Surgical specimen identification errors: A new measure of quality in surgical care

Martin A. Makary, MD, MPH,[a,d,e] Jonathan Epstein, MD,[b] Peter J. Pronovost, MD, PhD,[a,c,d,e] E. Anne Millman, MS,[a,d,e] Emily C. Hartmann, MS,[a] *and* **Julie A. Freischlag, MD,[a]** *Baltimore, Md*

Background. *Communication errors are the primary factor contributing to all types of sentinel events including those involving surgical patients. One type of communication error is mislabeled specimens. The extent to which these errors occur is poorly quantified. We designed a study to measure the incidence and type of specimen identification errors in the surgical patient population.*

Methods. *We performed a prospective cohort study that included all patients who underwent surgery in an outpatient clinic or hospital operating room and for whom a pathology specimen was sent to the laboratory. The study took place during a 6-month period (October 2004 to April 2005) at an urban, academic medical center. The study's main end-points were the incidence and type of specimen labeling errors in the hospital operating room and the outpatient clinic. The specimen was the unit of analysis. All specimens were screened for "identification errors," which, for the purposes of this study, were defined as any discrepancy between information on the specimen requisition form and the accompanying labeled specimen received in the laboratory. Errors were stratified by the type of identification error, source, location, and type of procedure.*

Results. *A total of 21,351 surgical specimens were included in the analysis. There were 91 (4.3/1000) surgical specimen identification errors (18, specimen not labeled; 16, empty container; 16, laterality incorrect; 14, incorrect tissue site; 11, incorrect patient; 9, no patient name; and 7, no tissue site). Identification errors occurred in 0.512% of specimens originating from an outpatient clinic (53/10,354 specimens) and 0.346% of specimens originating from an operating room (38/10,997 specimens). Procedures involving the breast were the most common type to involve an identification error (breast = 11, skin = 10, colon = 8); in addition, 59.3% (54/91) of errors were associated with a biopsy procedure. Follow-up was complete in all cases found to have an identification error.*

Conclusions. *Surgical specimen identification errors are common and pose important risks to all patients. In our study, these events occurred in 4.3 per 1000 surgical specimens or an annualized rate of occurrence of 182 mislabeled specimens per year. Given the frequency with which these errors occur and their potential effect on patients, the rate of surgical specimen identification errors may be an important measure of patient safety. Strategies to reduce the rate of these errors should be a research priority. (Surgery 2007;141:450-5.)*

From the Departments of Surgery,[a] Surgical Pathology,[b] and Anesthesiology,[c] Center for Surgical Outcomes Research, John Hopkins University School of Medicine; Department of Health Policy and Management,[d] Johns Hopkins Bloomberg School of Public Health; and the Johns Hopkins Quality and Safety Research Group,[e] Johns Hopkins Medical Institutions, Baltimore, Md

Though communication failures are a major contributor to all types of sentinel events,[1,2] we lack scientifically sound and feasible methods with which to measure communication.[3-5] Communication failures are particularly problematic among surgical team members and result in preventable morbidity, mortality, and high costs of care. In studies of surgical patients, communication failures have been identified as the root cause in 80% of sentinel events, 77% of wrong-site operations, and other medical errors in the operating room.[6-8] One type of communication failure that poses risks to patients is the mislabeling of a surgical specimen prior to its arrival in a pathology laboratory. Specifically, an error in verbal communication and transcription during the hand-off increases the risk

Accepted for publication September 9, 2006.

Reprint requests: Martin A. Makary, MD, MPH, Assistant Professor of Surgery and Health Policy & Management, Johns Hopkins University, 4940 Eastern Ave, Building A-5, Baltimore, MD 21224. E-mail: mmakary1@jhmi.edu.

0039-6060/$ - see front matter

doi:10.1016/j.surg.2006.08.018

of mislabeling the specimen. Indeed, the relay of vital surgical specimen information occurs within a local culture of teamwork and communication between a physician and nurse or technician.[9] Although this context of information transfer is associated with rates of medical errors, the local culture of teamwork is difficult to quantify.[5] One consequence of poor teamwork and communication that can harm patients and can be measured is the occurrence of specimens being labeled incorrectly.

Despite the known risk to patient safety, the rate of mislabeled specimens remains poorly defined, and few hospitals routinely monitor these rates as a measure of patient safety. We hypothesize that mislabeled surgical specimens occur commonly and that the rate of these occurrences can be used to measure patient safety in the surgical patient population. The specific aim of this study was to define and measure the frequency of mislabeled specimens in the setting of surgical procedures.

METHODS

We used a prospective cohort design. All surgical pathology specimens were screened by the surgical pathology department over a 6-month period (October 2004 to April 2005) at an urban, academic medical center. The primary dependent variable was the rate of mislabeled surgical specimens, and the specimen was the unit of analysis. All specimens were screened for errors, which were defined as any discrepancy between information on the specimen requisition form and the accompanying labeled specimen received in the laboratory. Errors were stratified by the type of identification error, source location, and type of procedure. We classified type of identification error by the type of discrepancy found between the specimen and accompanying requisition form. The following categories were identified: "specimen not labeled," which included specimens without a label present on the container; "empty container," which included containers with no tissue specimen; "laterality incorrect," which included specimens for which the left versus right designation was inconsistent with the label or requisition; "incorrect tissue site," which included any organ- or body-site–designation discrepancy between the label and requisition, excluding laterality discrepancies; "incorrect patient name," which included any wrong patient label on the specimen or requisition form; "no patient name," which included specimens missing a patient name on the container and/or requisition; and "no tissue site," which included any specimen for which the site designation of the specimen was absent.

Source location was classified as clinic versus operating room. Type of procedure was classified using the source organ of the specimen (eg, breast procedure, skin procedure, prostate procedure, etc.). The results were descriptive, and we presented the data as proportions. Comparisons were performed using a chi-square test using SPSS version 12.0 (SPSS, Chicago, Ill). Institutional Review Board (IRB) exemption approval was obtained from the Johns Hopkins University for the publication of these findings as quality measurement data.

RESULTS

Of 21,351 specimens studied, 91 (4.3 per 1000 specimens) were associated with an identification error. These included the following types of identification errors: 18, specimen not labeled; 16, empty container; 16, laterality incorrect; 14, incorrect tissue site; 11, incorrect patient; 9, no patient name; and 7, no tissue site identified (Fig 1).

Of 10,354 specimens originating from the operating room, 38 contained errors (3.7 per 1000 specimens). Of the 10,997 specimens received from a clinic, 53 contained errors (4.8/1000) ($P = .062$). Compared with specimens received from the operating room, specimens from a clinic had an increased proportion of identification errors involving an empty container (1.00 vs 0.48/1000), incorrect laterality (1.00 vs 0.48/1000), incorrect tissue site (0.82 vs 0.48/1000), and incorrect patient (0.64 vs 0.38/1000); whereas specimens received from an operating room were more likely to have a specimen not labeled (1.16 vs 0.55/1000), no patient name (0.58 vs 0.28/1000), and no tissue site (0.97 vs 0.55/1000) (Table I).

Among the 91 mislabeled surgical specimens, the largest number came from surgical procedures involving a biopsy (n = 54), followed by procedures associated with an excision (n = 24), resection (n = 3), or other procedure (n = 10). Among the 91 mislabeled surgical specimens, breast tissue was the most common (n = 11), followed by skin (n = 10), and colon (n = 9) (Table II).

CONCLUSION

In this study, we found that mislabeled surgical specimens are common, occurring at an overall rate of 4.3 per 1000 specimens, with 4.8 per 1000 occurring in a clinic and 3.7 per 1000 in an operating room; when annualized, the rate of occurrence at this single institution is 182 mislabeled specimens each year, all posing significant risks to patient safety. Given the high frequency with which these events occurred and the feasibility and validity of measuring them, mislabeled surgical speci-

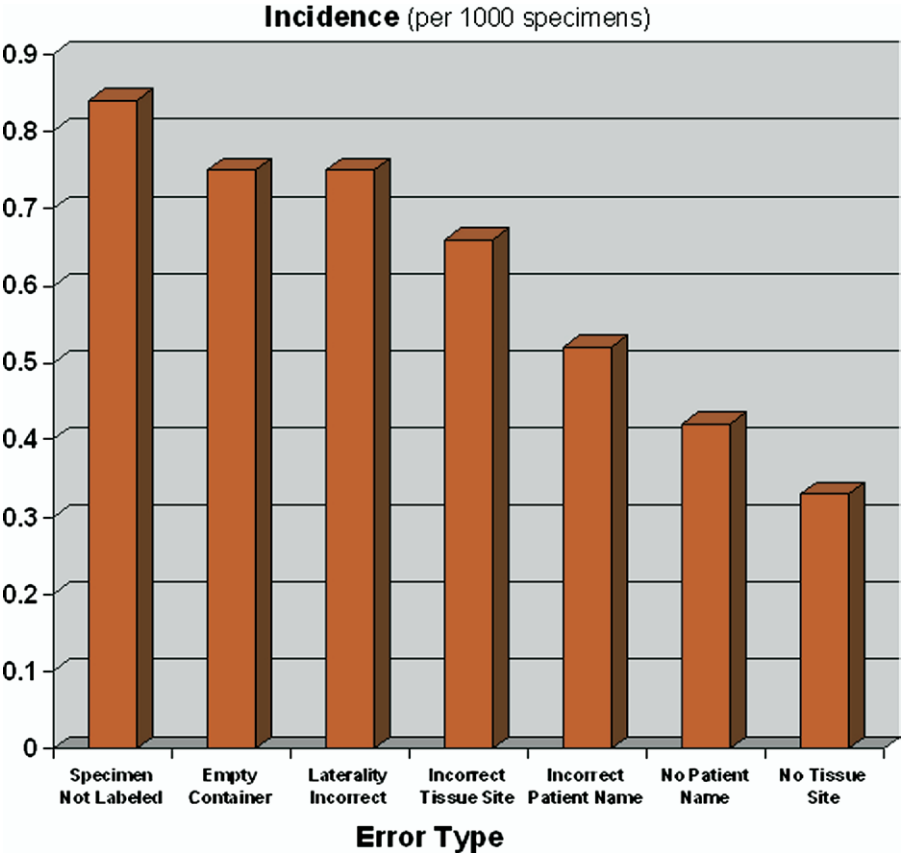

Figure. Incidence of identification errors observed per 1000 specimens (n = 21,351).

Table I. Type of specimen identification error by source location (clinic vs operating room [OR])

	Total		*Clinic*		*OR*	
Error type	*Events (% of total)*	*Proportion per 1000 specimens (n = 21,351)*	*Events (% of total)*	*Proportion per 1000 specimens (n = 10,997)*	*Events (% of total)*	*Proportion per 1000 specimens (n = 10,354)*
Specimen not labeled	18 (19.8)	0.84	6 (11.3)	0.55	12 (31.6)	1.16
Empty container	16 (17.6)	0.75	11 (20.8)	1.00	5 (13.2)	0.48
Laterality incorrect	16 (17.6)	0.75	11 (20.8)	1.00	5 (13.2)	0.48
Incorrect tissue site	14 (15.4)	0.66	9 (17.0)	0.82	5 (13.2)	0.48
Incorrect patient	11 (12.1)	0.52	7 (13.2)	0.64	4 (10.5)	0.38
No patient name	9 (9.9)	0.42	3 (5.7)	0.27	6 (15.8)	0.58
No tissue site	7 (7.7)	0.33	6 (11.3)	0.55	1 (2.6)	0.97
Total	91 (100)	4.26	53 (100)	4.82	38 (100)	3.67

mens may serve as a useful measure for patient safety in surgical patients. In fact, the data for this study were collected as part of routine quality audits in the Division of Surgical Pathology. Furthermore, *not* investigating these errors may represent a failure to recover from an error.

In the study presented, a full-time, dedicated quality improvement nurse was notified about all mislabeled surgical specimens. The nurse immediately undertook an aggressive investigation to find the correct information needed to resolve the discrepancy. This individual searched medical records to identify the intended procedure and tissue sought for biopsy or resection. In addition, operating room personnel were interviewed, and the surgeon who performed the procedure was contacted

Table II. Common procedure types associated with mislabeled specimens (n = 91)

Procedure type	*Events*	*(% total)*
Breast	11	(12.1)
Skin	10	(11.0)
Colon	9	(9.9)
Prostate	6	(6.6)
Uterus	4	(4.4)
Bone marrow	4	(4.4)
Vertebrae	4	(4.4)
Small bowel	4	(4.4)
Brain	3	(3.3)
Cervix	3	(3.3)
Liver	2	(2.2)
Lymph node	2	(2.2)
Tonsil	2	(2.2)
Thyroid	2	(2.2)
Larynx	2	(2.2)
Anus	2	(2.2)
Other	23	(25.3)

by pager to discuss the discrepancy. All incidents of mislabeled surgical specimens were resolved without patient harm. However, the extent of patient harm prior to this program at our hospital, or any hospital, is unknown and has not been reported in the literature. Nevertheless, errors involving specimen identification could have resulted in delay in care, the need for an additional biopsy or therapy, failure to use appropriate therapy, or therapy administered to the wrong body site, side, or patient. These system failures can have serious implications to patients and providers and result in significant harm to the patient, costs to the institution, and distrust by a community. In one study of the clinical impact of errors in diagnosing surgical specimens at 4 hospitals, 39% to 45% of errors were associated with patient harm.[10]

Errors involving surgical specimens are recognized to occur during 1 of 3 discrete phases: the preanalytical phase (transfer of information from physician to nurse during a procedure, and subsequent specimen labeling, packaging, and transport), the analytical phase (handling a specimen within a surgical pathology laboratory and interpretation by a pathologist), and the postanalytical phase (recording and relaying the interpretation for the clinician). The latter two phases have been well-studied and have error rates of 0.1% to 9%[11-15] and 0.12% to 3.4%, respectively.[16,17] In this study, we chose to focus on the preanalytical phase because it is poorly understood and may represent a significant and highly variable source of harm, given the paucity of checks and balances routinely used for this phase and the differences in local practices regarding specimen handling, labeling, and transport.[18] Although the benefit of second opinions and tracking systems within pathology departments have been established, a standardized protocol for surgeon-to-nurse/technician communication of specimen identification information in labeling and handling the specimen, and an independent check of specimen information in the same way that blood products are checked, are lacking.

The new Joint Commission on Accreditation of Healthcare Organization safety goal—requiring labeling specimens with at least 2 patient identifiers and a preoperative verification step in the presence of the physician (www.jointcommission.org/PatientSafety/NationalPatientSafetyGoals/)—may help standardize this process. In addition, recent efforts to improve communication in operating rooms, such as briefing and debriefing protocols,[5,19,20] and systems to evaluate adverse events, such as comprehensive safety programs,[21] are aimed at improving the transfer of such critical information and the process used to evaluate mistakes. The range of procedure types and sources (operating room and outpatient clinic) for mislabeled specimens included in our study indicates that the problem is prevalent across surgical care and a function of systems, not particular individuals. To address the issues of poor communication in the operating room, which can lead to mislabeled specimens and other potential errors, we have instituted a routine briefing[19] and debriefing[20] with every operating room procedure. The briefing and debriefing (in the form of checklists) is prompted by the circulator and reviewed by the surgeon, anesthesia provider, and nurse. The debriefing includes the question "Has the surgical specimen been verified?" as a double check for accuracy in the same way that blood products are verified by 2 persons. We are also promoting briefings and debriefings for all clinic and bedside procedures to identify and mitigate potential issues that can result in patient harm. These efforts are aimed at promoting a culture of safety, which we measure among all our staff annually using a psychometrically validated survey instrument.[5]

Measuring safety rates in surgical care has proved challenging. We generally lack standardized definitions of harm (ie, what constitutes a surgical complication); we are unsure of the appropriate denominator (ie, the population at risk); and we often rely on self-reported events that are notoriously biased because we lack an independent surveillance mechanism. Even if we could obtain a rate of harm, variation in case mix, methods of data collection, and random error would likely far ex-

ceed variation in safety.[22,23] Each type of quality measure has its weakness. For example, quality as measured by surgical site infection rates may be a function of a hospital's case mix of colorectal operations (which more commonly result in wound infection secondary to intestinal spillage of bacteria) versus nongastrointestinal laparoscopic operations (which involve incisions often too small to manifest a wound infection). Furthermore, the infection data is difficult and expensive to collect because it requires short- and long-term follow-up and uniform expertise to define what constitutes an infection. Other measures of quality, such as rates of postoperative deep venous thrombosis and myocardial infarctions, are also confounded by variability in detection and selection bias given the high rate of asymptomatic occurrences. Thus, the realized incidence of these events can simply be a function of the tools and vigilance used to search for them, not their true incidence.[24] As a result, hospitals that are good at collecting complication data are punished for having the worst results, and hospitals that do a poor job of capturing events are rewarded for having the best outcomes. We found that tracking the number of specimens that are mislabeled in the operating room or other procedure room is easily performed using a simple screening process at the receiving desk in the Surgical Pathology Department. As such, this method of monitoring quality is inexpensive, not subject to interpretation, and has a high capture rate, which makes it a scientifically sound, yet feasible, measure of patient safety in surgical patients.

We recognize several limitations in our findings. First, we did not obtain detailed data regarding the tissue or procedure type for all specimens. As a result, we cannot provide rates of error stratified by tissue or procedure type, although the hospital studied is representative of large academic facilities with a broad spectrum of surgical services. Second, we only captured errors that were identified. We did not evaluate whether there were additional errors, namely lost specimens that were not received by surgical pathology, or what percentage of errors resulted in harm. Considering these events, our rate is likely a conservative estimate of the true error rate and may understate the extent of the problem.

Though we have not yet reduced the rates of these errors, safety science provides some guidance on how we might accomplish this. In our safety efforts, we reduce errors by the following measures: (1) standardizing what we do and when we do it, (2) creating independent checks for key processes, and (3) learning from mistakes when they occur.[19,21,23,25] By applying these principles to mislabeled surgical specimens, we can standardize requisition forms, the process for completing the forms, and the training for staff who complete the forms. We also can create a checklist that a nurse or other member of the care team would complete prior to sending a specimen to the laboratory. Finally, when a mislabeled surgical specimen is discovered, we can review how and why it happened and whether we need to change the system. In this way, mislabeled surgical specimen events can be used both as a surrogate marker of patient safety and as a quality improvement flag that could signal the need to redesign the process of specimen identification in surgical care. In conclusion, mislabeled surgical specimens are common, represent a significant threat to patient safety, and should be a focus of patient safety improvement efforts.

The authors would like to thank Christine Holzmueller of the Johns Hopkins Quality and Safety Research Group for her assistance in the preparation of this manuscript, and Doris Deane, surgical pathology quality assurance coordinator, for her assistance with data management.

REFERENCES

1. Joint Commission on Accreditation of Healthcare Organizations. Sentinel events: evaluating cause and planning improvement. Oakbrook Terrace, IL: Joint Commission on Accreditation of Healthcare Organizations; 1998.
2. Lingard L, Espin S, Whyte S, et al. Communication failures in the operating room: An observational classification of recurrent types and effects. Qual Saf Health Care 2004;13:330-4.
3. Firth-Cozens J. Why communication fails in the operating room. Qual Saf Health Care 2004;13:327.
4. Espin S, Levinson W, Regehr G, Baker GR, Lingard L. Error or "act of god"? A study of patients' and operating room team members' perceptions of error definition, reporting, and disclosure. Surgery 2006;139:6-14.
5. Makary MA, Sexton JB, Freischlag JA, et al. Patient safety in surgery. Ann Surg 2006;243:628-35.
6. Edmonds CR, Liguori GA, Stanton MA. Two cases of a wrong-site peripheral nerve block and a process to prevent this complication. Reg Anesth Pain Med 2005;30:99-103.
7. Gawande AA, Studdert DM, Orav EJ, Brennan TA, Zinner MJ. Risk factors for retained instruments and sponges after surgery. N Engl J Med 2003;348:229-35.
8. Gawande AA, Zinner MJ, Studdert DM, Brennan TA. Analysis of errors reported by surgeons at three teaching hospitals. Surgery 2003;133:614-21.
9. Makary MA, Sexton JB, Freischlag JA, et al. Teamwork in the operating room: teamwork in the eye of the beholder. J Am Coll Surg 2006;202:848-52.
10. Raab SS, Grzybicki DM, Janosky JE, et al. Clinical impact and frequency of anatomic pathology errors in cancer diagnoses. Cancer 2005;104:2205-13.
11. Safrin RE, Bark CJ. Surgical pathology sign-out. routine review of every case by a second pathologist. Am J Surg Pathol 1993;17:1190-2.

12. Furness PN, Lauder I. A questionnaire-based survey of errors in diagnostic histopathology throughout the United Kingdom. J Clin Pathol 1997;50:457-60.
13. Renshaw AA, Young ML, Jiroutek MR. How many cases need to be reviewed to compare performance in surgical pathology? Am J Clin Pathol 2003;119:388-91.
14. McBroom HM, Ramsay AD. The clinicopathological meeting. A means of auditing diagnostic performance. Am J Surg Pathol 1993;17:75-80.
15. Raab SS, Nakhleh RE, Ruby SG. Patient safety in anatomic pathology: Measuring discrepancy frequencies and causes. Arch Pathol Lab Med 2005;129:459-66.
16. Nakhleh RE, Zarbo RJ. Amended reports in surgical pathology and implications for diagnostic error detection and avoidance: A college of American pathologists Q-probes study of 1,667,547 accessioned cases in 359 laboratories. Arch Pathol Lab Med 1998;122:303-9.
17. Raab SS, Grzybicki DM, Zarbo RJ, Meier FA, Geyer SJ, Jensen C. Anatomic pathology databases and patient safety. Arch Pathol Lab Med 2005;129:1246-51.
18. Zarbo RJ, Meier FA, Raab SS. Error detection in anatomic pathology. Arch Pathol Lab Med 2005;129:1237-45.
19. Makary MA, Holzmueller CG, Thompson D, et al. Operating room briefings. Jt Comm Jour Qual Saf 2006;32:351-5.
20. Makary MA, Holzmueller CG, Sexton JB, et al. Operating room debriefings. Jt Comm Jour Qual Saf 2006; 32:407-10.
21. Pronovost PJ, Weast B, Rosenstein B, et al. Implementing and validating a comprehensive unit-based safety program. J Patient Saf 2005;1:33-40.
22. Pronovost PJ, Nolan T, Zeger S, Miller M, Rubin H. How can clinicians measure safety and quality in acute care? Lancet 2004;363:1061-7.
23. Lilford R, Mohammed MA, Spiegelhalter D, Thomson R. Use and misuse of process and outcome data in managing performance of acute medical care: avoiding institutional stigma. Lancet 2004;363:1147-54.
24. 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 Inform Assoc 1998;5:305-14.
25. Pronovost PJ, Holzmueller CG, Martinez E, et al. A practical tool to investigate defects in patient care. Jt Comm Jour Qual Saf 2006;32:102-8.