


TECHNICAL ADVANCE

Open Access



Improving handover competency in preclinical medical and health professions students: establishing the reliability and construct validity of an assessment instrument

Meghan Michael^{1*} , Andrew C. Griggs², Ian H. Shields³, Mozhdeh Sadighi⁴, Jessica Hernandez⁵, Chrissy Chan⁵, Mary McHugh⁵, Blake E. Nichols⁶, Kavita Joshi⁵, Daniel Testa⁵, Sonika Raj⁵, Richard Preble³, Elizabeth H. Lazzara² and Philip E. Greilich¹

Abstract

Background: As part of the worldwide call to enhance the safety of patient handovers of care, the Association of American Medical Colleges (AAMC) requires that all graduating students “give or receive a patient handover to transition care responsibly” as one of its Core Entrustable Professional Activities (EPAs) for Entering Residency. Students therefore require educational activities that build the necessary teamwork skills to perform structured handovers. To date, a reliable instrument designed to assess teamwork competencies, like structured communication, throughout their preclinical and clinical years does not exist.

Method: Our team developed an assessment instrument that evaluates both the use of structured communication and two additional teamwork competencies necessary to perform safe patient handovers. This instrument was utilized to assess 192 handovers that were recorded from a sample of 229 preclinical medical students and 25 health professions students who participated in a virtual course on safe patient handovers. Five raters were trained on utilization of the assessment instrument, and consensus was established. Each handover was reviewed independently by two separate raters.

Results: The raters achieved 72.22% agreement across items in the reviewed handovers. Krippendorff's alpha coefficient to assess inter-rater reliability was 0.6245, indicating substantial agreement among the raters. A confirmatory factor analysis (CFA) demonstrated the orthogonal characteristics of items in this instrument with rotated item loadings onto three distinct factors providing preliminary evidence of construct validity.

* Correspondence: Meghan.Michael@utsouthwestern.edu

¹Department of Anesthesiology and Pain Management, University of Texas Southwestern Medical Center, 5323 Harry Hines Blvd, Mail Code 9068, Dallas, TX 75390, USA

Full list of author information is available at the end of the article



© The Author(s). 2021 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Conclusions: We present an assessment instrument with substantial reliability and preliminary evidence of construct validity designed to evaluate both use of structured handover format as well as two team competencies necessary for safe patient handovers. Our assessment instrument can be used by educators to evaluate learners' handoff performance as early as their preclinical years and is broadly applicable in the clinical context in which it is utilized. In the journey to optimize safe patient care through improved teamwork during handovers, our instrument achieves a critical step in the process of developing a validated assessment instrument to evaluate learners as they seek to accomplish this goal.

Keywords: Teamwork, Handovers, Measurement, Reliability, Validity

Background

Handover of patient care, defined as the transition of information and responsibility from one individual or team to another [1, 2], is a critical time in a patient's medical journey. Handovers are high-frequency, error-prone events that have been implicated in nearly 80% of serious adverse events [3]. Given the criticality of handovers, organizational bodies have put forth mandates to strengthen handovers. The World Health Organization [4, 5] and Joint Commission [6] have mandated a more structured process to mitigate the risk of communication failures on patient safety. Although there is meta-analytic evidence that suggests structure within handovers improves patient outcomes [7], the specific structure is left to individual institutions to adopt.

Governing bodies for undergraduate [8–10] and graduate [11] education identified handovers as a core professional skill that future healthcare workforce must be prepared to perform. Accredited training programs must prepare and validate a healthcare trainee's behavioral competencies in conducting safe patient handovers. These teamwork behaviors are essential non-technical skills required to achieve highly reliable interprofessional communication in a variety of ambulatory and inpatient settings [12, 13].

Despite these educational mandates, there is little guidance on when training in teamwork competencies should begin. Currently, early exposure of undergraduate learners to the skills needed for interprofessional patient handovers is largely ignored and reflects a significant gap in pre-clinical education [2, 14–16]. Emerging evidence suggests that introduction of interprofessional educational activities in preclinical learners can improve attitudes and skills needed for team-based communication during handovers [17–19]. In addition, this strategy helps to mitigate the risk of introducing this core entrustable professional activity during clinical rotations, which can be unpredictable and prone to inconsistency, being untestable, and frequently unsustainable [19–22]. Given this background, an essential component of a longitudinal curriculum designed to improve competencies required for handovers is a psychometrically sound

assessment instrument that is appropriate for preclinical learners with minimal or no clinical experience [22–24].

The development of a reliable and valid instrument capable of assessing constructs critical to safe patient handovers is essential. Such an instrument would not only permit the assessment of learners' proficiency in the required behaviors, but also, secondarily, the curriculum itself. Ideally, an assessment instrument needs to be generalizable and subjected to rigorous methodological design [25–27]. Despite the high number of available teamwork and communication-focused interventions [2, 28–31], a literature review found that assessment instruments for discrete handover competencies were rare and had widespread heterogeneity depending on specialty, profession, handover context, and institution [32]. In fact, consensus among systematic reviews suggests ongoing gaps in available evidence regarding high-quality assessment instruments of competency-based training in handovers, communication behaviors, and teamwork behaviors [33–38].

Previous systematic reviews by Gordon, et al. [2, 34], found reliability and/or validity of assessment instruments for handover competencies were reported in only a few studies even though conclusions on their efficacy and usability were common. Muller, et al. [16] reported similar findings in handover instruments used to assess educational curricula designed to teach structured communication using the "Situation-Background-Assessment-Recommendation" (SBAR) handover tool [39]. In studies reporting reliability or validity data, Davis, et al. [38] found evaluator training, experimental conditions, and study bias were poorly controlled. The available evidence underscores the need to develop reliable instruments for assessing curricula designed to teach teamwork and communication in pre-clinical learners.

The purpose of this study is to establish the reliability and construct validity of an assessment instrument capable of evaluating communication competencies for handovers (i.e., structured communication, closed-loop communication, and question clarification) in preclinical medical and health professions students. We use the clinical handover as a pedagogical vehicle to accomplish

assessment of learners and leverage the SBAR framework to enable pre-clinical learners to engage in a clinical simulation targeting communication skills during handovers.

Methods

Study design and setting

We developed a curriculum to teach preclinical students with limited to no clinical experience multiple competencies as they relate to handovers (i.e., structured communication, closed-loop communication, and asking clarifying questions). The curriculum assumed that learners had three semesters of didactic basic science education but no prior clinical science training. The curriculum also involved prereading, videos, and activities that were facilitated by clinical educators and hosted in Microsoft Teams (Microsoft Corporation, version 4.2.4.0) [40] in which groups of preclinical students acted as handover senders and receivers for four simulated patient scenarios. Students reviewed the four patient scenarios and performed simulated handovers for each patient during the activity. Audiovisual recordings of these handovers were collected and reviewed by multiple raters. The raters utilized a seven-item behavioral assessment instrument developed by our team to rate students' performance based on the degree to which they handed over pertinent clinical information and whether they leveraged clarifying questions and closed-loop communication during the handover. Data collected by multiple raters reviewing these recordings were collated and analyzed for inter-rater reliability and construct validity.

Participants

During the handover activity, multiple groups of four to six interprofessional students performed handovers for each of the four simulated patient scenarios facilitated by a clinical educator. Each handover occurred between one sender and one receiver assigned by the educator, while the rest of the team played an observer role. In total, 192 handovers were recorded from a sample of 229 medical students and 25 health profession students who participated in the activity ($N = 254$ students). After the handover activity, a team of five raters reviewed recordings of the students' handovers to evaluate their performance ($N = 180$ reviewed handovers).

Materials

Patient scenarios

The simulated patient scenarios (Scenario A, B, C, and D) contained all the information to be handed over during the activity. The patient scenarios were pre-recorded in a simulation center. Each scenario featured a patient describing their symptoms, a physician assessing the patient and asking them questions, and a nurse checking

the patient's vitals. Students were assigned to review either scenarios A and B, or C and D before joining the handover activity and were instructed to utilize SBAR as a structured communication tool to organize the information from the patient scenarios during their handovers. Table 1 provides a summary of the patient scenarios including the patients' gender, age, and chief complaint.

Behavioral measure

To facilitate the raters' measurement of students' handover performance during the activity, we developed a behavioral assessment instrument with items related to three handover communication competencies: structured communication via Situation, Background, Assessment, and Recommendation (SBAR; 4 items), closed-loop communication (2 items), and asking clarifying questions (1 item). The inclusion of these seven items was primarily based on three considerations.

The first consideration centered on the first four items of the tool (i.e., the items related to SBAR). SBAR was utilized due to its substantial empirical support as well as it being widely used and recommended [16, 41]. To elaborate, SBAR has a long history across a variety of countries, domains, and provider types. In fact, a systematic review of handoff protocols found that a majority of handoff mnemonics were leveraging the SBAR structure [41]. Additionally, there is evidence to suggest that SBAR is associated with enhanced perceptions as well as patient outcomes [42]. Due to such evidence, SBAR is recognized as an effective communication tool by The Joint Commission, the Agency for Healthcare Research and Quality, Institute for Healthcare Improvement, and the World Health Organization [42] as well as the Royal College of Physicians [43]. Such recommendations are rooted in the notion that structured communication enhances shared understanding and the potential to mitigate memory lapses.

The second consideration focused on closed-loop communication, which relates to the next two items on the tool. Closed-loop communication has been widely touted as a useful component within effective teamwork (e.g., Salas, et al. [44, 45]) and even patient care [44, 46]. As further support, a recent narrative review stated that closed-loop communication is one of the most useful

Table 1 Patient scenarios

Scenario Label	Patient Gender	Patient Age	Chief Complaint
A	Male	50	Shortness of breath
B	Male	75	Foul smelling urine
C	Female	55	Hypertension
D	Female	35	Acute onset headache

strategies for strengthening communication for specific medical teams [47]. Similar to structured communication, closed-loop communication fosters accuracy and ameliorates potential errors.

The last consideration targeted the remaining item – asking clarifying questions. That item was included due to its criticality in patient care and even handovers specifically [48–50]. As such, the Joint Commission and the Agency for Healthcare Research and Quality have explicitly stated that asking questions should be formally incorporated during handovers [51, 52]. These mandates are grounded in the idea that asking questions strengthens resiliency [53]. Table 2 provides a list of all the items on this behavioral assessment instrument. The three aforementioned considerations served as the theoretical basis for the development of the seven items in our assessment instrument.

Each item could receive one of three scores based on pre-defined behavioral anchors: No/Never, Sometimes/Somewhat, or Yes/Always. We developed a scoring guide to provide to the raters before reviewing any videos and later revised the guide to bolster inter-rater reliability. The scoring guide provided guidance concerning when each scoring option should be used for all items as well as the most pertinent information to be handed over for the structured communication items in each patient scenario. Students did not have access to the assessment instrument or the scoring guide. However, detailed instructions and informational material on both structured handovers and team competencies, supplemented with examples, were shared with the students in the format of pre-reading documents and videos. Table 3 provides a summary of this scoring guide.

Procedure

One hour before the handover activity, students were provided with a Microsoft Teams meeting invitation and links to access the patient scenarios to be handed over during the activity via email. Students were assigned two patient scenarios to review (A and B, or C and D) before using the provided link to join the handover activity in Microsoft Teams. After joining the activity, students

were assigned to act as either a handover sender for one of the patients in the scenarios they were instructed to review or a receiver for one of the patients that they did not review. Each handover had one sender and one receiver. Recordings of these handovers were stored in a secure, network-enabled repository for later analysis.

We developed a scoring guide to aid raters in assessing students' performance during the activity when reviewing the recordings. We then assembled a multidisciplinary team of five raters from varying clinical backgrounds, including faculty from the Emergency Medicine and Pediatrics departments as well a medical student (Dean's scholar), that were familiar with SBAR handovers and the clinical competencies being assessed. We held a meeting with the raters to describe the handovers that were performed as well as the scoring guide and behavioral assessment items that were developed. Following this meeting, the raters were tasked with reviewing twelve handovers across three video recordings to serve as training before reviewing the full sample of videos. Inter-rater reliability statistics were calculated based on the data collected from these three videos, and the scoring guide was revised for clarity based on rater feedback. Next, the raters were tasked with reviewing the remaining videos in pairs such that each video was reviewed by at least two different raters. Finally, once all videos were reviewed by the team of raters, their data were collated, and their inter-rater reliability was assessed.

Statistical analysis

Following training, data from the team of five raters were collated and imported into IBM's SPSS software (Version 27) [54] for analysis. We identified instances where all raters were in total agreement or at least one rater was in disagreement for each item across handovers under analysis and calculated percent agreement by comparing the proportion of each response. As all raters reviewed all subjects at this stage (i.e., handovers contained in videos) and their responses were recorded as categorical variables (i.e., No/Never, Sometimes/Somewhat, or Yes/Always), we used Fleiss' kappa to

Table 2 Behavioral assessment instrument items

Associated Competencies	Items
Structured Communications	1. Sender described <i>situation</i> succinctly 2. Sender provided additional <i>background</i> info 3. Sender offered <i>assessment</i> based on situation and background 4. Sender stated clear <i>recommendation</i> about what should happen
Closed-Loop Communications	5. Receiver repeated what they heard (not just "I understand") 6. Sender clearly confirmed or corrected the check back
Asking Clarifying Questions	7. Receiver asked clarifying questions from sender when receiving handover

Table 3 Summary of scoring guide

Competency	Item Number(s)	No/Never Scoring Requirements	Sometimes/Somewhat Scoring Requirements	Yes/Always Scoring Requirements
Structured Communications	1–4	No pertinent SBAR information handed over	A portion of pertinent SBAR information handed over	All pertinent SBAR information handed over
Closed-Loop Communications	5	Receivers only stated “I understand” or failed to repeat any information from <i>any</i> components of SBAR	Receivers repeated at least one piece of information from <i>at least one</i> component of SBAR	Receivers repeated at least one piece of information from <i>each</i> component of SBAR
	6	Senders fail to confirm or correct <i>any</i> the information given in the checkback	Senders confirmed or correct <i>a portion</i> of the information given in the checkback	Senders confirmed or correct <i>all</i> information given in the checkback
Asking Clarifying Questions	7	Receivers asked <i>no</i> clarifying questions	Receivers asked at least <i>one</i> clarifying question	Receivers asked at least <i>two or more</i> clarifying questions

assess inter-rater reliability. Fleiss’ kappa is a measure of inter-rater reliability that can be used with categorical data collected from two or more raters [55]. Table 4 provides common cut-off values denoted by Landis and Koch [56] that are applicable to Fleiss’ kappa.

Once all videos were reviewed by the rating team, we collated and imported the data for the full sample of videos into SPSS to calculate percent agreement and overall measures of inter-rater reliability. At this stage, since not all raters reviewed all subjects (i.e., five raters for each training video and two raters per remaining video), we utilized the Krippendorff’s alpha SPSS macro developed by Andrew Hayes [57] to calculate an inter-rater reliability coefficient from the entire data set. Krippendorff’s alpha can be used with two or more raters and response categories, ordinal data, or missing data, and the same cut-off values provided in Table 4 can be applied to Krippendorff’s alpha [55]. We also utilized Cronbach’s alpha to serve as a secondary measure of reliability.

Prior researchers have noted that what constitutes acceptable inter-rater reliability based on these qualitative cutoffs is often debated [58, 59]. Krippendorff [60] has provided more conservative interpretations based on their work in content analysis (i.e., “rely only on variables with reliabilities above $\alpha = .8$,” p. 241). Values below this cut-off are often retained in research, however, and what constitutes acceptable inter-rater reliability can vary depending on the hypotheses or research questions under observation [58]. Concerning healthcare

research, McHugh [59] has argued that kappa values below 0.60 are unacceptable.

Following assessment of inter-rater reliability, we performed a confirmatory factor analysis (CFA) in IBM’s SPSS software (Version 27) [54] to establish a preliminary measure of this instrument’s construct validity. A single rating was determined for each item across handovers in this analysis using the mode of raters’ responses when possible. In instances where a mode response could not be calculated (i.e., for items in which two raters provided non-matching responses), a single rater’s data were used. As the items in this instrument were based on three distinct clinical competencies needed for effective handovers, a fixed number of three factors were selected a priori to be extracted in our analysis.

Results

Reliability

Results from rater training

We collated data from twelve handovers reviewed by the team of raters during training and calculated percent agreement and Fleiss’ kappa as preliminary measures of inter-rater reliability at this stage. From the handovers reviewed during training (i.e., three videos containing four handovers each, with each handover rated using seven items), the raters achieved 31.25% agreement across the 84 items. We calculated Fleiss’ kappa and achieved a coefficient of 0.339, indicating fair agreement among the raters. Based on these data, we revised the scoring guide for clarity and tasked the five raters with reviewing the remaining videos in pairs.

Results from full assessment

Following training and the review of all remaining videos, we recalculated percent agreement and ran Krippendorff’s alpha to assess inter-rater reliability. Based on the data from the full sample of handovers (i.e., 45 videos containing 4 handovers each that were rated using 7 items), the raters achieved 72.22% agreement across the 1,260 items. We calculated Krippendorff’s alpha and

Table 4 Kappa Value Interpretation Cut-Offs Provided by Landis and Koch [56]

Kappa Value	Interpretation
< 0	Poor Agreement
0.01–0.20	Slight Agreement
0.21–0.40	Fair Agreement
0.41–0.60	Moderate Agreement
0.61–0.80	Substantial Agreement
0.81–1.00	Almost Perfect Agreement

achieved a coefficient of 0.6245, indicating substantial agreement among the raters. We also calculated Cronbach's alpha and achieved a coefficient of 0.63.

Construct validity

As a preliminary measure of construct validity, we ran a CFA on the 7-items in our instrument that measured three competencies related to handovers: structured communications, closed-loop communications, and asking clarifying questions. We assessed the suitability of CFA prior to analysis; inter-item correlations between -0.093 and 0.797 were observed, the overall Kaiser-Meyer-Olkin (KMO) measure was 0.552 with individual KMO measures between 0.359 and 0.977 , and Bartlett's test of sphericity was statistically significant ($p < .001$).

CFA revealed that three factors had eigenvalues of approximately one or greater which explained 27.7%, 21.6%, and 14.2% of the total variance, respectively. Visual inspection of the scree plot also indicated that three factors should be retained. The three-factor structure explained 63.5% of the total variance. We used a Varimax orthogonal rotation to support interpretability. Items regarding closed-loop communication had strong loadings on Factor 1, items concerning SBAR exhibited moderate to weak loadings on Factor 2, and the item regarding clarifying questions exhibited a moderate loading onto Factor 3. A summary of rotated factor loadings is presented in Table 5. These loadings demonstrate preliminary evidence of the orthogonal nature of constructs measured using the seven items in this instrument.

Discussion

Our team has described an assessment instrument that demonstrates substantial agreement between raters as well as some preliminary evidence towards construct validity in discerning proficiency in the use of structured handover format as well as two communication competencies required for safe patient handovers. In addition to filling the void of reliable and construct valid assessments that address both structured handovers and teamwork competencies, this instrument addresses an

important gap in preclinical medical education for teamwork competencies given the lack of psychometrically sound instruments for assessments that can be applied to the preclinical and interprofessional population.

In addition to filling the aforementioned gaps, our instrument offers many features that make it useful for evaluators and instructors. First, it is appropriate for learners of all levels, including those in their preclinical years of training. Because all learners throughout the medical education curriculum will need to demonstrate proficiency in handovers and teamwork, it is imperative that all learners are assessed with an appropriate, yet scientifically sound, instrument. Second, our instrument is generalizable to any clinical context. Because teamwork is necessary in any facet of care and handovers are conducted by every clinical role in any department, it is critical that learners are competent independent of clinical context. Third, our instrument is generalizable for any handover. Handovers are extremely heterogeneous; that is, the same information is not necessarily exchanged across contexts. For example, the information exchanged in a handover between two anesthesia providers during an anesthetic will contain very different pertinent information than a handover between two care providers in a primary care clinic. Because our instrument is not contingent upon a specific type of handover, it applies to various types of handovers by simply exchanging the SBAR tool for one that is applicable to the given context. Fourth, our instrument has the potential to be employed during real-time ratings as well as during video review. Being flexible to different modalities is advantageous because the availability of resources fluctuates. In other words, it may not be possible to have raters available to attend and observe every educational session; conversely, it may not be possible to have raters review numerous hours of video footage. Similarly, raters can select their modality based on convenience; therefore, it is beneficial to have an instrument that can be adapted for different modalities. Fifth, our instrument has the potential to be utilized by raters with varying skill levels. To expand, raters with extensive medical expertise (e.g., attending

Table 5 Rotated factor matrix

Items	Factor		
	1	2	3
Sender clearly confirmed or corrected the check back	0.932		
Receiver repeated what they heard (not just "I understand")	0.928		
Sender offered assessment based on situation and background		0.725	
Sender described situation succinctly		0.671	
Sender stated clear recommendation about what should happen		0.624	-0.418
Receiver asked clarifying questions from sender when receiving handover			0.774
Sender provided additional background info		0.377	0.484

physicians) may be cost-intensive and often have limited time to dedicate to such tasks. Meanwhile, raters with less medical expertise (e.g., residents or interns) are more cost-effective and have different time demands. Finally, because this instrument only has three categories for ratings, it has the potential to require less training compared to more sophisticated ratings. Even though less sophistication would seemingly appear to be disadvantageous, such instruments require extensive training to achieve appropriate inter-rater reliability. Less sophisticated instruments, although not able to be as diagnostic as their more sophisticated counterparts, also necessitate less resources for training. Considering that many raters often have clinical or administrative responsibilities, having an instrument that potentially needs less training can be especially worthwhile. Further simplification, if desired, could bring the assessment to a dichotomous yes/no rating that could be utilized as a pass/fail assessment.

We posit that our instrument has merit; however, our study does have some limitations that are worth noting. First, while we analyzed a substantial number of handovers, this was an early phase, single institutional experience which limits its generalizability. Despite it being only a single experience, we would argue that such an event is necessary for establishing some of the psychometric properties of measurement instruments before a large-scale roll out becomes integrated into the curriculum. In addition, while we attempted to include a larger number of interprofessional students, only 25 health professional students versus 229 medical students participated. Therefore, this may impact the appropriateness of the use of this assessment instrument outside of the medical student population. Furthermore, when considering which form of structured communication framework to utilize, we selected SBAR due to its prevalence as well as its simplicity relative to other communication frameworks such as Illness Severity- Patient Summary-Action List-Situation Awareness and Contingency Planning- Synthesis by Receiver (IPASS) [61]. Given the inexperience of preclinical students and our goal to avoid cognitive overload of learners and raters, we felt the SBAR format met multiple needs. Our instrument is currently limited to the application of SBAR only; however, the overarching competency is 'structured communication', so the instrument could be modified such that structured communication represents a different handover framework (e.g., IPASS). Finally, there is room for improvement in this instrument regarding its diagnosticity and specificity. For example, there could be a greater number of assessment items for each competency. As another example, there could be greater granularity about the quality of the behaviors exhibited. Adding more items or categories would likely strengthen the

diagnosticity of the instrument, but it can diminish the simplicity and usability of the instrument. A simplistic and usable instrument is advantageous for raters compared to more complicated instruments as it reduces the cognitive load required by raters. Complicated instruments can be cumbersome to use, are more cognitively taxing, and require more extensive training.

Given some of these limitations, we present several avenues for future research. Our study primarily focused on reliability, so one avenue future work should investigate additional types of validity. Another direction is to expand beyond a single educational experience and investigate handovers and competence longitudinally. A third idea is to examine additional student samples outside of medical students. All health professionals conduct handovers, so all of them need to perform handovers effectively. Consequently, all health professional students should be assessed accordingly. A fourth area for future expansion would be to modify the structured communication component to other handover frameworks (e.g., IPASS). Even though there have been mandates to strengthen handover standardization, handovers remain heterogeneous. Therefore, there would be utility in investigating structured communication within other handover frameworks. Finally, future work could examine additional items or even categories to strengthen the diagnosticity and granularity of the instrument and its reliability in the hands of evaluators with even more varied clinical expertise. Because handovers rely on safe behaviors, diagnostic insights into how to perform and improve handovers would be informative. That is, it is insufficient to simply do handovers, they must be performed safely and effectively.

Conclusions

Handovers remain to be frequent, yet problematic events that necessitate multiple competencies to perform effectively. To determine if individuals are competent in conducting handovers, they must be assessed appropriately. Assessment instruments, therefore, need to demonstrate reliability and validity. The assessment instrument described in this study demonstrated substantial agreement, acceptable reliability, and preliminary construct validity in assessing competencies required for safe patient handover in pre-clinical learners. Further development of this instrument could be helpful in assessing entrustable professional activities (EPA #8) and Interprofessional Education Collaborative (IPEC) sub-competencies (Teamwork & Communication) in undergraduate medical and health professions students.

Abbreviations

AAMC: Association of American Medical Colleges; EPAs: Entrustable Professional Activities; SBAR: Situation-Background-Assessment-Recommendation; IPASS: Illness Severity- Patient Summary-Action List-

Situation awareness and Contingency Planning- Synthesis by receiver;
IPEC: Interprofessional Education Collaborative

Acknowledgements

The authors would like to acknowledge the Offices of Medical Education and Quality, Safety, and Outcomes Education for their support of the Quality Enhancement Plan-Team FIRST and Scott Tannenbaum, PhD for his guidance on the overall design of this longitudinal educational program. The authors would also like to acknowledge the contribution of Aditee Ambardekar, MD, MEd in the conception of the idea for this work. In addition, the authors would like to acknowledge the support provided by the staff of the UT Southwestern Simulation Center. Finally, the authors would like to acknowledge David Klocko, MPAS, DMSC and Bau Tran, PharmD from the School of Health Professions for their assistance in recruiting health professions students for the learning activity, which allowed us to assess their handovers and teamwork competencies.

Authors' contributions

MMM assisted in the development and execution of the safe patient handover course, creating research design, conception of the idea for this manuscript, writing portions of the manuscript, and managing the drafts. ACG supported the development of the assessment instrument and scoring guide, facilitated rater training, led statistical analysis efforts, and authored sections. IHS assisted in the development and execution of the safe patient handover course, creating research design, conduct and validation of the assessment evaluations, providing feedback on the assessment design, conception of the idea for this manuscript, writing portions, and managing the drafts. MS contributed to the development of the assessment instrument and guide, supported the statistical analysis efforts, and authored sections of the manuscript. JH assisted in the development and execution of the safe patient handover course, creating research design, conception of the idea for this manuscript, writing portions of the manuscript and revisions. RP contributed to the development and execution of the safe patient handover course and all integral elements as well as development of this manuscript. EHL led the development of the assessment instrument and scoring guide, supported statistical analysis efforts, and authored sections of the manuscript. CC assisted in the design, development, and execution of the safe patient handover course. She also reviewed and revised this manuscript. MM assisted in the design, development, and execution of the safe patient handover course. She also reviewed and revised this manuscript. KJ assisted in the execution of the safe patient handover course, acted as a rater for participant performance, and provided feedback on scoring instrument to improve inter-rater reliability. She also reviewed and edited the manuscript. BEN assisted in the analysis and interpretation of data for the work by acting as a rater and working to provide feedback to alter the instrument to improve inter-rater reliability. He also revised the manuscript critically and approved the final draft. DT assisted in the execution of the safe patient handover course, acted as a rater for participant performance, and provided feedback on scoring instrument to improve inter-rater reliability. He also reviewed the manuscript. SR assisted in the execution of the safe patient handover course, acted as a rater for participant performance, and provided feedback on scoring instrument to improve inter-rater reliability. She also reviewed and edited the manuscript. PG assisted in the development and execution of the safe patient handover course, creating research design, conception of the idea for this manuscript, writing portions of the manuscript, and managing and editing the drafts. All authors have read and approved the final manuscript.

Funding

This work was partially funded by the University of Texas Southwestern Office of Medical Education to support activities within its Simulation Center and effort by several of its contributors.

Availability of data and materials

The datasets (gradings from rater training and full assessment) used and analyzed during this study are available from the corresponding author on request. The student-led handover video recordings are not publicly available due to FERPA and are only available on reasonable request with permission of UT Southwestern.

Declarations

Ethics approval and consent to participate

The UT Southwestern Human Research Protection Program (HRPP) has reviewed this project and determined that it does not meet the definition of research under 45 CFR 46.102 and therefore does not require IRB approval or oversight.

We obtained written and verbal consent to record student interactions of standard educational activities for the purposes of quality improvement and medical education research. The study and use of participant data comply with the Family Educational Rights and Privacy Act (FERPA) (20 U.S.C. § 1232 g; 34 CFR Part 99).

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

¹Department of Anesthesiology and Pain Management, University of Texas Southwestern Medical Center, 5323 Harry Hines Blvd, Mail Code 9068, Dallas, TX 75390, USA. ²Department of Human Factors and Behavioral Neurobiology, Embry-Riddle Aeronautical University, 1 Aerospace Blvd, Daytona Beach, FL 32114, USA. ³Office of Quality, Safety, and Outcomes Education, University of Texas Southwestern Medical Center, 5323 Harry Hines Blvd, Dallas, TX 75390, USA. ⁴Department of Undergraduate Medical Education, University of Texas Southwestern Medical Center, 5323 Harry Hines Blvd, Dallas, TX 75390, USA. ⁵Department of Emergency Medicine, University of Texas Southwestern Medical Center, 5323 Harry Hines Blvd, Dallas, TX 75390, USA. ⁶Department of Pediatrics, University of Texas Southwestern Medical Center, 5323 Harry Hines Blvd, Mail Code 9063, Dallas, TX 75390, USA.

Received: 25 June 2021 Accepted: 15 September 2021

Published online: 02 October 2021

References

- Solet DJ, Norvell JM, Rutan GH, Frankel RM. Lost in translation: challenges and opportunities in physician-to-physician communication during patient handoffs. *Acad Med.* 2005;80(12):1094–9.
- Gordon M, Hill E, Stojan JN, Daniel M. Educational Interventions to Improve Handover in Health Care: An Updated Systematic Review. *Acad Med.* 2018; 93(8):1234–44.
- Joint Commission Center for Transforming. Healthcare releases targeted solutions tool for hand-off communications. *Jt Comm Perspect.* 2012;32(8):1, 3.
- World Health Organization. Communication during patient hand-overs. 2007. Available from: <https://www.who.int/teams/integrated-health-services/patient-safety/research/patient-safety-solutions> Accessed.
- World Health Organization. Framework for action on interprofessional education & collaborative practice. Geneva: WHO; 2010. WHO reference number WHO/HRH/HPN/10.3. Available from: https://www.who.int/hrh/resources/framework_action/en/ Accessed 24 Apr 2021.
- Joint Commission. Provision of Care standard PC.02.02.01. Element of performance 22010.
- Keebler JR, Lazzara EH, Patzer BS, Palmer EM, Plummer JP, Smith DC, et al. Meta-Analyses of the Effects of Standardized Handoff Protocols on Patient, Provider, and Organizational Outcomes. *Hum Factors.* 2016;58(8):1187–205.
- Association of American Medical Colleges. Core Entrustable Professional Activities for Entering Residency. 2014. <https://www.aamc.org/what-we-do/mission-areas/medical-education/cbme/core-epas>. Accessed 13 Jan 2021.
- Interprofessional Education Collaborative. Core competencies for interprofessional collaborative practice: 2016 update. Washington, D.C.: Interprofessional Education Collaborative; 2016. Available from: <https://www.ipecollaborative.org/ipec-core-competencies> Accessed 27 Apr 2021.
- Interprofessional Education Collaborative Expert Panel. Core competencies for interprofessional collaborative practice: Report of an expert panel. Washington, D.C.: Interprofessional Education Collaborative; 2011. Available from: <https://www.ipecollaborative.org/ipec-core-competencies> Accessed 27 Apr 2021.

11. Accreditation Council for Graduate Medical Education. ACGME Common program requirements (Residency). 2011. Available from: <https://www.acgme.org/What-We-Do/Accreditation/Common-Program-Requirements/> Accessed 5 Jan 2019.
12. Philibert I. Use of strategies from high-reliability organisations to the patient hand-off by resident physicians: practical implications. *Qual Saf Health Care*. 2009;18(4):261–6.
13. Rosen MA, DiazGranados D, Dietz AS, Benishek LE, Thompson D, Pronovost PJ, et al. Teamwork in healthcare: Key discoveries enabling safer, high-quality care. *Am Psychol*. 2018;73(4):433–50.
14. Gore A, Leasure AR, Carithers C, Miller B. Integrating hand-off communication into undergraduate nursing clinical courses. *J Nurs Educ Pract*. 2015;5(4):70–6.
15. Eddy K, Jordan Z, Stephenson M. Health professionals' experience of teamwork education in acute hospital settings: a systematic review of qualitative literature. *JBI Database System Rev Implement Rep*. 2016;14(4):96–137.
16. Muller M, Jurgens J, Redaelli M, Klingberg K, Hautz WE, Stock S. Impact of the communication and patient hand-off tool SBAR on patient safety: a systematic review. *BMJ Open*. 2018;8(8):e022202.
17. Berger-Estilita J, Chiang H, Stricker D, Fuchs A, Greif R, McAleer S. Attitudes of medical students towards interprofessional education: a mixed-methods study. *PloS one*. 2020;15(10):e0240835.
18. Banerjee A, Slagle JM, Mercado ND, Booker R, Miller A, France DJ, et al. A simulation-based curriculum to introduce key teamwork principles to entering medical students. *BMC Med Educ*. 2016;16(1):295.
19. Surmon L, Bialocerkowski A, Hu W. Perceptions of preparedness for the first medical clerkship: a systematic review and synthesis. *BMC Med Educ*. 2016;16(1):89.
20. Hughes HK, Serwint JR, O'Toole JK, Spector ND, Ngo TL. I-PASS Adherence and Implications for Future Handoff Training. *J Grad Med Educ*. 2019;11(3):301–6.
21. Shahian D. I-PASS handover system: a decade of evidence demands action. *BMJ Qual Saf*. 2021;30(10):769–74.
22. Pinilla S, Lenouvel E, Cantisani A, Kloppel S, Strik W, Huwendiek S, et al. Working with entrustable professional activities in clinical education in undergraduate medical education: a scoping review. *BMC Med Educ*. 2021;21(1):172.
23. Chen HC, McNamara M, Teherani A, Cate OT, O'Sullivan P. Developing Entrustable Professional Activities for Entry Into Clerkship. *Acad Med*. 2016;91(2):247–55.
24. Sheng J, Manjunath S, Michael M, Gajera P, Wang E, Sendelbach D, et al. Integrating handover curricula in medical school. *Clin Teach*. 2020;17(6):661–8.
25. Praslova L. Adaptation of Kirkpatrick's four level model of training criteria to assessment of learning outcomes and program evaluation in Higher Education. *Educ Assess Eval Account*. 2010;22(3):215–25.
26. Cook DA, Hatala R. Validation of educational assessments: a primer for simulation and beyond. *Adv Simul (Lond)*. 2016;1(1):31.
27. Webster KLW, Tan AC, Unger N, Lazzara EH. Considerations and Strategies for Assessing: Simulation-Based Training in Interprofessional Education. In: Paige JT, Sonesh SC, Garbee DD, Bonanno LS, editors. *Comprehensive Healthcare Simulation: interProfessional Team Training and Simulation*. Comprehensive Healthcare Simulation. Cham: Springer; 2020. p. 121–33.
28. Welsch LA, Hoch J, Poston RD, Parodi VA, Akpınar-Elci M. Interprofessional education involving didactic TeamSTEPPS(R) and interactive healthcare simulation: A systematic review. *J Interprof Care*. 2018;32(6):657–65.
29. Lee J, Kim H, Kim KH, Jung D, Jowsey T, Webster CS. Effective virtual patient simulators for medical communication training: A systematic review. *Med Educ*. 2020;54(9):786–95.
30. O'Donovan R, McAuliffe E. A systematic review exploring the content and outcomes of interventions to improve psychological safety, speaking up and voice behaviour. *BMC Health Serv Res*. 2020;20(1):101.
31. Gelis A, Cervello S, Rey R, Llorca G, Lambert P, Franck N, et al. Peer Role-Play for Training Communication Skills in Medical Students: A Systematic Review. *Simul Healthc*. 2020;15(2):106–11.
32. Desmedt M, Ulenaers D, Grosemans J, Hellings J, Bergs J. Clinical handover and handoff in healthcare: a systematic review of systematic reviews. *Int J Qual Health Care*. 2021;33(1):22–30.
33. Busby JS. An assessment of post-project reviews. *Project Manage J*. 1999;30(3):23–9.
34. Gordon M, Farnan J, Grafton-Clarke C, Ahmed R, Gurbutt D, McLachlan J, et al. Non-technical skills assessments in undergraduate medical education: a focused BEME systematic review: BEME Guide No. 54. *Med Teach*. 2019;41(7):732–45.
35. Rege NN. Assessment in competency-based medical education: a paradigm shift. *J Postgrad Med*. 2020;66(4):184–6.
36. Wooding EL, Gale TC, Maynard V. Evaluation of teamwork assessment tools for interprofessional simulation: a systematic literature review. *J Interprof Care*. 2020;34(2):162–72.
37. Coyne E, Calleja P, Forster E, Lin F. A review of virtual-simulation for assessing healthcare students' clinical competency. *Nurse Educ Today*. 2021;96:104623.
38. Davis J, Roach C, Elliott C, Mardis M, Justice EM, Riesenber LA. Feedback and assessment tools for handoffs: a systematic review. *J Grad Med Educ*. 2017;9(1):18–32.
39. Institute for Healthcare Improvement. SBAR Tool: Situation-Background-Assessment-Recommendation. 2016. <http://www.ihl.org/resources/Pages/Tools/SBARToolkit.aspx>. Accessed 10 May 2021.
40. Microsoft Corporation. Microsoft Teams. Version 4.7.15.0. 2020. Redmond, WA, USA:Microsoft Corporation. Available from: <https://www.microsoft.com/en-us/microsoft-teams/download-app> Accessed 4 Jan 2021.
41. Riesenber LA, Leitzsch J, Little BW. Systematic review of handoff mnemonics literature. *Am J Med Qual*. 2009;24(3):196–204.
42. Shahid S, Thomas S. Situation. Background, Assessment, Recommendation (SBAR) Communication Tool for Handoff in Health Care – A Narrative Review. *Safety in Health*. 2018;4(1).
43. Royal College of Physicians. Improving teams in healthcare: Resource 3 - Team communication. London, RCP policy: patient safety, Improving teams in healthcare; 2017. Available from: <https://www.rcplondon.ac.uk/projects/outputs/improving-teams-healthcare-resource-3-team-communication> Accessed 9 Jul 2021.
44. Salas E, Wilson KA, Murphy CE, King H, Salisbury M. Communicating. Coordinating, and cooperating when lives depend on it: tips for teamwork. *Joint Comm J Qual Patient Saf*. 2008;34(6):333–41.
45. Salas E, Shuffler ML, Thayer AL, Bedwell WL, Lazzara EH. Understanding and Improving Teamwork in Organizations: A Scientifically Based Practical Guide. *Hum Resour Manag*. 2015;54(4):599–622.
46. El-Shafy IA, Delgado J, Akerman M, Bullaro F, Christopherson NAM, Prince JM. Closed-Loop Communication Improves Task Completion in Pediatric Trauma Resuscitation. *J Surg Educ*. 2018;75(1):58–64.
47. Etherington N, Wu M, Cheng-Boivin O, Larrigan S, Boet S. Interprofessional communication in the operating room: a narrative review to advance research and practice. *Can J Anaesth*. 2019;66(10):1251–60.
48. Agarwala A, Lane-Fall M. The Evidence Base for Optimal Conduct of Handoffs. *APSF Newsletter*. 2017; 32(2):36–9. Available from: <https://www.apsf.org/article/the-evidence-base-for-optimal-conduct-of-handoffs/> Accessed 9 Jul 2019.
49. 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):203–9.
50. O'Brien CM, Flanagan ME, Bergman AA, Ebright PR, Frankel RM. "Anybody on this list that you're more worried about?" Qualitative analysis exploring the functions of questions during end of shift handoffs. *BMJ Qual Saf*. 2016;25(2):76–83.
51. Joint Commission. Inadequate hand-off communication. Sentinal Event Alert. 2017;58:1–6.
52. Agency for Healthcare Research and Quality. Handoffs and Signouts. 2019. <https://psnet.ahrq.gov/primer/handoffs-and-signouts>. Accessed 9 Jul 2021.
53. Kannampallil T, Abraham J. Listening and question-asking behaviors in resident and nurse handoff conversations: a prospective observational study. *J Am Med Inform Assoc*. 2020;3(1):87–93.
54. IBM Corp. IBM SPSS Statistics for Windows. Version 27.0. 2020. Armonk: IBM Corp. Available from: <https://www.ibm.com/products/spss-statistics> Accessed 2 May 2021.
55. Zapf A, Castell S, Morawietz L, Karch A. Measuring inter-rater reliability for nominal data - which coefficients and confidence intervals are appropriate? *BMC Med Res Methodol*. 2016;16(1):93.
56. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33(1):159–74.
57. Hayes AF, Krippendorff K. Answering the call for a standard reliability measure for coding data. *Commun Methods Measures*. 2007;1(1):77–89.

58. Hallgren KA. Computing inter-rater reliability for observational data: an overview and tutorial. *Tutor Quant Methods Psychol.* 2012;8(1):23–34.
59. McHugh ML. Interrater reliability: the kappa statistic. *Biochem Med (Zagreb).* 2012;22(3):276–82.
60. Krippendorff K. *Content analysis: an introduction to its methodology.* Thousand Oaks: Sage; 2004.
61. Starmer AJ, Landrigan C, Srivastava R, Wilson KM, Allen A, Mahant S, et al. I-PASS Handoff curriculum: faculty observation tools. *MedEdPORTAL.* 2013; 9(1):1–5.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

