

ORIGINAL RESEARCH

Using Standardized Videos to Validate a Measure of Handoff Quality: The Handoff Mini-Clinical Examination Exercise

Vineet M. Arora, MD, MAPP^{1,2*}, Saba Berhie, BA², Leora I. Horwitz, MD, MSH³, Mark Saathoff, MA¹, Paul Stasiunas, BA⁴, Jeanne M. Faman, MD, MHPE^{1,2}

¹Department of Medicine, University of Chicago, Chicago, Illinois; ²Pritzker School of Medicine, University of Chicago, Chicago, Illinois;

³Department of Medicine, Yale University, New Haven, Connecticut; ⁴Marquette University Law School, Milwaukee, Wisconsin.

BACKGROUND: The most recent iteration of the Accreditation Council for Graduate Medical Education duty-hour regulations includes language mandating handoff education for trainees and assessments of handoff quality by residency training programs. However, there is a lack of validated tools for the assessment of handoff quality and for use in trainee education.

METHODS: Faculty at 2 sites (University of Chicago and Yale University) were recruited to participate in a workshop on handoff education. Video-based scenarios were developed to represent varying levels of performance in the domains of communication, professionalism, and setting. Videos were shown in a random order, and faculty were instructed to use the Handoff Mini-Clinical Examination Exercise (CEX), a paper-based instrument with qualitative anchors defining each level of performance, to rate the handoffs.

RESULTS: Forty-seven faculty members (14 at site 1; 33 at site 2) participated in the validation workshops, providing a total of 172 observations (of a possible 191 [96%]). Reliability testing revealed a Cronbach α of 0.81 and Kendall coefficient of concordance of 0.59 ($>0.6 =$ high reliability). Faculty were able to reliably distinguish the different levels of performance in each domain in a statistically significant fashion (ie, unsatisfactory professionalism mean 2.42 vs satisfactory professionalism 4.81 vs superior professionalism 6.01, $P < 0.001$ trend test). Two-way analysis of variance revealed no evidence of rater bias.

CONCLUSIONS: Using standardized video-based scenarios highlighting differing levels of performance, we were able to demonstrate evidence that the Handoff Mini-CEX can draw reliable and valid conclusions regarding handoff performance. Future work to validate the tool in clinical settings is warranted. *Journal of Hospital Medicine* 2014;9:441–446. © 2014 Society of Hospital Medicine

Over the last decade, there has been an unprecedented focus on physician handoffs in US hospitals. One major reason for this are the reductions in residency duty hours that have been mandated by the American Council for Graduate Medical Education (ACGME), first in 2003 and subsequently revised in 2011.^{1,2} As residents work fewer hours, experts believe that potential safety gains from reduced fatigue are countered by an increase in the number of handoffs, which represent a risk due to the potential miscommunication. Prior studies show that critical patient information is often lost or altered during this transfer of clinical information and professional responsibility, which can result in patient harm.^{3,4} As a result of these concerns, the ACGME now requires residency programs to “ensure and monitor effective, structured hand-over processes to facilitate both continuity of care and patient safety. Programs must ensure that

residents are competent in communicating with team members in the hand-over process.”² Moreover, handoffs have also been a major improvement focus for organizations with broader scope than teaching hospitals, including the World Health Organization, Joint Commission, and the Society for Hospital Medicine (SHM).^{5–7}

Despite this focus on handoffs, monitoring quality of handoffs has proven challenging due to lack of a reliable and validated tool to measure handoff quality. More recently, the Accreditation Council of Graduate Medical Education’s introduction of the Next Accreditation System, with its focus on direct observation of clinical skills to achieve milestones, makes it crucial for residency educators to have valid tools to measure competence in handoffs. As a result, it is critical that instruments to measure handoff performance are not only created but also validated.⁸

To help fill this gap, we previously reported on the development of a 9-item Handoff Clinical Examination Exercise (CEX) assessment tool. The Handoff CEX, designed for use by those participating in the handoff or by a third-party observer, can be used to rate the quality of patient handoffs in domains such as professionalism and communication skills between the receiver and sender of patient information.^{9,10} Despite prior demonstration of feasibility of use, the initial tool was perceived as lengthy and redundant.

*Address for correspondence and reprint requests: Vineet Arora, MD, 5841 South Maryland Ave., MC 2007, AMB W216, Chicago, IL 60637; Telephone: 773-702-8157; Fax: 773-834-2238; E-mail: varora@medicine.bsdc.uchicago.edu

Additional Supporting Information may be found in the online version of this article.

Received: November 20, 2013; Revised: February 25, 2014; Accepted: February 28, 2014

2014 Society of Hospital Medicine DOI 10.1002/jhm.2185

Published online in Wiley Online Library (Wileyonlinelibrary.com).

In addition, although the tool has been shown to discriminate between performance of novice and expert nurses, the construct validity of this tool has not been established.¹¹ Establishing construct validity is important to ensuring that the tool can measure the construct in question, namely whether it detects those who are actually competent to perform handoffs safely and effectively. We present here the results of the development of a shorter Handoff Mini-CEX, along with the formal establishment of its construct validity, namely its ability to distinguish between levels of performance in 3 domains of handoff quality.

METHODS

Adaption of the Handoff CEX and Development of the Abbreviated Tool

The 9-item Handoff CEX is a paper-based instrument that was created by the investigators (L.I.H., J.M.F., V.M.A.) to evaluate either the sender or the receiver of handoff communications and has been used in prior studies (see Supporting Information, Appendix 1, in the online version of this article).^{9,10} The evaluation may be conducted by either an observer or by a handoff participant. The instrument includes 6 domains: (1) setting, (2) organization and efficiency, (3) communication skills, (4) content, (5) clinical judgment, and (6) humanistic skills/professionalism. Each domain is graded on a 9-point rating scale, modeled on the widely used Mini-CEX (Clinical Evaluation Exercise) for real-time observation of clinical history and exam skills in internal medicine clerkships and residencies (1–3 = unsatisfactory, 4–6 = marginal/satisfactory, 7–9 = superior).¹² This familiar 9-point scale is utilized in graduate medical education evaluation of the ACGME core competencies.

To standardize the evaluation, the instrument uses performance-based anchors for evaluating both the sender and the receiver of the handoff information. The anchors are derived from functional evaluation of the roles of senders and receivers in our preliminary work at both the University of Chicago and Yale University, best practices in other high-reliability industries, guidelines from the Joint Commission and the SHM, and prior studies of effective communication in clinical systems.^{5,6,13}

After piloting the Handoff CEX with the University of Chicago's internal medicine residency program ($n = 280$ handoff evaluations), a strong correlation was noted between the measures of content (medical knowledge), patient care, clinical judgment, organization/efficiency, and communication skills. Moreover, the Handoff CEX's Cronbach α , or measurement of internal reliability and consistency, was very high ($\alpha = 0.95$). Given the potential of redundant items, and to increase ease of use of the instrument, factor analysis was used to reduce the instrument to yield a shorter 3-item tool, the Handoff Mini-CEX, that

assessed 3 of the initial items: setting, communication skills, and professionalism. Overall, performance on these 3 items were responsible for 82% of the variance of overall sign-out quality (see Supporting Information, Appendix 2, in the online version of this article).

Establishing Construct Validity of the Handoff Mini-CEX

To establish construct validity of the Handoff Mini-CEX, we adapted a protocol used by Holmboe and colleagues to report the construct validity of the Handoff Mini-CEX, which is based on the development and use of video scenarios depicting varying levels of clinical performance.¹⁴ A clinical scenario script, based on prior observational work, was developed, which represented an internal medicine resident (the sender) signing out 3 different patients to colleagues (intern [postgraduate year 1] and resident). This scenario was developed to explicitly include observable components of professionalism, communication, and setting. Three levels of performance—superior, satisfactory, and unsatisfactory—were defined and described for the 3 domains. These levels were defined, and separate scripts were written using this information, demonstrating varying levels of performance in each of the domains of interest, using the descriptive anchors of the Handoff Mini-CEX.

After constructing the superior, or gold standard, script that showcases superior communication, professionalism, and setting, individual domains of performance were changed (eg, to satisfactory or unsatisfactory), while holding the other 2 constant at the superior level of performance. For example, superior communication requires that the sender provides anticipatory guidance and includes clinical rationale, whereas unsatisfactory communication includes vague language about overnight events and a disorganized presentation of patients. Superior professionalism requires no inappropriate comments by the sender about patients, family, and staff as well as a presentation focused on the most urgent patients. Unsatisfactory professionalism is shown by a hurried and inattentive sign-out, with inappropriate comments about patients, family, and staff. Finally, a superior setting is one in which the receiver is listening attentively and discourages interruptions, whereas an unsatisfactory setting finds the sender or receiver answering pages during the handoff surrounded by background noise. We omitted the satisfactory level for setting due to the difficulties in creating subtleties in the environment.

Permutations of each of these domains resulted in 6 scripts depicting different levels of sender performance (see Supporting Information, Appendix 3, in the online version of this article). Only the performance level of the sender was changed, and the receivers of the handoff performance remained consistent, using

TABLE 1. Script Matrix

	Unsatisfactory	Satisfactory	Superior
Communication	Script 3 (n = 36)* Uses vague language about overnight events, missing critical patient information, disorganized. "Look in the record; I'm sure it's in there. And oh yeah, I need you to check enzymes and finish ruling her out."	Script 2 (n = 13) Insufficient level of clinical detail, directions are not as thorough, handoff is generally on task and sufficient. "So the only thing to do is to check labs; you know, check CBC and cardiac enzymes."	Script 1 (n = 13) Anticipatory guidance provided, rationale explained; important information is included, highlights "sick" patients. "So for today, I need you to check post-transfusion hemoglobin to make sure it's back to the baseline of 10. If it's under 10, then transfuse her 2 units, but hopefully it will be bumped up. Also continue to check cardiac enzymes; the next set is coming at 2 PM, and we need to continue the rule out. If her enzymes are positive or she has other ECG changes, definitely call the cardio fellow, since they'll want to take her to the CCU."
Professionalism	Script 5 (n = 39)* Hurried, inattentive, rushing to leave, inappropriate comments (re: patients, family, staff). "[D]efinitely call the cards fellow, since they'll want to take her to the CCU. And let me tell you, if you don't call her, she'll rip you a new one."	Script 4 (n = 22)* Some tangential comments (re: patients, family, staff). "Let's breeze through them quickly so I can get out of here, I've had a rough day. I'll start with the sickest first, and oh my God she's a train wreck!"	Script 1 Appropriate comments (re: patients, family, staff), focused on task.
Setting	Script 6 (n = 13) Answering pages during handoff, interruptions (people entering room, phone ringing).		Script 1 Attentive listening, no interruptions, pager silenced.

NOTE: Abbreviations: CBC, complete blood count; CCU, coronary care unit; ECG, electrocardiogram.

*Denotes video scenario seen by Yale University raters. All videos were seen by University of Chicago raters.

best practices for receivers, such as attentive listening, asking questions, reading back, and taking notes during the handoff. The scripts were developed by 2 investigators (V.M.A., S.B.), then reviewed and edited independently by other investigators (J.M.F., P.S.) to achieve consensus. Actors were recruited to perform the video scenarios and were trained by the physician investigators (J.M.F., V.M.A.). The part of the sender was played by a study investigator (P.S.) with prior acting experience, and who had accrued over 40 hours of experience observing handoffs to depict varying levels of handoff performance. The digital video recordings ranged in length from 2.00 minutes to 4.08 minutes. All digital videos were recorded using a Sony XDCAM PMW-EX3 HD camcorder (Sony Corp., Tokyo, Japan).

Participants

Faculty from the University of Chicago Medical Center and Yale University were included. At the University of Chicago, faculty were recruited to participate via email by the study investigators to the Research in Medical Education (RIME) listhost, which includes program directors, clerkship directors, and medical educators. Two sessions were offered and administered. Continuing medical education (CME) credit was provided for participation, as this workshop was given in conjunction with the RIME CME conference. Evaluations were deidentified using a unique identifier

for each rater. At Yale University, the workshop on handoffs was offered as part of 2 seminars for program directors and chief residents from all specialties. During these seminars, program directors and chief residents used anonymous evaluation rating forms that did not capture rater identifiers. No other incentive was provided for participation. Although neither faculty at the University of Chicago nor Yale University received any formal training on handoff evaluation, they did receive a short introduction to the importance of handoffs and the goals of the workshop. The protocol was deemed exempt by the institutional review board at the University of Chicago.

Workshop Protocol

After a brief introduction, faculty viewed the tapes in random order on a projected screen. Participants were instructed to use the Handoff Mini-CEX to rate whichever element(s) of handoff quality they believed they could suitably evaluate while watching the tapes. The videos were rated on the Handoff Mini-CEX form, and participants anonymously completed the forms independently without any contact with other participants. The lead investigators proctored all sessions. At University of Chicago, participants viewed and rated all 6 videos over the course of an hour. At Yale University, due to time constraints in the program director and chief resident seminars, participants reviewed 1 of the videos in seminar 1 (unsatisfactory

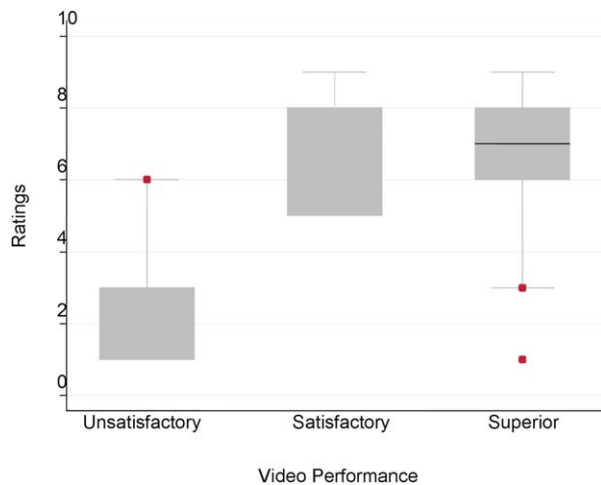


FIG. 1. Faculty ratings of communication by performance. The handoff Clinical Examination Exercise ratings are a 9-point scale: 1–3 = unsatisfactory, 4–6 = satisfactory, 7–9 = superior.

professionalism) and 2 in the other seminar (unsatisfactory communication, unsatisfactory professionalism) (Table 1).

Data Collection and Statistical Analysis

Using combined data from University of Chicago and Yale University, descriptive statistics were reported as raw scores on the Handoff Mini-CEX. To assess internal consistency of the tool, Cronbach α was used. To assess inter-rater reliability of these attending physician ratings on the tool, we performed a Kendall coefficient of concordance analysis after collapsing the ratings into 3 categories (unsatisfactory, satisfactory, superior). In addition, we also calculated intraclass correlation coefficients for each item using the raw data and generalizability analysis to calculate the number of raters that would be needed to achieve a desired reliability of 0.95. To ascertain if faculty were able to detect varying levels of performance depicted in the video, an ordinal test of trend on the communication, professionalism, and setting scores was performed.

To assess for rater bias, we were able to use the identifiers on the University of Chicago data to perform a 2-way analysis of variance (ANOVA) to assess if faculty scores were associated with performance level after controlling for faculty. The results of the faculty rater coefficients and P values in the 2-way ANOVA were also examined for any evidence of rater bias. All calculations were performed in Stata 11.0 (StataCorp, College Station, TX) with statistical significance defined as $P < 0.05$.

RESULTS

Forty-seven faculty members (14 = site 1; 33 = site 2) participated in the validation workshops (2 at the University of Chicago, and 2 at Yale University), which were held in August 2011 and September 2011, providing a total of 172 observations of a possible 191 (90%).

The overall handoff quality ratings for the superior, gold standard video (superior communication, profes-

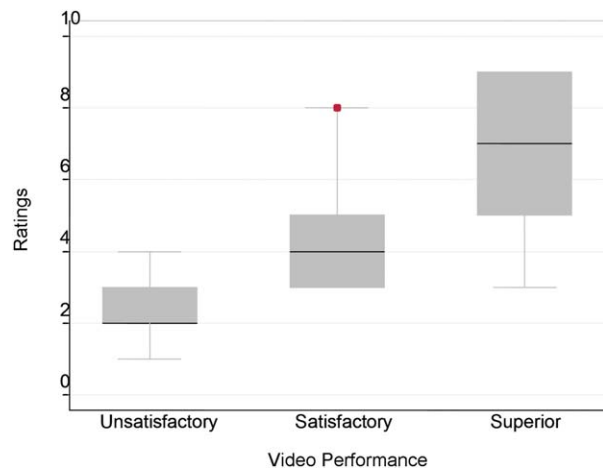


FIG. 2. Faculty ratings of professionalism by performance. The handoff Clinical Examination Exercise ratings are a 9-point scale: 1–3 = unsatisfactory, 4–6 = satisfactory, 7–9 = superior.

sionalism, and communication) ranged from 7 to 9 with a mean of 8.5 (standard deviation [SD] 0.7). The overall ratings for the video depicting satisfactory communication (satisfactory communication, superior professionalism and setting) ranged from 5 to 9 with a mean of 7.3 (SD 1.1). The overall ratings for the unsatisfactory communication (unsatisfactory communication, superior professionalism and setting) video ranged from 1 to 7 with a mean of 2.6 (SD 1.2). The overall ratings for the satisfactory professionalism video (satisfactory professionalism, superior communication and setting) ranged from 4 to 8 with a mean of 5.7 (SD 1.3). The overall ratings for the unsatisfactory professionalism (unsatisfactory professionalism, superior communication and setting) video ranged from 2 to 5 with a mean of 2.4 (SD 1.03). Finally, the overall ratings for the unsatisfactory setting (unsatisfactory setting, superior communication and professionalism) video ranged from 1 to 8 with a mean of 3.1 (SD 1.7).

Figure 1 demonstrates that for the domain of communication, the raters were able to discern the unsatisfactory performance but had difficulty reliably distinguishing between superior and satisfactory performance. Figure 2 illustrates that for the domain of professionalism, raters were able to detect the videos' changing levels of performance at the extremes of behavior, with unsatisfactory and superior displays more readily identified. Figure 3 shows that for the domain of setting, the raters were able to discern the unsatisfactory versus superior level of the changing setting. Of note, we also found a moderate significant correlation between ratings of professionalism and communication ($r = 0.47$, $P < 0.001$).

The Cronbach α , or measurement of internal reliability and consistency, for the Handoff Mini-CEX (3 items plus overall) was 0.77, indicating high internal reliability and consistency. Using data from University of Chicago, where raters were labeled with a unique identifier, the Kendall coefficient of concordance was

calculated to be 0.79, demonstrating high inter-rater reliability of the faculty raters. High inter-rater reliability was also seen using intraclass coefficients for each domain: communication (0.84), professionalism (0.68), setting (0.83), and overall (0.89). Using generalizability analysis, the average reliability was determined to be above 0.9 for all domains (0.99 for overall).

Last, the 2-way ANOVA ($n = 75$ observations from 13 raters) revealed no evidence of rater bias when examining the coefficient for attending rater ($P = 0.55$ for professionalism, $P = 0.45$ for communication, $P = 0.92$ for setting). The range of scores for each video, however, was broad (Table 2).

DISCUSSION

This study demonstrates that valid conclusions on handoff performance can be drawn using the Handoff CEX as the instrument to rate handoff quality. Utilizing standardized videos depicting varying levels of performance communication, professionalism, and setting, the Handoff Mini-CEX has demonstrated potential to discern between increasing levels of performance, providing evidence for the construct validity of the instrument.

We observed that faculty could reliably detect unsatisfactory professionalism with ease, and that there was a distinct correlation between faculty ratings and the internally set levels of performance displayed in the videos. This

trend demonstrated that faculty were able to discern different levels of professionalism using the Handoff Mini-CEX. It became more difficult, however, for faculty to detect superior professionalism when the domain of communication was permuted. If the sender of the handoff was professional but the information delivered was disorganized, inaccurate, and missing crucial pieces of information, the faculty perceived this ineffective communication as unprofessional. Prior literature on professionalism has found that communication is a necessary component of professional behavior, and consequently, being a competent communicator is necessary to fulfill one's duty as a professional physician.^{15,16}

This is of note because we did find a moderate significant correlation between ratings of professionalism and communication. It is possible that this distinction would be made clearer with formal rater training in the future prior to any evaluations. However, it is also possible that professionalism and communication, due to a synergistic role between the 2 domains, cannot be separated. If this is the case, it would be important to educate clinicians to present patients in a concise, clear, and accurate way with a professional demeanor. Acknowledging professional responsibility as an integral piece of patient care is also critical in effectively communicating patient information.⁵

We also noted that faculty could detect unsatisfactory communication consistently; however, they were unable to differentiate between satisfactory and superior communication reliably or consistently. Because the unsatisfactory professionalism, unsatisfactory setting, and satisfactory professionalism videos all demonstrated superior communication, we believe that the faculty penalized communication when distractions, in the form of interruptions and rude behavior by the resident giving the handoff, interrupted the flow of the handoff. Thus, the wide ranges in scores observed by some raters may be attributed to this interaction between the Handoff Mini-CEX domains. In the future, definitions of the anchors, including at the middle spectrum of performance, and rater training may improve the ability of raters to distinguish performance between each domain.

The overall value of the Handoff Mini-CEX is in its ease of use, in part due to its brevity, as well as

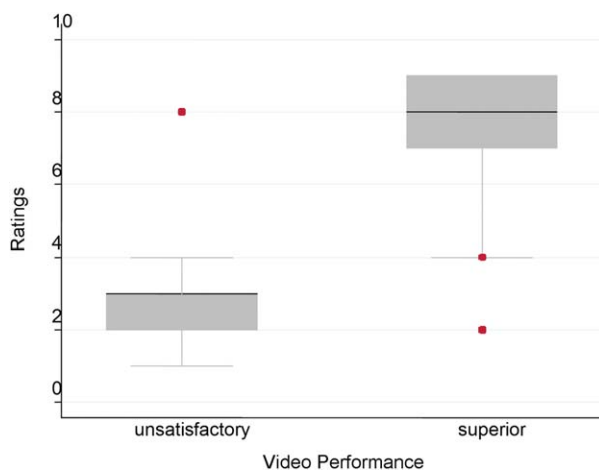


FIG. 3. Faculty ratings of setting by performance. The handoff Clinical Examination Exercise ratings are a 9-point scale: 1–3 = unsatisfactory, 4–6 = satisfactory, 7–9 = superior.

TABLE 2. Faculty's Mini-Handoff Clinical Examination Exercise Ratings by Level of Performance Depicted in Video*

	Unsatisfactory			Satisfactory			Superior			P^{\dagger}
	Mean	Median	Range	Mean	Median	Range	Mean	Median	Range	
Professionalism	2.3	2	1–4	4.4	4	3–8	7.0	7	3–9	0.026
Communication	2.8	3	1–6	7	8	5–9	6.6	7	1–9	0.005
Setting	3.1	3	1–8				7.5	8	2–9	0.005

*NOTE: Clinical Examination Exercise ratings are on a 9-point scale: 1–3 = unsatisfactory, 4–6 = satisfactory, 7–9 = superior.

[†] P value is from 2-way analysis of variance examining the level of performance on rating of that construct controlling for rater.

evidence for its validity in distinguishing between varying levels of performance. Given the emphasis on monitoring handoff quality and performance, the Handoff Mini-CEX provides a standard foundation from which baseline handoff performance can be easily measured and improved. Moreover, it can also be used to give individual feedback to a specific practicing clinician on their practices and an opportunity to improve. This is particularly important given current recommendations by the Joint Commission, that handoffs are standardized, and by the ACGME, that residents are competent in handoff skills. Moreover, given the creation of the SHM's handoff recommendations and handoffs as a core competency for hospitalists, the tool provides the ability for hospitalist programs to actually assess their handoff practices as baseline measurements for any quality improvement activities that may take place.

Faculty were able to discern the superior and unsatisfactory levels of setting with ease. After watching and rating the videos, participants said that the chaotic scene of the unsatisfactory setting video had significant authenticity, and that they were constantly interrupted during their own handoffs by pages, phone calls, and people entering the handoff space. System-level fixes, such as protected time and dedicated space for handoffs, and discouraging pages to be sent during the designated handoff time, could mitigate the reality of unsatisfactory settings.^{17,18}

Our study has several limitations. First, although this study was held at 2 sites, it included a small number of faculty, which can impact the generalizability of our findings. Implementation varied at Yale University and the University of Chicago, preventing use of all data for all analyses. Furthermore, institutional culture may also impact faculty raters' perceptions, so future work aims at repeating our protocol at partner institutions, increasing both the number and diversity of participants. We were also unable to compare the new shorter Handoff Mini-CEX to the larger 9-item Handoff CEX in this study.

Despite these limitations, we believe that the Handoff Mini-CEX, has future potential as an instrument with which to make valid and reliable conclusions about handoff quality, and could be used to both evaluate handoff quality and as an educational tool for trainees and faculty on effective handoff communication.

Disclosures: This work was supported by the National Institute on Aging Short-Term Aging-Related Research Program (5T35AG029795), Agency for Healthcare Research and Quality (1 R03HS018278-01), and the University of Chicago Department of Medicine Excellence in Medical Education Award. Dr. Horwitz is supported by the National Institute on Aging (K08 AG038336) and by the American Federation for Aging Research through the Paul B. Beeson Career Development Award Program. Dr. Arora is funded by National Institute on Aging Career Development Award K23AG033763. Prior presentations of these data include the 2011 Association of American Medical Colleges meeting in Denver, Colorado, the 2012 Association of Program Directors of Internal Medicine meeting in Atlanta, Georgia, and the 2012 Society of General Internal Medicine Meeting in Orlando, Florida.

References

1. Nasca TJ, Day SH, Amis ES. The new recommendations on duty hours from the ACGME task force. *New Engl J Med*. 2010;363(2):e3.
2. ACGME common program requirements. Effective July 1, 2011. Available at: [http://www.acgme.org/acgmeweb/Portals/0/PDFs/Common_Program_Requirements_07012011\[2\].pdf](http://www.acgme.org/acgmeweb/Portals/0/PDFs/Common_Program_Requirements_07012011[2].pdf). Accessed February 8, 2014.
3. Horwitz LI, Moin T, Krumholz HM, Wang L, Bradley EH. Consequences of inadequate sign-out for patient care. *Arch Intern Med*. 2008;168(16):1755-1760.
4. 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 Healthcare*. 2005;14(6):401-407.
5. Arora VM, Manjarrez E, Dressler DD, Basaviah P, Halasyamani L, Kripalani S. Hospitalist handoffs: a systematic review and task force recommendations. *J Hosp Med*. 2009;4(7):433-440.
6. Arora V, Johnson J. A model for building a standardized hand-off protocol. *Jt Comm J Qual Patient Saf*. 2006;32(11):646-655.
7. World Health Organization Collaborating Centre for Patient Safety. Solutions on communication during patient hand-overs. 2007; Volume 1, Solution 1. Available at: <http://www.who.int/patientsafety/solutions/patientsafety/PS-Solution3.pdf>. Accessed February 8, 2014.
8. Patterson ES, Wears RL. Patient handoffs: standardized and reliable measurement tools remain elusive. *Jt Comm J Qual Patient Saf*. 2010;36(2):52-61.
9. Horwitz L, Rand D, Staisiunas P, et al. Development of a handoff evaluation tool for shift-to-shift physician handoffs: the handoff CEX. *J Hosp Med*. 2013;8(4):191-200.
10. Farnan JM, Paro JAM, Rodriguez RM, et al. Hand-off education and evaluation: piloting the observed simulated hand-off experience (OSHE). *J Gen Intern Med*. 2010;25(2):129-134.
11. Horwitz LI, Dombroski J, Murphy TE, Farnan JM, Johnson JK, Arora VM. Validation of a handoff tool: the Handoff CEX. *J Clin Nurs*. 2013;22(9-10):1477-1486.
12. Norcini JJ, Blank LL, Duffy FD, Fortna GS. The mini-CEX: a method for assessing clinical skills. *Ann Intern Med*. 2003;138(6):476-481.
13. Patterson ES, Roth EM, Woods DD, Chow R, Gomes JO. Handoff strategies in settings with high consequences for failure: lessons for health care operations. *Int J Qual Health Care*. 2004;16(2):125-132.
14. Holmboe ES, Huot S, Chung J, Norcini J, Hawkins RE. Construct validity of the miniclinical evaluation exercise (miniCEX). *Acad Med*. 2003;78(8):826-830.
15. Reddy ST, Farnan JM, Yoon JD, et al. Third-year medical students' participation in and perceptions of unprofessional behaviors. *Acad Med*. 2007;82(10 suppl):S35-S39.
16. Hafferty FW. Professionalism—the next wave. *N Engl J Med*. 2006;355(20):2151-2152.
17. Chang VY, Arora VM, Lev-Ari S, D'Arcy M, Keysar B. Interns overestimate the effectiveness of their hand-off communication. *Pediatrics*. 2010;125(3):491-496.
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(3):203-209.