and internal surveys of satisfaction. The hospital uses an externally benchmarked customer satisfaction survey for patients seen in the emergency department, inpatient services, and outpatient services (Press Ganey Associates, South Bend, Indiana). One question asks patients to rate their satisfaction with the time spent waiting for a radiograph. Before the processes were changed satisfaction was rated in the 20th centile. Internal data from private physicians and emergency physicians showed that there was considerable dissatisfaction with the process, especially with the difficulties in locating films needing to be reviewed.

The hospital managers commissioned an interdisciplinary team to redesign the process; it was composed of a staff radiologist, two radiology residents, a staff emergency physician, the managers of the emergency department and the diagnostic radiology department, the executive director of radiology, and the medical directors of the emergency department and radiology. The team's goal was to improve patients' satisfaction by significantly shortening the time they spent in the department waiting for the interpretation of their radiographs to become available. The team also aimed to further reduce the number of errors made in interpreting radiographs.

Changes

A system was developed for interpreting radiographs that would be followed regardless of the day of the week or the time of day. All standard radiographs were to be brought directly to the emergency physician for immediate interpretation. A radiologist would provide an interpretation within 12 hours as a quality control measure. When a clinically significant misinterpretation was found by the radiologist, staff from the emergency department would contact patients and ask them to return.

This process clearly assigned the primary, initial responsibility for interpreting plain radiographs to the emergency physician. The new system reduced the confusion that had been caused by ambiguously defined roles and engendered collaboration between the radiology and emergency departments to redesign other related processes to make the new system function reliably.

Other improvements to support the new system were also made. The process of communication between the radiologists and the emergency physicians was enhanced to better address situations in which a clinically significant error was detected. A new form was designed to provide feedback to the physicians, and a process for using the form was implemented. Potentially significant discrepancies are logged by the radiologist. These are then formally sent to the emergency physician on duty who is responsible for reviewing the films and the patient's record. If a discrepancy is identified, the emergency physician is responsible for contacting the patient and documenting the contact. This process embeds the ongoing training of emergency physicians into the normal process of care. Additionally, radiologists can call the emergency department and speak to an emergency physician when a discrepancy is identified.

Predictions of error rates obtained by reliability modelling. Results were multiplied by 1000 to define rates per 1000 cases

Process	Potential adverse events/1000 cases	Unmitigated potential adverse events/1000 cases (95% recall rate)	Unmitigated potential adverse events/1000 cases (99% recall rate)
Radiologist interprets radiograph	3 ^(a)	3 ^(a)	3 ^(a)
Before initial improvements:			
40% of cases interpreted by radiologists; 60% of cases interpreted by emergency physician with radiologist doing subsequent reading	19 ^(b)	2.2 ^(c)	1.4
After initial improvements:			
40% of cases interpreted by radiologists; 60% of cases interpreted by emergency physician with radiologist doing subsequent reading	8.4 ^(d)	1.6 ^(e)	1.3
Redesigned system:			
Emergency physician interprets all radiographs; radiologist does subsequent reading	3 ^(f)	0.16 ^(g)	0.039

(a) See assumption 2. The rate of potential adverse events and the rate of adverse events are the same since there is no subsequent reading by a radiologist.

(b) The weighted average of the error rates of 0.03 for emergency physicians and 0.003 for radiologists (assumption 4).

(c) The sum of three components: $0.4(0.003) + 0.6(0.03)(0.003) + 0.6(0.03)(0.997)(0.05)$. The first is the contribution of an error by the radiologist. The second is the contribution of an error by the emergency physician that is not identified on the subsequent reading. The third is the case in which an error made by the emergency physician is identified by the radiologist but the patient does not return for appropriate care (assumption 6).

(d) Same as (b) with the improved error rate of 0.012 for the emergency physician substituted for the initial rate of 0.03.

(e) Same as (c) with the exception of the new error rate of 0.012.

(f) Error rate for emergency physicians after the system was redesigned.

(g) Sum of two components: the first, $(0.003)(0.003)$, represents an error by the emergency physician that is undetected by the radiologist. The second, $(0.003)(0.997)(0.05)$, represents the case in which an error by the emergency physician is identified by the radiologist but the patient does not return for appropriate treatment.

The team redesigned the processes in the radiology film room to accommodate changes made in the emergency department. Additionally, a film alternator was placed in the emergency department, which allows the outputs of a series of radiographs to be viewed simultaneously and facilitates interpretation.

Results

From November 1996 to December 1999 the rate of false negatives for the 67 111 cases studied was 0.3% (0.26% to 0.34%). This result is consistent with the 0.4% rate of clinically significant misinterpretation of radiographs reported by Preston et al after a similar improvement effort.¹¹

Substantial improvements in patient satisfaction and a shortening of the turnaround time for interpretations also occurred as a result of the redesign. Improvements included a 50% reduction in the time it took from ordering plain films to having them returned to the emergency department. The centile ranking of the items on the patient satisfactions survey relating to radiology rose above the 90th centile. There was also a 50% reduction in the time it took for patients presenting to the emergency department with trauma to an extremity to be discharged. There was an improvement in the overall satisfaction of patients with their visit to the emergency department: satisfaction ratings rose from the 18th centile to above the 95th centile.

Reliability modelling

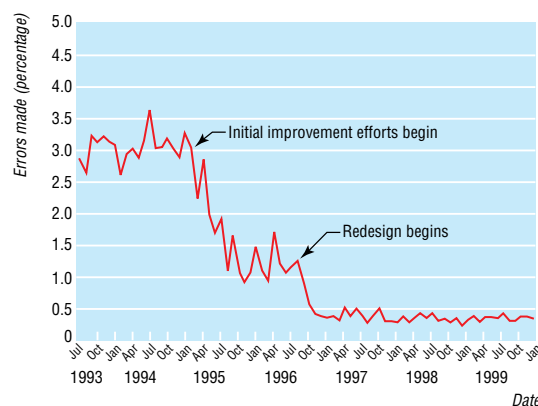
The safety of the system with respect to the misinterpretation of radiographs depends on the reliability of the interpretations of the emergency physician and the radiologist, the effectiveness of the recall procedure when a clinically significant error is identified, and the structure of the interaction among the components. The figure shows the error rate for the interpretation of radiographs by emergency physicians. Data on radiologists' error rates and on the effectiveness of the recall procedure were not available. In the absence of these data and to provide a quantitative prediction of what had been accomplished as a result of the improvements we modelled the reliability of the

system. Reliability modelling is a common statistical and engineering technique used to ascertain the reliability and safety of alternative structures of systems.¹² (Further information on reliability modelling is available in an appendix on the *BMJ's* website.)

For the purposes of this analysis, a potential adverse event was defined as a case in which a clinically significant error of misinterpretation is made on the first reading of the radiograph either by the emergency physician or the radiologist. An unmitigated potential adverse event is defined as a case in which an error of misinterpretation is made and the error goes undetected or the error is detected but the patient does not return for treatment. The rates of potential adverse events and unmitigated potential adverse events provide a measure of the reliability of the system.

The model was based on the following assumptions:

- (1) The error rates for emergency physicians can be obtained from the data in the figure
- (2) The error rates for the radiologists are assumed to be stable at 0.3%, which was the final error rate achieved by the emergency physicians



Trends in the rate of errors (clinically significant false negative interpretations) made by emergency physicians in interpreting radiographs. Details of the changes are described in the text

(3) If a radiologist makes an error of interpretation, the emergency physician will not identify it
 (4) Before the new system was implemented, a radiologist interpreted 40% of the radiographs and an emergency physician interpreted 60% with a subsequent reading by a radiologist occurring within 12 hours
 (5) The radiologist's interpretation was not influenced by the emergency physician's interpretation
 (6) Not all patients whose radiographs were found to have a clinically significant misinterpretation would return to the emergency department. A success rate of 95% of patients returning was assumed but the model was also run for a 99% success rate for comparison. The predictions obtained by the model are shown in the table.

Nolan has outlined a strategy that includes three components for increasing the safety and reliability of a system.¹³ The first is to design the system to prevent errors. The second is to design the procedures to make errors visible when they do occur so that they can be intercepted. And the third is to design procedures for mitigating the adverse effects of errors when they are not detected and intercepted.

The table illustrates the effect of each of these three components on the reliability of the system. The effect of the decrease in the emergency physicians' error rate on the rate of potential adverse events can be seen by comparing the weighted average of the error rates for emergency physicians and radiologists with the improved error rate for emergency physicians and the error rate for emergency physicians after the system was redesigned. By comparing the rate of potential adverse events with the associated rate of unmitigated adverse events for each process the impact of the identification of errors by the radiologist can be seen.

The mitigation strategy in this study was to have the patients who represented potential adverse events return to the emergency department for treatment. The effect of the mitigation (the recall process) on the rate of unmitigated potential adverse events can be seen by comparing the third and fourth columns in the table. The success rate for the recall process has a substantial effect on the rate of unmitigated potential adverse events when the interpretation errors are few (table).

Conclusions

Efforts to reduce errors in health care can be directed at a wide variety of processes. We have described efforts to reduce the rate of error among emergency physicians interpreting radiographs. The validation of such efforts will largely depend on replication by other centres. This change has been sustained over time, which suggests that the changes were sufficiently robust to meet the challenge of a complex and changing environment.

The error rate for emergency physicians who are interpreting radiographs can be reduced to the level of incidents per thousand. Systems for interpreting radiographs that are designed to optimise the skills of both emergency physicians and radiologists and that contain reliable processes for mitigating errors can reduce the rate of unmitigated potential adverse events to the level of parts per hundred thousand. This rate is substantially lower than the rate that could be achieved by either an emergency physician or a radiologist acting alone. As error rates are reduced to the level of

What is already known on this topic

Most previous studies have estimated the rates of misinterpreting radiographs

These studies have usually focused on one professional group, such as emergency physicians or radiologists, over a short interval

What this study adds

This six year, longitudinal study takes a systems approach to the problem of misinterpreting radiographs

Rather than comparing the performance of two groups of professionals, this study shows the impact of cooperation between emergency physicians and radiologists in reducing errors and the potential adverse events that result from them

parts per thousand, the process of mitigating errors has a substantial effect and further reduces the incidence of unmitigated potential adverse events.

This approach seems to have generalisable features. The design principles might also apply to reducing other adverse events in emergency departments, such as the preventable adverse events associated with the management of patients with chest pain. When any large effort at redesigning systems is undertaken or significant clinical change is made, it is prudent to study errors as an important outcome measure of undesirable side effects.

Contributors: JAE provided leadership for the changes and performed the study at his hospital. TWN recognized the broader systems issues addressed by the study and developed the approach to reliability modelling used to understand the broader systems impact of the work. Both authors shared equally in writing the paper. JAE is guarantor.

Funding: None.

Competing interests: None declared.

- 1 Robinson PJ, Wilson D, Coral A, Murphy A, Verow P. Variation between experienced observers in the interpretation of accident and emergency radiographs. *Br J Radiol* 1999;72:323-30.
- 2 Lufkin KC, Smith SW, Matticks CA, Brunette DD. Radiologists' review of radiographs interpreted confidently by emergency physicians infrequently lead to changes in patient management. *Ann Emerg Med* 1998;31:202-7.
- 3 Scott WW Jr, Bluemke DA, Mysko WK, Weller GE, Kelen GD, Reichle RL, et al. Interpretation of emergency department radiographs by radiologists and emergency medicine physicians: teleradiology workstation versus radiograph readings. *Radiology* 1995;195:223-9.
- 4 Fleischer G, Ludwig S, McSorley M. Interpretation of pediatric x-ray films by emergency department pediatricians. *Ann Emerg Med* 1983;12:153-8.
- 5 Rhea JT, Potsaid MS, DeLuca SA. Errors of interpretation as elicited by a quality audit of an emergency department facility. *Radiology* 1979;132:277-80.
- 6 George JE, Espinosa JA. Legal issues in emergency radiology: Practical strategies to reduce risk. *Emerg Med Clin North Am* 1992;10:179-203
- 7 Traudlein JJ. Malpractice in the emergency department-review of 200 cases. *Ann Emerg Med* 1984;13:709-11.
- 8 Fastow, JS. Medico-legal risks: identification and reduction. In: *American College of Emergency Physicians comprehensive guide to effective practice management*. Dallas: ACEP, 1986.
- 9 O'Leary MK, Smith M, Olmsted WW, Curtis DJ. Physicians' assessment of practice patterns in emergency department radiograph interpretation. *Ann Emerg Med* 1988;17:1019-23.
- 10 Berwick DM. A primer on leading the improvement of systems. *BMJ* 1996;312:619-22.
- 11 Preston CA, Marr JJ, Amaraneni KK, Suthar BS. Reduction of "call-backs" to the emergency department due to discrepancies in the plain radiograph interpretation. *Am J Emerg Med* 1998;16:160-2.
- 12 Ushakov I. *Handbook of reliability engineering*. New York: John Wiley, 1994.
- 13 Nolan TW. System changes to improve patient safety. *BMJ* 2000;320:771-3.

(Accepted 23 February 2000)