Changes in Medical Errors after Implementation of a Handoff Program


BACKGROUND
Miscommunications are a leading cause of serious medical errors. Data from multicenter studies assessing programs designed to improve handoff of information about patient care are lacking.

METHODS
We conducted a prospective intervention study of a resident handoff-improvement program in nine hospitals, measuring rates of medical errors, preventable adverse events, and miscommunications, as well as resident workflow. The intervention included a mnemonic to standardize oral and written handoffs, handoff and communication training, a faculty development and observation program, and a sustainability campaign. Error rates were measured through active surveillance. Handoffs were assessed by means of evaluation of printed handoff documents and audio recordings. Workflow was assessed through time–motion observations. The primary outcome had two components: medical errors and preventable adverse events.

RESULTS
In 10,740 patient admissions, the medical-error rate decreased by 23% from the preintervention period to the postintervention period (24.5 vs. 18.8 per 100 admissions, P<0.001), and the rate of preventable adverse events decreased by 30% (4.7 vs. 3.3 events per 100 admissions, P<0.001). The rate of nonpreventable adverse events did not change significantly (3.0 and 2.8 events per 100 admissions, P = 0.79). Site-level analyses showed significant error reductions at six of nine sites. Across sites, significant increases were observed in the inclusion of all prespecified key elements in written documents and oral communication during handoff (nine written and five oral elements; P<0.001 for all 14 comparisons). There were no significant changes from the preintervention period to the postintervention period in the duration of oral handoffs (2.4 and 2.5 minutes per patient, respectively; P=0.55) or in resident workflow, including patient–family contact and computer time.

CONCLUSIONS
Implementation of the handoff program was associated with reductions in medical errors and in preventable adverse events and with improvements in communication, without a negative effect on workflow. (Funded by the Office of the Assistant Secretary for Planning and Evaluation, U.S. Department of Health and Human Services, and others.)
Preventable adverse events — injuries due to medical errors — are a major cause of death among Americans. Although some progress has been made in reducing certain types of adverse events, overall rates of errors remain extremely high. Failures of communication, including miscommunication during handoffs of patient care from one resident to another, are a leading cause of errors; such miscommunications contribute to two of every three “sentinel events,” the most serious events reported to the Joint Commission. The omission of critical information and the transfer of erroneous information during handoffs are common. As resident work hours have been reduced, handoffs between residents have increased in frequency.

Improving handoffs has become a priority in efforts to improve patient safety. The Accreditation Council for Graduate Medical Education now requires training programs to provide formal instruction in handoffs and to monitor handoff quality. However, few studies have rigorously evaluated the effectiveness of handoff-improvement programs.

In a single-center study, we found that the implementation of a handoff program was associated with a reduction in medical-error rates and improvements in communications between residents at change of shift. After performing this study, we developed a bundle of interventions around a refined mnemonic, I-PASS (illness severity, patient summary, action list, situation awareness and contingency plans, and synthesis by receiver). We hypothesized that multicenter implementation of the I-PASS Handoff Bundle would lead to improvements in communication and patient safety.

STUDY INSTITUTIONS
Nine pediatric residency training programs, ranging in size from 36 to 182 residents, were identified as data-collection sites through professional academic networks, as described elsewhere. Each site determined which study unit (all non–intensive care units) to include in the intervention. There was heterogeneity across sites with regard to medical complexity among patients. At baseline, no sites had a standardized handoff program in place.

INTERVENTION
We developed the I-PASS Handoff Bundle through an iterative process based on the best evidence from the literature, our previous experience, and our previously published conceptual model. The I-PASS Handoff Bundle included the following seven elements: the I-PASS mnemonic, which served as an anchoring component for oral and written handoffs and all aspects of the curriculum; a 1-hour workshop to teach TeamSTEPPS teamwork and communication skills, as well as I-PASS handoff techniques, which was highly rated; a 1-hour role-playing and simulation session for practicing skills from the workshop; a computer module to allow for independent learning; a faculty development program; direct-observation tools used by faculty to provide feedback to residents; and a process-change and culture-change campaign, which included a logo, posters, and other materials to ensure program adoption and sustainability. A detailed description of all curricular elements and the I-PASS mnemonic have been published elsewhere and are provided in Table S1 in the Supplementary Ap-
We used a well-established surveillance process to measure our two-component primary outcome: rates of medical errors (preventable failures in processes of care) and preventable adverse events (unintended consequences of medical care that lead to patient harm). We also assessed nonpreventable adverse events, which were not expected to change after the intervention. At each site, a research nurse reviewed all medical records and orders on the study unit 5 days per week (Monday reviews included a review of the weekend), formal incident reports from the hospital incident-reporting system, solicited reports from nurses working on the study unit, and daily medical-error reports from residents, collected through daily postshift surveys. Two physician investigators who were unaware of whether a given incident occurred before or after the intervention classified each suspected incident as an adverse event (i.e., harm due to medical care), a near miss or error with little potential for harm, or an exclusion (i.e., an incident determined to be neither a medical error nor an adverse event) (70% agreement; kappa, 0.47; 95% confidence interval [CI], 0.44 to 0.50). Physician reviewers further classified all adverse events as preventable (i.e., due to a medical error) or nonpreventable (i.e., due to a medical intervention with no error in the medical care delivery process) (72% agreement; kappa, 0.44; 95% CI, 0.36 to 0.52). Discordant classifications were reconciled by discussion between the paired reviewers. Examples of errors and events are provided in Tables S2A and S2B in the Supplementary Appendix.

Assessment of Written and Oral Handoffs
Each handoff consisted of both a written document and an in-person oral communication between residents. We collected copies of all written handoff documents on each weekday morning and evening at each site and audiotaped evening oral handoffs when a research assistant was present conducting time–motion observations (further details are given below). Research nurses who were aware of the intervention period evaluated a random sample of written handoff documents (a total of 432, or 24 per study period per site [half from the morning, half from the evening]) and audio recordings of oral handoffs (a total of 207, or approximately 12 per study period per site) for the presence of key handoff data elements. We compared the rates of inclusion of these elements within the document or recording for each patient before and after the intervention.

Resident Workflow Patterns and Satisfaction
We conducted time–motion observations throughout the preintervention and postintervention periods to measure the time spent by residents in various activities. Our primary interest was the time spent at the computer, conducting handoffs, and in direct patient care. To collect these data, research assistants followed individual residents for 8 to 12 hours, recording start and stop times for all activities with the use of a Microsoft Access database that included 12 major and 114 minor possible activities. Observation blocks included a representative ratio of hours from all 24 hours of the day and weekdays versus weekends. In addition, an end-of-rotation survey was administered to each resident to assess perceptions of handoff training.

Statistical Analysis
We compared medical-error rates before and after the intervention by means of Poisson regression, with a dichotomous covariate for before versus after the intervention and a fixed effect for site. We compared the percentage of written and oral handoffs (individual patient entries and discussions) that included key data elements with the use of generalized-estimating-equation z-tests that accounted for clustering based on the date of the handoff discussion or document with a fixed effect for site. To compare time–motion
data before and after the intervention, we used a generalized-estimating-equation z-test, accounting for clustering according to observation session with a fixed effect for site. This approach was based on a Dirichlet distribution, which is a distribution for the percentage of time that a continuous variable (in this case, time) is in each category. When the Bonferroni correction for multiple testing was used, two-sided P values of less than 0.025 were considered to indicate statistical significance for the two-part primary hypothesis test (postintervention change in rates of overall medical errors and postintervention change in rates of preventable adverse events across all sites). Because the other tests of hypotheses (for the main outcomes within each site as well as other outcomes overall and within each site) were more exploratory in nature, the Bonferroni correction was not used, and two-sided P values of less than 0.05 were considered to indicate statistical significance. All analyses were completed with the use of SAS/STAT software, version 9.2 (SAS Institute).

On the basis of data from our single-site study, we determined that 6 months of data collection at each site would be sufficient for more than 90% power to detect a 20% relative reduction in overall error rates and for 80% power to detect a 28% relative reduction in the rate of preventable adverse events at each site (alpha level of 0.025 with the use of a Bonferroni correction).

RESULTS

STUDY PATIENTS AND RESIDENT PHYSICIANS

We reviewed 10,740 patient admissions (5516 preintervention and 5224 postintervention) for the presence of medical errors. Length of stay, medical complexity, and the sex and age of patients did not differ significantly between the preintervention and postintervention periods, nor did the respective proportions of patients who were white (41.2% and 41.4%, P=0.38) and who were enrolled in public insurance programs (55.1% and 54.2%, P=0.61) (Table 1).

A total of 875 residents (representing 95.4% of those approached) provided written informed consent to participate. Response rates for postshift surveys used as part of medical-error surveillance were similar in the preintervention and postintervention periods (93.1% [1729 completed surveys] and 93.3% [1489 completed surveys], respectively; P=0.88).

MEDICAL ERRORS AND ADVERSE EVENTS

From the preintervention period to the postintervention period, significant reductions were observed for both components of our primary outcome: the I-PASS Handoff Bundle was associated with a 23% relative reduction in the overall medical-error rate across all sites combined (24.5 vs. 18.8 errors per 100 admissions, P<0.001) and a 30% relative reduction in the rate of preventable adverse events (4.7 vs. 3.3 events per 100 admissions, P<0.001). The rate of near misses and nonharmful medical errors decreased by 21% (19.7 vs. 15.5 near misses and nonharmful errors per 100 admissions, P<0.001) (Table 2). There was no significant change in the rate of nonpreventable adverse events (3.0 and 2.8 events per 100 admissions, P=0.79). Rates of errors that were diagnostic, related to medical history or physical examination, multifactorial, and related to therapies other than medications and procedures decreased significantly, whereas rates of errors related to medications, procedures, falls, and nosocomial infections did not change. In site-level analyses, significant reductions in error rates were observed in six of the nine participating institutions (Table 3).

WRITTEN AND ORAL HANDOFF QUALITY

The 432 written handoff documents examined yielded 5752 unique patient handoffs for evaluation (Fig. 1), and the 207 oral handoff sessions yielded 2281 unique patient handoffs (Fig. 2). I-PASS implementation was followed by significant improvements in the inclusion of all nine written handoff elements evaluated and all five oral handoff elements evaluated (see Tables S3 and S4 in the Supplementary Appendix for site-level data). The mean duration of in-person oral handoff sessions did not change significantly after the intervention (duration before and after the intervention, 2.4 and 2.5 minutes per patient, respectively; P=0.55).

RESIDENT WORKFLOW PATTERNS AND SATISFACTION

We collected 8128 hours of time-motion data (preintervention period, 3510 hours; postintervention period, 4618 hours). For all sites combined, there was no significant change in the percentage of time in a 24-hour period spent in contact with patients and families (before and after the intervention, 11.8% and 12.5%, respectively; P=0.41), creating or editing the computerized handoff document (1.6% and 1.3%, P=0.54),
Table 1. Characteristics of Patients before and after Implementation of the I-PASS Handoff Bundle, According to Site.*

<table>
<thead>
<tr>
<th>Site No.</th>
<th>Length of Stay</th>
<th>Medical Complexity†</th>
<th>Female Sex</th>
<th>Age</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Before After P Value</td>
<td>Before After P Value</td>
<td>No. of Children (%)</td>
<td>No. of Children (%)</td>
</tr>
<tr>
<td>1</td>
<td>7.9±13.7 8.1±15.0 0.22</td>
<td>316 (61.6) 272 (67.0) 0.09</td>
<td>7.0±6.2 7.7±6.5 0.11</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>13.0±25.0 11.0±18.4 0.34</td>
<td>184 (62.8) 165 (67.9) 0.22</td>
<td>4.5±5.1 5.3±5.6 0.17</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>6.0±10.6 6.2±10.5 0.80</td>
<td>293 (55.8) 317 (59.9) 0.30</td>
<td>6.7±6.0 6.4±5.8 0.24</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>2.8±5.6 3.9±13.2 0.85</td>
<td>126 (21.9) 100 (20.3) 0.52</td>
<td>4.8±5.3 4.0±4.9 0.003</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>5.0±9.7 4.1±8.3 0.29</td>
<td>583 (61.6) 567 (58.3) 0.15</td>
<td>3.7±5.4 4.1±5.7 0.58</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>3.2±5.3 3.1±5.3 0.31</td>
<td>268 (29.4) 264 (29.6) 0.95</td>
<td>7.7±6.1 7.6±6.2 0.82</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>5.1±9.6 4.9±6.4 0.88</td>
<td>266 (51.3) 205 (47.7) 0.27</td>
<td>5.2±5.7 5.0±6.1 0.23</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>3.6±5.8 3.3±6.2 0.78</td>
<td>160 (46.4) 170 (48.7) 0.54</td>
<td>7.0±6.7 6.8±6.8 0.80</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>3.2±8.7 3.1±5.9 0.51</td>
<td>180 (20.7) 215 (24.0) 0.10</td>
<td>5.2±5.7 5.1±5.8 0.89</td>
<td></td>
</tr>
<tr>
<td>Overall</td>
<td>4.9±0.13 4.8±0.14 0.59</td>
<td>2376 (43.2) 2275 (43.6) 0.40</td>
<td>5.7±0.08 5.7±0.08 0.90</td>
<td></td>
</tr>
</tbody>
</table>

* Plus–minus values are means ±SD. Within each site, demographic characteristics before and after the intervention were compared with the use of the Pearson chi-square test for dichotomous variables and the Wilcoxon rank-sum (two-sample) test for continuous variables. For all sites combined, demographic characteristics before and after the intervention were compared with the use of the Cochran-Mantel-Haenszel test for dichotomous variables and a stratified Wilcoxon test for continuous variables, to account for site effects. Data on age, length of hospital stay, sex, insurance status, and race for all patients admitted to the study unit were obtained from hospital administrative databases at each site.

† Medical complexity was defined to be present for each patient whose condition could be classified as belonging to one of three commonly published categories based on International Classification of Diseases, 9th Revision, diagnostic and procedural codes: a complex chronic condition, neurologic impairment, or a condition for which technological assistance was required.
Working at the computer (16.2% and 16.5%, \(P=0.81\)), or writing on printed copies of the handoff document (0.5% and 0.6%, \(P=0.19\)).

Significantly more residents reported having received handoff training after the intervention (60.3% before the intervention vs. 98.9% after the intervention, \(P<0.001\)). The proportion of residents who rated the overall quality of their handoff training as very good or excellent increased significantly after the intervention (27.8% before the intervention vs. 72.2% after the intervention, \(P<0.001\)).

**Discussion**

We found that implementation of the I-PASS Handoff Bundle across nine academic hospitals was associated with a 23% relative reduction in the rate of all medical errors and a 30% relative reduction in the rate of preventable adverse events. We also found significant decreases in rates of specific types of medical errors, including diagnostic errors. Site-level reductions in the overall rate of medical errors were observed at six of the nine participating sites. As anticipated, the rate of nonpreventable adverse events did not change. The quality of written and oral handoff communications significantly improved, which provided evidence that the I-PASS Handoff Bundle was successfully implemented across multiple sites and was likely to have accounted for the observed reduction in medical errors. This error reduction occurred without an increase in the time required to complete handoffs or a decrease in residents' direct contact time with patients. These findings support calls from professional and federal bodies to improve the patient-handoff process.7-9
This work builds substantially on our previous single-institution study, in which we found that implementing a prototype handoff-improvement program was associated with reductions in medical errors. We designed our current study to address several limitations of the single-center study. First, we performed a multicenter study to improve study generalizability. Second, we collected data on preintervention and postintervention error rates at the same time of year at each site, to control for potential time-of-year confounding. Third, with the help of experts at multiple sites, we simplified the mnemonic and developed a more robust curriculum to enhance the generalizability, implementation, and sustainability of the intervention.

One of the major concerns about resident duty-hour limits is that although sleep deprivation increases the risk of performance failures and medical errors, reducing work hours leads to more patient handoffs and the potential for more handoff-related errors. However, our study shows that the risk of handoff-related errors can be significantly reduced. Implementing handoff-improvement programs such as the I-PASS Handoff Bundle may potentiate the effectiveness of work-hour reductions, because doing both together may concurrently reduce both fatigue and handoff-related errors.

Our study design precludes definitively establishing a causal link between implementation of the I-PASS Handoff Bundle and improved patient safety. However, we believe it most likely that the safety improvements were due to our intervention because we saw parallel improvements in handoff processes, which was a plausible mech-
After intervention

1810 cant increases were observed in the frequency of inclusion of illness severity. Improvements were seen in every category. In the site-level analyses, significant reductions in written and oral handoff processes were observed at all sites. Because error rates are the product of numerous interacting hospital structures and processes, it is possible that institution-specific factors, such as variation in the ascertainment of error data, inconsistent implementation of the program, or other unmeasured factors, were responsible for the lack of improvement in error rates at some sites. Our study may also have been underpowered to detect improvements in error rates at some sites. Further research on the role of site-specific factors might explain these variations.

Our intervention focused on pediatric inpatient units; the extent to which the I-PASS Handoff Bundle is applicable to other disciplines, specialties, and settings is not yet known. Future studies will be required to determine the broader applicability of the intervention.

We chose to combine several educational and process changes into a single bundled intervention because numerous successful patient-safety interventions have used this approach. Although bundling appears to have been effective in this instance, it prevents us from determining which elements of the intervention were most essential.

In conclusion, we found that implementation of the I-PASS Handoff Bundle was associated with significant reductions in medical errors and preventable adverse events. Site-level changes in error rates were observed at most participating institutions.
The opinions and conclusions expressed herein are solely those of the authors and should not be construed as representing the opinions or policy of any agency of the federal government.

Supported by grants from the U.S. Department of Health and Human Services, Office of the Assistant Secretary for Planning and Evaluation (1R18AE000029-01), the Agency for Healthcare Research and Quality/Oregon Comparative Effectiveness Research K12 Program (1K12HS194956-01, to Dr. Starmer), the Medical Research Foundation of Oregon, and the Physician Services Incorporated Foundation (of Ontario, Canada) and by an unrestricted medical education grant from Pfizer.

Disclosure forms provided by the authors are available with the full text of this article at NEJM.org.

APPENDIX


From the Department of Medicine, Division of General Pediatrics, Boston Children’s Hospital (A.I.S., T.C.S., C.P.L., A.D.A., E.L.N., L.T.J.); Harvard Medical School (A.I.S., A.K.D., S.R.L., J.M.R., T.C.S., C.P.L.), Center for Patient Safety Research, Division of General Medicine (A.K.D., C.A.K., J.M.R., S.R.L., M.F.W., C.S.Y., K.R.Z.) and Division of Sleep Medicine (C.P.L.), Brigham and Women’s Hospital, and CRICO(Risk Management Foundation (C.A.K.))—all in Boston; the Department of Pediatrics, Doernbecher Children’s Hospital, Oregon Health and Science University, Portland (A.I.S., M.A.); the Department of Pediatrics, Section of General Pediatrics (N.D.S.) and Section of Critical Care (S.C.), St. Christopher’s Hospital for Children, Drexel University College of Medicine, Philadelphia; the Departments of Pediatrics (R.S., J.B.S., A.T.S.) and Neurology (F.B.), Primary Children’s Hospital, Intermountain Healthcare, University of Utah School of Medicine, and Institute for Health Care Delivery Research, Intermountain Healthcare (R.S.), Salt Lake City; the Department of Pediatrics, Benioff Children’s Hospital, University of California, San Francisco (D.C.W., G.R.), and the Department of Pediatrics, Division of General Pediatrics, Lucile Packard Children’s Hospital Stanford, Stanford University School of Medicine, Palo Alto (R.L.B., L.A.D., J.L.E., S.I.P.)—both in California; the Department of Pediatrics, Children’s Hospital Colorado, University of Colorado School of Medicine, Aurora (K.M.W.); the Department of Pediatrics, Cincinnati Children’s Hospital Medical Center, University of Cincinnati College of Medicine, Cincinnati (J.K.O., L.G.S.); the Department of Paediatics (B.Z., M.C., S.M.), Centre for Quality Improvement and Patient Safety (M.C.), and Institute for Health Policy, Management and Evaluation (S.M.), Hospital for Sick Children and University of Toronto, Toronto; the Department of Pediatrics, Division of General Pediatrics, Kapi‘olani Medical Center for Women and Children and University of Hawai‘i John A. Burns School of Medicine, Honolulu, HI (S.I.P.); the Edward Mallinckrodt Department of Pediatrics, St. Louis Children’s Hospital and Washington University School of Medicine, St. Louis (F.S.C.); the Department of Pediatrics, Texas Children’s Hospital and Baylor College of Medicine, Houston (D.F.B.); the Department of Pediatrics, Walter Reed National Military Medical Center (I.H.H., J.O.L., C.E.Y.) and the Val G. Hemmings Simulation Center (J.O.L.), Uniformed Services University of the Health Sciences, Bethesda, MD.

REFERENCES


Copyright © 2014 Massachusetts Medical Society.