

Final Progress Report

1. Title Page

Title: Development and Validation of a Tool to Evaluate Handoff Quality

Principal Investigator and Team Members:

Vineet Arora, MD, MAPP (Principal Investigator); Jeanne Farnan, MD, MHPE (Co-investigator);
Leora Horwitz, MD, MHS (Co-investigator, Yale University); Paul Stasiunas, BA (Project
Manager)

Organization: University of Chicago

Inclusive Dates of the Project: Project Period: 09/30/2009 - 09/29/2011

Federal Project Officer:

Project Officer: DENISE BURGESS, (301) 427-1318, denise.burgess@ahrq.hhs.gov

Acknowledgment of Agency Support and Grant Number: This research was supported by
AHRQ R03 HS18278-02 Development and Validation of a Tool to Assess Handoff Quality

2. Structured Abstract (200-words maximum)

Purpose: To develop and validate a tool to assess real-time handoff quality among inpatient physicians

Scope: Although numerous organizations recommend handoff monitoring to ensure safe communication, no reliable or valid tool exists with which to guide handoff observations.

Methods: Trained personnel observed hospitalist and resident handoffs at University of Chicago and Yale University using the Handoff CEX. Senders and receivers performed peer evaluation. Internal consistency (Cronbach's alpha) and inter-rater reliability were examined.

Standardized video scenarios depicting varying levels of handoff quality along three domains (communication, professionalism, and setting) were filmed. Faculty from both sites viewed films in random order, rating performance using the Handoff CEX. Kendall's coefficient of concordance was used to assess if raters could distinguish levels of performance. Two-way ANOVA was used to detect rater bias.

Results: In 673 evaluations representing 149 handoffs involving 98 different subjects (34 UChicago; 64 Yale), external observer ratings were lower than peer ratings. Faculty were able to discern varying levels of performance on setting and professionalism without evidence of rater bias. Discerning superior and satisfactory communication remains challenging. In summary, the three-item Handoff CEX shows promise as a reliable and valid tool to assess varying levels of videos depicting handoff performance.

Key Words: handoffs, patient safety

Word count: 199

3. Purpose: The objectives of this proposal have not been modified and are the following:

Specific Aim 1: To assess the feasibility of using a paper-based tool, the “Handoff CEX (Clinical Evaluation Exercise)” during the performance of actual patient handoffs between house staff and between hospitalist physicians.

Specific Aim 2: To establish the construct validity and inter-rater reliability of the Handoff CEX using standardized video scenarios displaying varying levels of performance during a mock handoff.

4. Scope

The patient handoff, or the temporary transition of care of patients between healthcare providers at the conclusion of their shift, is a time that is recognized as a threat to patient safety due to potential for miscommunication. Critical safety information is often lost or altered during this transfer of clinical details and professional responsibility.¹ Previous research has shown that an increased frequency of handoffs increases the risk of miscommunication and errors.²

Due to these concerns, numerous organizations have recognized the need for vigilance with respect to handoff education and training. In 2006, The Joint Commission made handoffs a “National Patient Safety Goal” and highlighted that “a standardized approach” should be adopted to handoff communications. That same year, the World Health Organization has made prevention of handover errors one of its top five patient safety solutions, which makes it as important as hand hygiene on the list of targets to improve patient safety.

More recent emphasis on handoffs in teaching hospitals has resulted from recent resident duty hour restrictions. For example, in 2011, the American Council for Graduate Medical Education (ACGME) standards state that “sponsoring institutions and programs must ensure and monitor effective, structured handover processes to facilitate both continuity of care and patient safety. Programs must ensure that residents are competent in communicating with team members in the handover process.”³

In addition to resident physicians, the growth of the hospitalist movement has also made handoffs more ubiquitous, as hospitalists often employ rotating shift work to care for patients. Recognizing this, the Society of Hospital Medicine has made handoffs a core competency for hospitalists and convened a task force to make evidence-based recommendations for handoffs. These recommendations, endorsed by the Society of Hospital Medicine, highlight the need to monitor handoffs to ensure that they are safe and effective.⁴

Despite this increasing focus on handoffs, improving handoffs is hindered by a lack of standardized instruments with which to measure handoff communication. To date, no standardized, validated metric exists to measure handoff quality. Because the handoff is a ubiquitous feature of clinical medicine, relevant to all specialties, it is critical that instruments to measure handoff performance are not only created but also validated.⁵

To address this crucial missing piece, we have developed the Handoff CEX, an evaluation instrument which rates the quality of patient handoffs in domains such as professionalism and communication between the receiver and sender of patient information. This tool is informed from other evaluative instruments, such as the Mini-CEX, as well as from hand-off procedures

in other shift-based industries, like aviation.^{6,7} Yet, the feasibility and validity of this tool had not been established. Validating the Handoff CEX may be integral in the training of residents and provide institutions with an instrument to further improve the quality and safety of patient care. Furthermore, the Handoff CEX could assist hospitals as they attempt to monitor and improve handoff quality.

5. Methods

Below, we expand in detail on how we met our objectives: 1) to assess the feasibility of using a paper-based tool, the “Handoff CEX (Clinical Evaluation Exercise)” during the performance of actual patient handoffs between house staff and between hospitalist physicians; and 2) to establish the construct validity and inter-rater reliability of the Handoff CEX using standardized video scenarios displaying varying levels of performance during a mock handoff.

Feasibility and Reliability

To assess the feasibility of using the Handoff CEX as a real-time observation tool in a clinical setting, we trained operators to use the instrument to conduct evaluations in two different settings (academic hospitalist shift change and internal medicine house staff handoffs) at two institutions. Both Yale University and the University of Chicago have an internal medicine residency ward-based services and a nonteaching hospitalist service staffed by physicians who participate daily in handoffs of hospitalized patients. We aimed to evaluate at least 60 handoffs in each of these two settings at the two institutions. Each handoff was evaluated simultaneously by both a participant in the handoff and a third-party observer. Inter-rater reliability between participant and observer was assessed using kappa statistics.

Upon completion, all evaluators were asked to rate their overall satisfaction with the evaluation using the same nine-point scale used in the Mini-CEX. We considered a satisfactory or superior level of overall satisfaction with the Handoff CEX by participants and observers as a marker of feasibility of the use of the Handoff CEX. In addition, we also reported the number of completed handoff evaluations out of those that were attempted. Descriptive statistics were used to summarize the quality of handoffs observed.

The instruments include an overall assessment of handoff competence, also scored on a nine-point scale. Because we expected that ratings of subdomains may be highly correlated, we also performed factor analysis to determine the principal components or most important domains contributing to the overall assessment of handoff quality for both attending and resident physicians. Finally, although not a primary aim of this proposal, comparison was made between the two study groups, academic hospitalists and house staff physicians, to determine the difference in handoff quality and the satisfaction with the Handoff CEX at differing stages of continuing practice.

Creation of Videotapes and Tool Validation

A clinical scenario, based on prior observational work (described above), was developed and represented an internal medicine resident (the sender) signing out three different patients to colleagues (intern [PGY-1] and resident). This scenario was developed to include components of professionalism, communication, and setting. Three levels of performance – superior, satisfactory, and unsatisfactory – were defined and described for the three domains. These

levels were defined and written using the descriptive anchors in the Handoff CEX. After constructing the superior script, or gold standard, showcasing superior communication, professionalism, and setting, we individually changed one domain, holding the other two constant. For example, for the script demonstrating unsatisfactory communication, setting and professionalism were both held at the superior level.

The varying levels of performance were written using previously determined markers of handoff performance that are also reflected in the anchors of the tool. 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 (see Table below). Superior professionalism requires appropriate comments by the sender and a presentation focused on the most urgent patients, whereas unsatisfactory professionalism is shown by a hurried and inattentive signout.

Finally, a superior environment is one in which the sender is listening attentively and discourages interruptions, whereas an unsatisfactory environment finds the sender answering pages during the handoff and surrounded by background noise. We omitted the satisfactory level for “setting” due to the difficulties in creating environmental subtleties. Permutations of each of these domains resulted in six scripts depicting different levels of sender performance. Only the performance level of the sender was changed, and the receivers of the handoff performance remained constant, demonstrated by their attentive listening, read-back, and note-taking during the handoff. Although the scope of this project is focused on the role of the sender, future projects will include permuting the quality of handoff reception by changing the behavior of the receiver.

The scripts were developed by three investigators (JF, SB, VA) and then reviewed and edited independently by two other investigators (VA, PS). The scripts were revised until consensus was reached. Standardized actors performed the video scenarios and were trained by the investigators to depict each level of performance handoff. The videotapes averaged 3 minutes in length.

Workshop Protocol

After a brief introduction, faculty viewed the tapes in random order on a projected screen. Participants were instructed to use the Handoff 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 CEX form, and participants anonymously completed the forms independently without any contact with other participants. The lead investigators proctored all sessions. The participants viewed and rated the six videos over the course of an hour.

Data Collection and Statistical Analysis

Descriptive statistics are reported on raw scores on the Handoff CEX. To assess internal consistency of the tool, Cronbach’s alpha was used. To assess inter-rater reliability of these attending ratings on the tool, we performed a Kendall’s coefficient of concordance. To ascertain if faculty were able to detect varying levels of performance depicted in the video, we did an ordinal test of trend on the scores given for communication, professionalism, and setting. Then, after clustering data by attending rater, a two-way analysis of variance (ANOVA) was performed

to assess if faculty's scores were associated with performance level after controlling for faculty. The results of the faculty coefficients in the two-way ANOVA were also examined for any evidence of rater bias.

6. Results

We have successfully met the initial aims of this project. A brief summary of the results follows. More details will be included in peer-reviewed publications that are expected to be submitted in 2012.

Feasibility and Reliability of the Handoff CEX

Overall, we collected 673 Handoff CEX evaluations: 343 conducted on handoff providers and 330 on handoff recipients. The study was designed to collect multiple evaluations per subject. Consequently, these 673 evaluations were collected during 149 unique signout sessions involving 98 different subjects. A total of 34 unique subjects were evaluated at the University of Chicago, and 64 unique subjects were evaluated at Yale. In total, 468 of the CEX evaluations (69.5%) assessed trainees or mid-level providers (including sub-interns, interns, residents and nurse practitioners); the remainder assessed hospitalist attendings.

Overall scores on the Handoff CEX for signout providers ranged from 1 to 9, demonstrating full use of the scale, with means per domain near 7 for all domains and an SD of roughly 1.5 for all domains. Mean scores were lowest for content and highest for professionalism.

For ratings of receivers, similar pattern was observed with full use of the scale but higher scores overall, with significant increases compared with ratings of senders overall and for setting, organization, and judgment. Mean scores were lowest for setting and highest for professionalism. Note that content is not listed on the evaluation form for receivers, because they are not transmitting content. There were significant differences by training level. Although attendings received significantly higher ratings for judgment than trainees when providing handoffs, attendings received significantly lower scores than trainees for judgment AND organization when receiving handoffs. Last, external evaluators gave significantly lower scores than peer evaluators in every domain for both handoff providers and recipients ($p < 0.0001$ for every domain using paired t test).

Validity of the Handoff CEX

From December 2010 to the present, we conducted 366 handoff observations amongst the University of Chicago Internal Medicine Residency Program. At this time, Drs. Arora, Farnan, and Horwitz began holding ongoing routine phone meetings to discuss plans to begin aim 2 of this study. In March 2010, Dr. Horwitz began piloting the Hand-off CEX amongst internal medicine residents at Yale.

In April 2011, a preliminary analysis revealed that three domains of the Handoff CEX instrument captured 82% of the variance of overall signout quality (communication, professionalism, and setting). Based on this information, the Handoff CEX was reduced to a four-item instrument with these 3 items and overall signout quality.

In May 2011, Drs. Arora and Farnan recruited an experienced videographer in medical education to purchase filming equipment and identify suitable actors and filming locations for aim 2. Script development and filming for the validation study was ongoing from June to mid-August 2011. As stated in specific aim 2, our goal was to establish the construct validity of the Handoff CEX – that is, the ability of the tool to differentiate between levels of performance and truly evaluate the construct [handoff quality] it intends to measure through the use of these standardized video scenarios. In developing these scenarios, we focused on the three main domains of handoff quality reflected in the new, shortened CEX instrument (communication, professionalism, and setting) to generate the video scenarios for the instrument validation.

Six video-based scenarios were developed to highlight varying levels of performance in the domains of communication skills, professionalism, and setting. Videos ranged in length from 3 to 5 minutes and permuted one domain of performance while holding the others constant. Scripts were based upon real-time clinical observations. Faculty were recruited via email to participate in a workshop on handoff education and evaluation both to pilot test the videos and for instrument validation. Videos were shown in a random order, and faculty were instructed to use the Handoff CEX to rate the performance. A debrief occurred immediately afterward to identify barriers and facilitators to the displayed behaviors. Descriptive statistics and two tests of reliability, Cronbach's alpha and Kendall's coefficient of concordance, were performed. Two tests of validity were performed: a test of trend across ordered groups and a two-way ANOVA to examine for rater bias.

Forty-seven faculty from two sites participated. In total, 172 (91%) of a possible 189 handoff observations were captured. Reliability testing revealed a Cronbach's alpha near 0.8 (optimal) and a Kendall's coefficient of concordance greater than 0.6 (indicating high reliability). Faculty were able to reliably distinguish the different levels of performance in professionalism and setting but had greater difficulty distinguishing between satisfactory and superior communication. Two-way ANOVA revealed no evidence of rater bias. Faculty participants commented on face validity of video scenarios, specifically those portraying setting and communication skills. In addition, robust discussion resulted in identifying the barriers and facilitators to the behaviors demonstrated in the video.

In summary, video-based scenarios, utilized to highlight differing levels of performance, with focused debrief are an effective way to observe specific domains and behaviors in handoff communication. Our findings demonstrate that the three-item Handoff CEX shows promise as a reliable and valid tool to assess varying levels of videos depicting handoff performance. We have not encountered any technical problems in completing this project.

Our findings and experience can help guide the adaptation of the Handoff CEX as a tool to evaluate handoff competency. We hope that the Handoff CEX can arm educators with an innovative, necessary, valid and feasible method for training health professionals to conduct safe and effective handoffs. Finally, the Handoff CEX may be a useful tool to assist hospitals in improving patient safety.

Development of a Tool to Assess Handoff Listening Skills

Because receiver ratings using the Handoff CEX exhibited a halo effect, our team was concerned with the ability of the instrument to correctly identify characteristics of listening. Therefore, using literature from the social sciences and informed by the observations obtained in this study, we worked on developing a new tool called the HEAR (Handoff Evaluation Assessing Receivers) Checklist based on active and passive listening behaviors that trained observers were able to use to rate the quality of listening at the handoff. Our results show that passive listening behaviors are more common than active ones (asking questions), and that listening is often hampered by a high rate of potentially modifiable interruptions, such as side conversations, arriving late to handoffs, and paging interruptions. The number of interruptions was also greater when more patients were discussed, highlighting that increased workload is associated with poor-quality handoffs and may need greater attention.

Dissemination Activities

Although this is our final progress report, we wish to highlight our intent to disseminate our findings for a broader audience of investigators, hospital-based clinicians, medical educators, and policymakers. We have presented this data in abstract format at the AHRQ 2010 National Meeting, the Society of Hospital Medicine 2011 National Meeting, and the 2011 Association of American Medical Colleges Annual Meeting in Denver, where it won the award for best research in medical education. We have also submitted an abstract to the Society of Hospital Medicine 2012 Annual Conference in May and are planning to submit to the Society of General Internal Medicine. We are planning to submit our VALID (Video Assessments of Level of Interactive Dialogue at Handoff) Training Videos to AAMC MedEdPortal, a free online portal of curricular resources for health professions educators.

7. List of Publications and Products (Bibliography of Outputs) from the study (Follow the AHRQ Citation Style Format at <http://dev.ahrq.gov/fund/refstyle.htm>)

Publications: To date, we have published two abstracts and submitted a manuscript for review. Several other manuscripts using more detailed data are also in preparation.

Farnan JM, Stasiunas PG, Banerjee SS, Greenstein EA, Horwitz LI, Arora VM. Development and Validation of a Tool to Evaluate Handoff Quality. *J Gen Int Med.* 2011;26(suppl 1):S299.

Arora VM, Stasiunas PG, Banerjee SS, Greenstein EA, Horwitz LI, Farnan, JM. Real-Time Ratings of Handoff Quality by Hospitalist Clinicians. *J of Hosp Med.* 2011;6;4(suppl 2):S8.

Greenstein EA, Arora VM, Stasiunas PG, Banerjee SS, Farnan JM. Characterizing Physician Listening Behavior During Hospitalist Handoffs using the HEAR Checklist. Submitted.

Project-Generated Resources: The Handoff CEX Instrument, three-item Hospitalist Handoff CEX, HEAR (Handoff Evaluation Assessing Receivers) Checklist, and VALID Handoff (Video

Assessments of Level of Interactive Dialogue at Handoff) scenarios are project-generated resources that can be utilized by other programs wishing to train faculty and residents in direct observation and handoff quality. Videos can be used by programs to train senders and receivers in handoff communication. We are in the process of submitting these resources for peer review to AAMC MedEd Portal.

References

1. Hinami K, Farnan JM, Meltzer DO, Arora VM. Understanding communication during hospitalist service changes: A mixed methods study. *J Hosp Med.* 2009;4(9):535-40.
2. 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(6):401-7.
3. ACGME Common Program Requirements. Available at: http://www.acgme.org/acwebsite/home/Common_Program_Requirements_07012011.pdf Accessed on December 28, 2011.
4. Arora VM, Manjarrez E, Dressler DD, Basaviah P, Halasyamani L, Kripalani S. Hospitalist hand-offs: A systematic review and task force recommendations. *J Hosp Med.* 2009;4(7):433-40.
5. Gakhar B, Spencer AL. Using direct observation, formal evaluation, and an interactive curriculum to improve the sign-out practices of internal medicine interns. *Acad Med.* 2010;85(7):1182-8.
6. Holmboe ES, Huot S, Chung J, Norcini J, Hawkins RE. Construct validity of the miniclinical evaluation exercise (miniCEX). *Acad Med.* 2003;78(8):826-30.
- 7 Hawkins RE, Margolis MJ, Durning SJ, Norcini JJ. Constructing a validity argument for the mini-clinical evaluation exercise: A review of the research. *Acad Med.* 2010;85(9):1453-61.