



Predicting and Communicating Risk of Clinical Deterioration: An Observational Cohort Study of Internal Medicine Residents

John T. Ratelle, M.D.¹, Diana J. Kelm, M.D.², Andrew J. Halvorsen, M.S.³, Colin P. West, M.D. Ph.D.^{4,5}, and Amy S. Oxentenko, M.D.⁶

¹Division of Hospital Internal Medicine, Department of Medicine, Mayo Clinic, Rochester, MN, USA; ²Division of Pulmonary and Critical Care Medicine, Department of Medicine, Mayo Clinic, Rochester, MN, USA; ³Internal Medicine Residency Office of Educational Innovations, Department of Medicine, Mayo Clinic, Rochester, MN, USA; ⁴Division of General Internal Medicine, Department of Medicine, Mayo Clinic, Rochester, MN, USA; ⁵Division of Biomedical Statistics and Informatics, Department of Health Sciences Research, Mayo Clinic, Rochester, MN, USA; ⁶Division of Gastroenterology and Hepatology, Department of Medicine, Mayo Clinic, Rochester, MN, USA.

BACKGROUND: Despite its importance, little is known about internal medicine (IM) residents' ability to assess and communicate a patient's overnight risk during the resident-to-resident handoff.

OBJECTIVE: To evaluate IM residents' ability to identify patients at risk for clinical deterioration using the Patient Acuity Rating (PAR) tool (scored on a 1–7 symmetric scale; 1="Extremely unlikely", 7="Extremely likely"), and to measure how well IM residents conveyed a patient's potential for clinical deterioration during day-to-night handoff.

DESIGN AND PARTICIPANTS: Observational cohort study of 46 postgraduate year 1 (PGY-1) and 32 postgraduate year 3 (PGY-3) internal medicine residents rotating on one of four general medicine services from October 2013 through January 2014.

MAIN MEASURES: Primary outcomes were (1) level of agreement between resident handoff giver and receiver regarding patients' clinical risk and (2) accuracy of resident-assigned PAR score in predicting a patient's risk of clinical deterioration over the subsequent 24 hours.

KEY RESULTS: Analysis of PGY-1 giver–receiver handoff agreement revealed an intraclass correlation coefficient (ICC) (95 % CI) of 0.51 (0.45–0.56), while PGY-3 giver–receiver agreement yielded an ICC (95 % CI) of 0.42 (0.36–0.47). Based on 865 ratings of 378 patients, PGY-1 handoff giver PAR scores of 5 and 6+ were significantly associated with increased odds of clinical deterioration within 24 hours (aOR=6.5 and 12.4; $P=0.03$ and 0.005 , respectively). For the 1,170 PAR ratings of 438 patients assigned by PGY-3 handoff givers, PAR scores of 4, 5, and 6+ were significantly associated with increased odds of an event within 24 hours (aORs=6.0, 9.6, and 18.1; $P=0.03$, 0.01 , and 0.0008 , respectively).

CONCLUSIONS: The PAR is a useful tool to quantify IM residents' judgment of patient stability, and may be particularly valuable during resident handoff, given that the level of agreement between giver and receiver regarding patient risk is only fair.

KEY WORDS: Medical education-graduate; Risk communication; Risk assessment; Medical education-communication skills.

Received June 18, 2014

Revised October 13, 2014

Accepted November 3, 2014

Published online December 2, 2014

J Gen Intern Med 30(4):448–53

DOI: 10.1007/s11606-014-3114-4

© Society of General Internal Medicine 2014

INTRODUCTION

The implementation of the new duty hour regulations from the Accreditation Council for Graduate Medical Education (ACGME) in 2011 has resulted in an increased number of patient handoffs among internal medicine (IM) residents.¹ Communication failures are known to be a major cause of adverse events, and handoffs are particularly vulnerable periods for these events to occur.^{2,3} IM trainees have recognized that discontinuity in handoffs can lead to uncertainty regarding patient care and an increased risk of medical errors.^{4,5} This has led to calls for improved education and standardization of the handoff process for IM residents.^{6–8}

One of the most important components of an effective handoff is being able to convey the acuity of a patient's illness.⁸ In particular, being able to identify which patients are at risk for clinical deterioration is an integral piece of this communication. The ability to differentiate "sick" from "not sick" is one of the important features distinguishing an intern from an upper-level resident.⁹ Several clinical prediction tools have been developed to aid physicians in predicting clinical deterioration in hospitalized medical patients.^{10–12} However, even if an IM resident is able to identify a patient at risk for decompensation, if that risk is not effectively communicated to the cross-covering colleague, the potential for clinical uncertainty and adverse patient outcomes still exists. One clinical prediction tool, the Patient Acuity Rating (PAR), has shown promise during simulated handoffs as a means to convey clinical risk, but to our knowledge, this tool has not been directly evaluated in clinical practice.¹³ To better understand and potentially improve the quality of patient handoffs, we evaluated how reliably IM residents conveyed a patient's potential for clinical deterioration during

the handoff and IM residents' ability to identify patients at risk for clinical deterioration using the PAR.

METHODS

This study took place at Mayo Clinic Hospital, Saint Mary's Campus, a 1,265-bed academic medical center, and involved the Mayo Rochester IM residency program, which includes 144 categorical residents and 24 preliminary residents. We specifically aimed to evaluate the handoff quality for IM residents rotating on general medicine services from October 2013 through January 2014. This site has four general medicine resident services, comprising one postgraduate year 3 (PGY-3) and three postgraduate year 1 (PGY-1) residents (2 categorical, 1 preliminary) per service, rotating monthly. Overnight coverage is provided by two PGY-1 residents, each covering two services, and a night-float supervising PGY-3 resident covering all four services. The day-to-night handoff takes place at 6:00 p.m., at which time PGY-1 residents hand off to one another and PGY-3 residents hand off to one another. The PGY-1 and PGY-3 handoffs occur separately.

To assess the information regarding patient risk provided during the handoff, we utilized the PAR, a validated tool used to quantify physician judgment of the stability of medical patients.¹² The PAR has been shown to be reliable and to provide an accurate, objective measurement of a patient's risk of clinical deterioration over the subsequent 24-hour period. The PAR is a symmetric seven-point scale (1='Extremely unlikely', 4='Neither likely nor unlikely', 7='Extremely likely'), where a higher PAR score reflects higher perceived risk. Immediately after the day-to-night handoff, both daytime and overnight PGY-1 and PGY-3 residents were asked to independently assign a PAR score for each patient. For PGY-1 and PGY-3 residents receiving handoff, this score was to be based solely on the information provided from the daytime residents, without supplementation from the electronic medical records or other sources. Residents were instructed not to discuss PAR scores before, during, or after the handoff.

Scores were placed into a private box, and results were entered into a secure database by one of two study authors (JTR or DJK). Retrospective chart review was performed for each patient in order to obtain demographic information (age, sex, hospital day, length of stay, previous ICU stay, previous code (cardiac and/or respiratory arrest), previous rapid response team [RRT] call) and to identify patients who had experienced a clinical deterioration event, defined as RRT calls, ICU transfers, or code calls within 24 hours of handoff. Criteria for calling an RRT included an acute and persistent change in physiologic parameters (heart rate, oxygenation, etc.), signs and symptoms suggestive of myocardial ischemia or stroke, or concern on the part of any care team member regarding the patient's clinical status. While calling an RRT was not included in the original PAR, at our institution the

RRT was developed in an effort to recognize signs of deterioration, and activation of the RRT is a common and preferred method to transfer a patient to the ICU. Therefore, RRT calls were included as a clinical deterioration outcome in this study.

The daily 6:00 p.m. handoff between daytime and nighttime interns was estimated to provide up to 120 unique resident pairings of PAR scores over the duration of the study, exceeding the 44 resident sign-out pairs that would be required to achieve 90 % power to detect a 0.5 SD difference in scores at the 0.05 alpha level. Based on the published findings of the original PAR study, the team structure of our four medicine services, a four-month study period, and recent ICU transfer/Code 45/RRT rates, we estimated that we would have 90 % power to detect an odds ratio of 1.4 for the association between PAR score and occurrence of a clinical deterioration event at the 0.05 alpha level.

This study was reviewed and deemed exempt by the Mayo Clinic Institutional Review Board.

Statistical analysis

To evaluate the agreement between the handoff giver and receiver, inter-rater reliability was assessed using intraclass correlation coefficients (ICCs) and their corresponding 95 % confidence intervals (CIs). The mean and standard deviation (SD) of PAR scores by the givers and receivers were reported, and paired t tests were used to assess magnitude and direction of disagreement.

The distribution of patient sex, prevalence of prior events, and giver PAR scores were reported as counts and percentages of patients, distinct hospitalizations, and patient-days, respectively. Patient age and length of stay (LOS) were summarized using median and interquartile range (IQR).

Event prevalence was calculated within each PAR score and reported as count (%) of patient-days. There were no observed clinical deterioration events on patient-days assigned a PAR of 7 by PGY-3 givers, so ratings of 6 or 7 were grouped as "6+" a posteriori to allow statistical estimation in subsequent associative analyses.

To assess potential associations between giver PAR score and observed clinical deterioration events, logistic regression models were fit using generalized estimating equations to estimate odds ratios (ORs) of an event within 24 hours of handoff for PAR scores of 6+, 5, 4, 3, and 2 vs. 1. Available patient demographics were adjusted for simultaneously, and repeated assessments of patients were accounted for via an exchangeable correlation structure. Area under receiver operating characteristic curves (AUROCs) were reported to assess performance of PAR score thresholds as predictors of subsequent clinical deterioration.

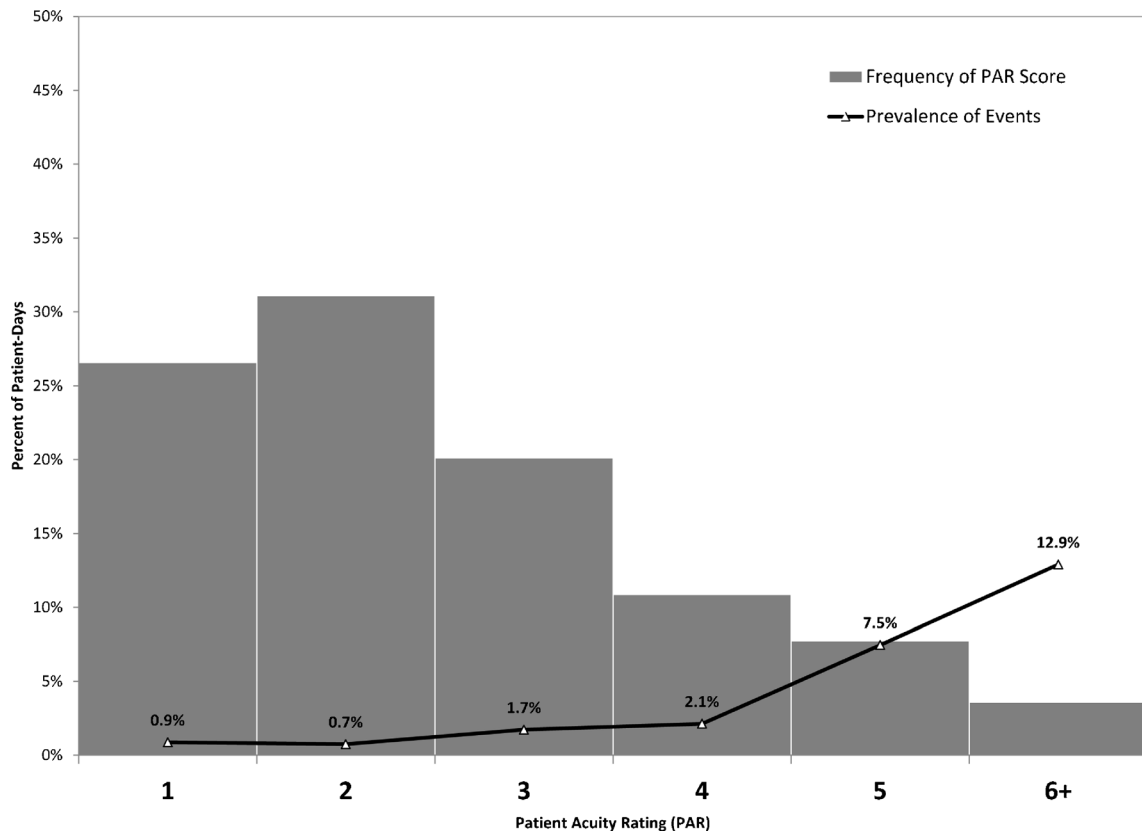


Fig. 1 Distribution of PAR scores assigned by PGY-1 residents and associated event rates

All analyses described above were performed separately by PGY (1 and 3) and reported as such below. The threshold for significance was set at 0.05 throughout. Calculations were made using SAS statistical software (version 9.3; SAS Institute Inc., Cary, North Carolina).

RESULTS

The study period included 3,967 patient-days incurred during 1,105 care episodes of 999 unique patients. There were 32 seniors and 46 interns who rotated on the four medicine teaching services during the study time frame.

Handoff agreement

PGY-1. For the interns, there were 1,197 patient-days with at least one PAR score assigned, 865 (72.2 %) by the handoff giver and 926 (77.4 %) by the handoff receiver. A total of 596 (49.8 %) had complete PAR score pairs on 295 patients, yielding an ICC (95 % CI) of 0.507 (0.445–0.564). The mean (SD) PAR scores were 2.64 (1.42) for the givers and 2.70 (1.45) for the receivers, a difference of +0.07 (1.44), which was not significant ($p=0.42$).

PGY-3. For the seniors, there were 1,794 patient-days with at least 1 PAR score assigned, 1,170 (65.2 %) by the handoff

giver and 1,520 (84.7 %) by the handoff receiver. A total of 896 (49.9 %) had complete PAR score pairs on 375 patients, yielding an ICC (95 % CI) of 0.420 (0.364–0.472). The mean (SD) PAR scores were 2.31 (1.20) for the givers and 2.49 (1.30) for the receivers, a difference of +0.18 (1.25), which was statistically significant ($p=0.003$).

Clinical deterioration

PGY-1 givers. There were 865 PAR ratings of 378 patients during 396 hospital stays by interns who were handing off. Scored patients included 195 (51.6 %) men and 183 (48.4 %)

Table 1 PAR scores and patient demographics, and the adjusted odds of clinical deterioration

	PAR	PGY-1 Givers (N=865)		PGY-3 Givers (N=1170)	
		aOR (95 % CI)*	P value	aOR (95 % CI)*	P value
	6+	12.4 (2.2–70.6)	0.005	18.1 (3.3–98.7)	0.0008
	5	6.5 (1.2–34.3)	0.03	9.6 (1.7–54.6)	0.01
	4	1.0 (0.1–8.3)	0.97	6.0 (1.2–30.8)	0.03
	3	1.4 (0.2–10.4)	0.76	2.2 (0.4–12.7)	0.40
	2	0.7 (0.09–5.5)	0.73	2.4 (0.5–11.7)	0.28
	1	–	–	–	–

*Adjusted odds ratios (aOR) from multivariable logistic regression analysis including age, sex, hospital day, length of stay, previous event (ICU stay, code (cardiac and/or respiratory arrest), rapid response team [RRT] call). Female sex reached significance for PGY-1 givers [aOR (95 % CI) = 3.0 (1.01–9.0); P value = 0.047]

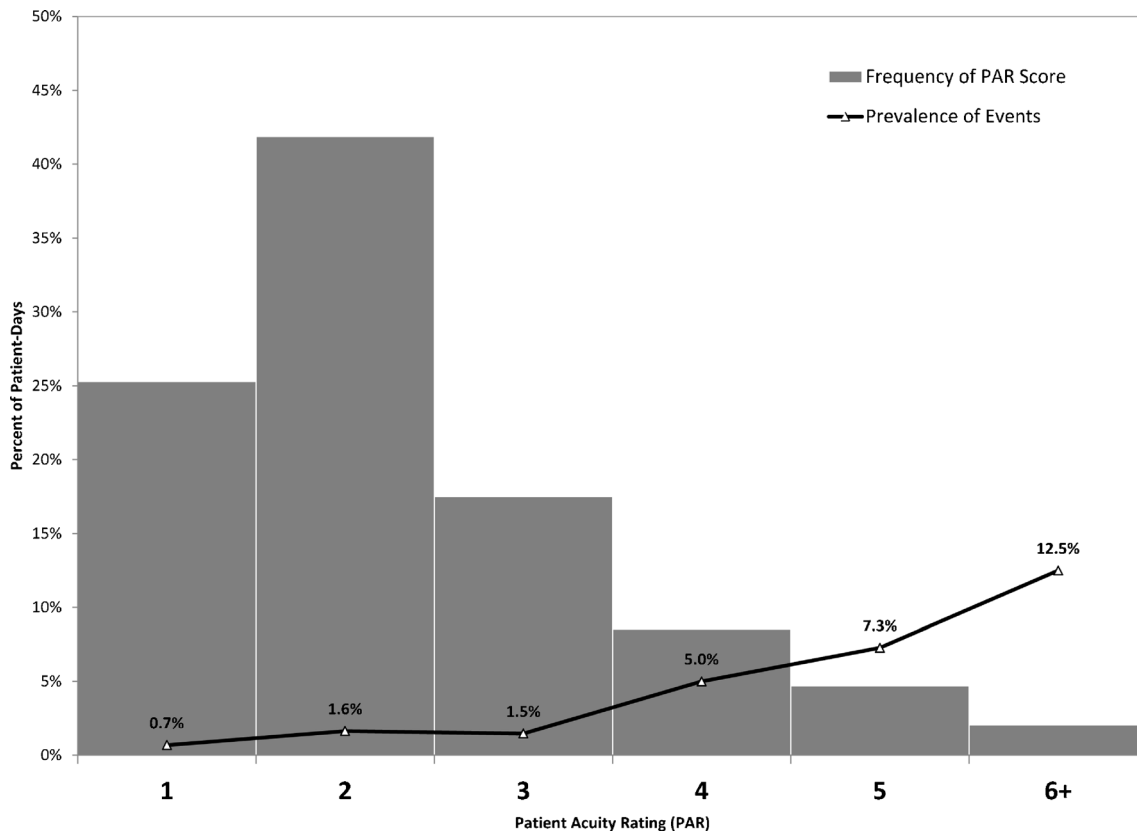


Fig. 2 Distribution of PAR scores assigned by PGY-3 residents and associated event rates.

women. The median (IQR) age was 67 (54–79) years. The median (IQR) LOS was five (3–9) days. There were 18 clinical deterioration events that occurred within 24 hours of handoff, representing 2.1 % of the 865 patient-days.

The distribution of PAR scores and their associated subsequent event rates are shown in Fig. 1. Frequency of PAR scores generally decreased as the perceived likelihood of an event increased: 230 (26.6 %) scores of 1; 269 (31.1 %) of 2; 174 (20.1 %) of 3; 94 (10.9 %) of 4; 67 (7.7 %) of 5; and 31 (3.6 %) scores of 6+. The clinical deterioration event rate generally increased with PAR score, was consistent with the sample prevalence (2.1 %) at a PAR score of 4, and showed an absolute increase of 5.4 % from both 4 to 5 (7.5 %) and 5 to 6+ (12.9 %).

For the unadjusted logistic regression model looking at event (0/1) and PAR score, PAR scores of 5 and 6+ (vs. 1) showed a significant association with the odds of a clinical deterioration event within 24 hours (ORs=9.2 and 16.9; $P=0.009$ and 0.002 , respectively). Accounting for all available variables simultaneously, as well as repeated measures within patients, a multiple logistic regression model found PAR scores of 5 and 6+ to be significantly associated with the odds of a clinical deterioration event within 24 hours (Table 1), consistent with the unadjusted findings. The AUROCs for the unadjusted and adjusted models were 0.753 and 0.821, respectively.

PGY-3 givers. There were 1,170 PAR ratings of 439 patients during 463 hospital stays by seniors who were handing off.

Scored patients included 221 (50.3 %) men and 218 (49.7 %) women. The median (IQR) age was 68 (52–81) years. The median (IQR) LOS was five (3–8) days. There were 25 clinical deterioration events that occurred within 24 hours of handoff, representing 2.1 % of the 1,170 patient-days.

The distribution of PAR scores and their associated event rates are shown in Fig. 2. The frequency of PAR scores generally decreased as the perceived likelihood of a clinical deterioration event increased: 296 (25.3 %) scores of 1; 490 (41.9 %) of 2; 205 (17.5 %) of 3; 100 (8.5 %) of 4; 55 (4.7 %) of 5; and 24 (2.1 %) scores of 6+. The clinical deterioration event rate generally increased with PAR score, ranging from 0.7 % for PAR scores of 1, to 12.5 % for scores of 6+.

For the unadjusted logistic regression model looking at clinical deterioration event (0/1) and PAR score, PAR scores of 4, 5, and 6+ (vs. 1) showed a significant association with the odds of a clinical deterioration event within 24 hours (ORs=7.7, 11.5, and 21.0; $P=0.02$, 0.005 , and 0.001 , respectively). Accounting for all variables simultaneously, as well as repeated measures within patients, a multiple logistic regression model found that PAR scores of 4, 5, and 6+ were significantly associated with the odds of a clinical deterioration event within 24 hours (Table 1), consistent with the unadjusted findings. The AUROCs for the unadjusted and adjusted models were 0.709 and 0.742, respectively.

DISCUSSION

This study demonstrated that the PAR score assigned to a patient by the day shift PGY-1 at handoff was associated with increased odds of patient deterioration within 24 hours. Specifically, PAR scores of 5 or higher were associated with a significant increase in the odds of a clinical deterioration event during the subsequent 24-hour period. Findings were similar for PAR scores of 4 or higher assigned by PGY-3 residents. Additionally, there was fair agreement between the PGY-1 day and night shift residents on risk of patient deterioration after the handoff, as demonstrated by the ICC of 0.507. However, the level of agreement was lower between PGY-3 handoff giver and receiver (ICC of 0.420), with the average PAR scores by the receiver significantly higher than those of the giver.

Our results regarding the accuracy of the PAR in predicting clinical deterioration are consistent with the findings in the original study on the PAR.¹² This provides external validation for the use of the PAR as a tool to objectively quantify IM resident perception of clinical risk.

The reasons for the lower level of agreement in patient risk assessment for PGY-3 residents relative to PGY-1 residents are not clear. One explanation for the difference in agreement between interns and seniors may be the effects of patient volume and workload. In our call structure, the overnight PGY-3 receives twice as many patient handoffs as the overnight PGY-1. Resident workload has been shown to have an influence on resident performance and patient outcomes.^{14–17} Additionally, a recent study of physician handoff practices showed that larger numbers of patient care handoffs were associated with a greater number of interruptions during handoff.¹⁸ These interruptions are associated with increased opportunity for overlooked or omitted patient information, and may explain the greater level of discordance between PGY-3 giver and receiver.¹⁹ Additional potential confounding factors include differences in handoff practices between interns and seniors, as well as the structure of the rotation call schedule. We plan to further evaluate these findings as we continue to improve and standardize the handoff practices of our residents.

Our project had several limitations. First, the format and content of patient handoffs was not evaluated, and is likely to vary among trainees in our program. Additionally, we did not assess for resident attitudes regarding the utility of the PAR, and whether quantifying a score influenced clinical decision-making for the overnight resident, as these scores were assigned independently by providers and not shared between them. However, a prior study of the PAR in simulated handoff experiences demonstrated that the addition of the PAR could influence cross-cover residents' clinical decision-making.¹³ An absence of observed deterioration events at the highest PAR value (7) in one of the groups (PGY-3 givers) necessitated a grouping of the two highest rating values (6 and 7) for associative analyses. Since the data were collected using the

original seven-point scale, the impact of this consolidation on the findings was likely small. We included activation of the RRT in the outcome of clinical deterioration, which was not done during the validation of the PAR. However, as previously noted, we deemed inclusion of RRT outcomes consistent with the intent of the PAR score, given their relationship to care team concerns about patient status. Lastly, this was a single-institution study, and event rates and practice standards (e.g., role of the RRT) will differ at other hospitals, which may limit the generalizability of our findings. Further study of the impact of the PAR tool on overnight resident attitudes, decision-making, and patient outcomes is needed.

In summary, this study provides evidence from clinical practice in support of the PAR as an accurate means to quantify patient stability. These findings help support its role in standardized handoff practices among internal medicine residents, particularly given that the level of agreement between resident handoff giver and receiver regarding clinical risk is only fair.

Acknowledgments: This study was supported in part by the Mayo Clinic Internal Medicine Residency Office of Educational Innovations as part of the ACGME Educational Innovations Project.

Conflicts of Interest: The authors each declare that they have no conflict of interest.

Corresponding Author: Amy S. Oxentenko, M.D.; Division of Gastroenterology and Hepatology, Department of Medicine Mayo Clinic, 200 First Street SW, Rochester, MN 55905, USA (e-mail: Oxentenko.Amy@mayo.edu).

REFERENCES

- Desai SV, Feldman L, Brown L, et al. Effect of the 2011 vs 2003 duty hour regulation-compliant models on sleep duration, trainee education, and continuity of patient care among internal medicine house staff: a randomized trial. *JAMA Intern Med.* 2013;173:649–55.
- The Joint Commission. Sentinel event data root causes by event type 2004 – June 2013 available at: http://www.jointcommission.org/assets/1/18/Root_Causes_by_Event_Type_2004-2Q2013.pdf. Accessed November 4, 2014.
- Horwitz LI, Moin T, Krumholz HM, Wang L, Bradley EH. Consequences of inadequate sign-out for patient care. *Arch Intern Med.* 2008;168:1755–60.
- Fletcher KE, Parekh V, Halasyamani L, et al. Work hour rules and contributors to patient care mistakes: a focus group study with internal medicine residents. *J Hosp Med.* 2008;3:228–37.
- Arora V, Johnson J, Lovinger D, Humphrey HJ, Meltzer DO. Communication failures in patient sign-out and suggestions for improvement: a critical incident analysis. *Qual Saf Health Care.* 2005;14:401.
- Wohlauer MV, Arora VM, Horwitz LI, Bass EJ, Mahar SE, Philibert I. The patient handoff: a comprehensive curricular blueprint for resident education to improve continuity of care. *Acad Med.* 2012;87:411–8.
- Cohen MD, Hilligoss PB. The published literature on handoffs in hospitals: deficiencies identified in an extensive review. *Qual Saf Health Care.* 2010;19:493–7.
- Vidarthi AR, Arora V, Schnipper JL, Wall SD, Wachter RM. Managing discontinuity in academic medical centers: strategies for a safe and effective resident sign-out. *J Hosp Med.* 2006;1:257–66.
- Wiese J. *Teaching in the Hospital.* Philadelphia: ACP Press; 2010.
- Rothman MJ, Rothman SI, Beals J 4th. Development and validation of a continuous measure of patient condition using the Electronic Medical Record. *J Biomed Inform.* 2013;46:837–48.
- Kirkland LL, Malinchoc M, O'Byrne M, et al. A clinical deterioration prediction tool for internal medicine patients. *Am J Med Qual.* 2013;28:135–42.

12. **Edelson DP, Retzer E, Weidman EK, et al.** Patient acuity rating: quantifying clinical judgment regarding inpatient stability. *J Hosp Med.* 2011;6:475–9.
13. **Phillips AW, Yuen TC, Retzer E, Woodruff J, Arora V, Edelson DP.** Supplementing cross-cover communication with the patient acuity rating. *J Gen Intern Med.* 2013;28:406–11.
14. **Coit MH, Katz JT, McMahon GT.** The effect of workload reduction on the quality of residents' discharge summaries. *J Gen Intern Med.* 2011;26:28–32.
15. **Ong M, Bostrom A, Vidyarthi A, McCulloch C, Auerbach A.** House staff team workload and organization effects on patient outcomes in an academic general internal medicine inpatient service. *Arch Intern Med.* 2007;167:47–52.
16. **Averbukh Y, Southern W.** The impact of the number of admissions to the inpatient medical teaching team on patient safety outcomes. *J Grad Med Educ.* 2012;4:307–11.
17. **Thanarajasingam U, McDonald FS, Halvorsen AJ, et al.** Service census caps and unit-based admissions: resident workload, conference attendance, duty hour compliance, and patient safety. *Mayo Clin Proc.* 2012;87:320–7.
18. **Greenstein EA, Arora VM, Staisiunas PG, Banerjee SS, Farnan JM.** Characterising physician listening behaviour during hospitalist handoffs using the HEAR checklist. *BMJ Qual Saf.* 2013;22:203–9.
19. **McGuillan A, Carthey J, Catchpole K, McCulloch P, Ridout DA, Goldman AP.** Creating a safe, reliable hospital at night handover: a case study in implementation science. *BMJ Qual Saf.* 2013.