

From Workshop to Simulation: Assessing I-PASS Handoff Skills in an Integrated Preclerkship Curriculum

Kenji Aoki, AB | Teresa Schiff-Elfalan, MD | Maya Maxym, MD | Leimomi Kanagusuku, MD | Kyle Ishikawa, MS | Jonathan Phisayavong, MS | Jannet Lee-Jayaram, MD

PRIMER. 2026;10:43.

Published: 10/21/2026 | DOI:[10.22454/PRIMER.2026.0065](https://doi.org/10.22454/PRIMER.2026.0065)

Abstract

Introduction: Effective communication between health care professionals is essential to patient care and outcomes yet is often underrepresented in preclerkship medical curricula. Efforts to introduce communication training early must balance clinical authenticity with learner readiness. In this workshop series, we explored how alignment with a problem-based learning (PBL) case can support the introduction and application of structured communication frameworks used in clerkships.

Methods: A total of 76 second-year medical students participated in a three-session curriculum culminating in a manikin-based simulation and associated I-PASS handoff. Student performances were video-recorded and independently evaluated by two physician-raters using a modified I-PASS assessment tool. Cohen's κ for interrater agreement was calculated for each component with a quadratic weighting applied to components with an ordinal scale.

Results: Across most assessed I-PASS components, ratings were distributed in the "moderate" and "great extent" categories. Students most consistently demonstrated patient summaries and action lists, while omissions in past medical history and task ownership were common. Interrater agreement was strongest for illness severity classification, on an ordinal scale, though interrater agreement was significant when using a dichotomized scale that distinguished high from low performance.

Conclusions: An aligned curriculum linking PBL, communication training, and simulation provides a feasible approach to introducing structured handoff frameworks in the preclerkship setting. Further work is needed to evaluate its impact on learner clerkship readiness and to improve the reliability of assessment tools.

Introduction

Communication failures among health care team members increase the number and severity of medical errors,¹ contributing to approximately two of every three sentinel events.² The Association of American Medical Colleges identified handoffs (Entrustable Professional Activity [EPA] 3) and team communication (EPA 6) as core competencies expected of entering residents. These expectations have prompted the introduction

of structured communication skill-building earlier in medical training.³⁻⁸ Specifically, the preclerkship phase has been identified as a place for integration of handoff training to establish longitudinal consistency and reinforcement, standardize training in formal settings, and foster student confidence upon entering their clerkship year.^{4-6,9,10} Prior interventions have shown improved confidence and knowledge, but have inconsistently examined learner performance in simulation.^{5,7,11,12}

One key barrier to early integration of clinical communication training is that the preclerkship curriculum emphasizes foundational scientific and medical knowledge, with limited exposure to the clinical complexity that frameworks like I-PASS are designed for.^{6,8,9} As a result, some curricula have simplified scenarios or used more basic frameworks like Simulation-Background-Assessment-Recommendation rather than the more comprehensive I-PASS,^{2,6} which standardizes handoffs and has been associated with improvements in patient safety.^{2,13-16}

To address this gap, we designed a communication curriculum for second-year medical students with no prior formal handoff training, aligning a problem-based learning (PBL) case with a subsequent workshop and related simulation. Using the same clinical content across three distinct modalities reduced the medical knowledge burden and allowed for increased complexity in communication training. The curriculum culminated in a team-based performance of an I-PASS handoff. We sought to describe student performance on I-PASS handoffs using video-based assessment with a modified assessment tool.

Methods

This study was reviewed by the University of Hawai'i Human Studies Program and determined not to meet the definition of human subjects research (Protocol number 2026-00149).

All 76 second-year medical students at a single institution, comprising the entire class, participated in three sessions at the end of their second year, prior to beginning clerkships. The first session started with a polysubstance intoxication PBL case designed by the Office of Medical Education (OME) faculty as part of the core curriculum. One week later, the students participated in a 2-hour communication workshop, designed and led by the block course director (a family physician and OME faculty) and director of medical student simulation education (a pediatric emergency physician and OME faculty), that translated clinical moments from the PBL case into communication with nurses, consultants, and residents. Workshop content was informed by a local needs assessment conducted by the designers of the workshop, including interviews with senior medical students, nurses, consulting physicians, and residents. Students rotated through small-group stations constructing and delivering nursing communication, consultation requests, and I-PASS handoffs. This was their first encounter with I-PASS in the preclerkship curriculum.

Three days later, students were divided into their 12 learning community groups and participated in a standing manikin simulation involving a polysubstance intoxication case, including acetaminophen and opioid misuse, like the PBL scenario. The simulation was designed by the director of medical student simulation education and modified that year to add a postsimulation I-PASS handoff. Afterward, each group collaboratively developed and presented an I-PASS handoff, which was received by a peer student performing the synthesis component and facilitated by the learning community faculty mentor, who also led the debrief. All sessions were video recorded and stored on a secure institutional server.

Two physician raters, who routinely use I-PASS in clinical practice and trainee supervision, independently evaluated all recordings. Raters were provided case details and rated each handoff as if assessing a medical student entering clerkships.¹⁰ Videos were independently rated to assess interrater reliability^{6,17} and viewed in a unique, randomized order to mitigate any learning effect.¹⁸

A modified I-PASS assessment tool was developed for this study (Appendix 1).¹⁹ The tool assessed the core I-PASS components: illness severity, patient summary, action list, situational contingency planning, and synthesis by receiver. Illness severity was scored as a binary variable (inaccurate/accurate). The remaining components were divided into subcomponents and rated on a 4-point ordinal scale (not at all, to a small extent, to a moderate extent, to a great extent). Raters also provided free-text comments to capture qualitative observations, detailed in Appendix 2. Interrater reliability was assessed using Cohen's κ with quadratic weighting using the psych package in R version 4.5.2 (R Foundation).²⁰ Agreement was calculated as absolute, near (within one category), and dichotomized: high (moderate/great extent) versus low (not at all/small extent).

Results

Two raters evaluated 12 simulated handoffs using the modified I-PASS assessment tool. Most ratings across nine of the 11 I-PASS subcomponents were in the high-performance categories (Table 1). Students performed strongest in providing to-do lists and summarizing, but frequently omitted pertinent past medical history and task ownership.

Absolute agreement and near agreement are presented in Tables 2 and 3, respectively. Cohen's κ was 0.8 for illness severity classification, while agreement for all other categories was low when using the 4-point ordinal scale (Table 4). Interrater agreement increased significantly when ratings were dichotomized into high/low categories (Table 5). Qualitative rater comments identified two recurring themes: inclusion of extraneous, nonclinically pertinent information; and difficulty distinguishing between action lists and contingency planning.

Discussion

This study described preclerkship learner handoff performance after curricular introduction of the I-PASS framework. Although interrater agreement was low across most elements using a 4-point scale, student ratings were in the "moderate" and "great extent" range for most subcomponents except "Patient summary: Provides an ongoing assessment and relevant plan" and "Action list: provides appropriate ownership and timeline for to-do list." This finding may reflect students' ability to generally meet baseline expectations for those matriculating into clerkship training. This discrepancy suggests that while the tool may have inherent limitations, such as ambiguous scoring criteria between ordinal categories, it may still serve as one evaluation tool to gauge student understanding of the handoff concept and its utility in patient care. By dichotomizing results into high and low, raters were able to reliably distinguish adequate execution from developing skills for most items. This finding suggests that the curriculum and assessment tool may have utility for early identification of common areas for continued communication development in medical students to support their transition into clerkship. Future iterations could improve reliability through structured rater calibration and clearer behavioral anchors.^{4,6,7}

While previous studies have described early integration of communication frameworks,^{3-8,10,11,17} this study highlights an aligned curricular design linking PBL, a communication workshop, and simulation through shared clinical context. Our approach may help integrate foundational preclerkship knowledge with the application of clinically contextualized communication, avoiding oversimplified clinical scenarios or substituting less comprehensive frameworks.^{4,6,11}

Recurrent qualitative comments regarding extraneous information and difficulty distinguishing between action items and contingency plans likely reflected the developmental stage of second-year medical students.

For example, students generally scored lower in the “provides appropriate ownership and timeline for to-do list” subcomponent of the action list, likely due to limited clinical exposure and time spent in an interprofessional care team. Future curricular efforts may benefit from an assessment of learner readiness, such as baseline familiarity of interprofessional roles, prior to introducing advanced communication frameworks. While students were introduced to the concept of designating ownership in the communication workshop session of the curriculum, a single session was likely insufficient to produce competency, particularly when this communication skill is not often emphasized in preclerkship curricula. Opportunities for structured oral communication within clinical contexts remain limited in the preclerkship curriculum; increasing scaffolded practice may support skill development as students transition to clerkship training.^{4-6,9,10}

This study was conducted at a single institution with a single cohort and lacked pre- and postintervention comparisons. The modified I-PASS tool was not formally validated, and minimal rater calibration likely contributed to low interrater reliability, especially for items that could have widely varying rater thresholds for desired clinical detail or concision, including subcomponents like providing an assessment, appropriate ownership, and effective questions. Qualitative analysis was limited to brief thematic grouping of rater comments.

Conclusions

Despite these limitations, this study provides a descriptive account of preclerkship student performance using the advanced I-PASS communication framework within a longitudinally aligned curriculum. Future work may explore whether similar curricular integration of foundational knowledge with clinical skills during the preclerkship curriculum impacts readiness for clerkship responsibilities and longitudinal student progress after curricular interventions.

Acknowledgments

Biostatisticians Kyle Masato Ishikawa and Jonathan Phisayavong are partially supported by the National Institute on Minority Health and Health Disparities (Grant Number U54MD007601) and the National Institute of General Medical Sciences, National Institutes of Health (NIH; Grant Number U54GM138062). The content of this study is solely the responsibility of the authors and does not necessarily represent the official views of the NIH.

Presentations

Biomedical Sciences Symposium, John A. Burns School of Medicine, April 25, 2026, Honolulu, HI.

Corresponding Author

Kenji Aoki

John A. Burns School of Medicine, University of Hawai‘i at Mānoa, Honolulu, HI

kenji292@hawaii.edu

Author Affiliations

Kenji Aoki, AB - John A. Burns School of Medicine, University of Hawai‘i at Mānoa, Honolulu, HI

Teresa Schiff-Elfalan, MD - John A. Burns School of Medicine, University of Hawai‘i at Mānoa, Honolulu, HI

| Department of Family Medicine and Community Health, John A. Burns School of Medicine, University of Hawai‘i, Honolulu, HI

Maya Maxym, MD - John A. Burns School of Medicine, University of Hawai'i at Mānoa, Honolulu, HI |

Department of Pediatrics, Kapi'olani Medical Center for Women & Children, Honolulu, HI

Leimomi Kanagusuku, MD - John A. Burns School of Medicine, University of Hawai'i at Mānoa, Honolulu, HI | Department of Family Medicine and Community Health, John A. Burns School of Medicine, University of Hawai'i, Honolulu, HI

Kyle Ishikawa, MS - Department of Quantitative Health Sciences, John A. Burns School of Medicine, University of Hawai'i at Mānoa, Honolulu, HI

Jonathan Phisayavong, MS - Department of Quantitative Health Sciences, John A. Burns School of Medicine, University of Hawai'i at Mānoa, Honolulu, HI

Jannet Lee-Jayaram, MD - John A. Burns School of Medicine, University of Hawai'i at Mānoa, Honolulu, HI

References

1. McCarthy S, Motala A, Lawson E, Shekelle PG. Use of structured handoff protocols for within-hospital unit transitions: a systematic review from Making Healthcare Safer IV. *BMJ Qual Saf.* 2025;34(10):680–690. doi:10.1136/bmjqs-2024-018385
2. Starmer AJ, Spector ND, O'Toole JK, et al; I-PASS SHM Mentored Implementation Study Group. Implementation of the I-PASS handoff program in diverse clinical environments: a multicenter prospective effectiveness implementation study. *J Hosp Med.* 2023;18(1):5–14. doi:10.1002/jhm.12979
3. Kulaylat AN, Saunders BD, Allen SR, Aarons CB, Hoffman RL. Evaluating the implementation of a medical student's handoff curriculum during the surgery clerkship. *J Surg Res.* 2023;282:262–269. doi:10.1016/j.jss.2022.10.005
4. Reyes JA, Greenberg L, Amdur R, Gehring J, Lesky LG. Effect of handoff skills training for students during the medicine clerkship: a quasi-randomized study. *Adv Health Sci Educ Theory Pract.* 2016;21(1):163–173. doi:10.1007/s10459-015-9621-1
5. Thaeter L, Schröder H, Henze L, et al. Handover training for medical students: a controlled educational trial of a pilot curriculum in Germany. *BMJ Open.* 2018;8(9):e021202. doi:10.1136/bmjopen-2017-021202
6. Michael M, Griggs AC, Shields IH, et al. Improving handover competency in preclinical medical and health professions students: establishing the reliability and construct validity of an assessment instrument. *BMC Med Educ.* 2021;21(1):518. doi:10.1186/s12909-021-02943-x
7. Carter K, Golden A, Martin S, et al. Results from the first year of implementation of CONSULT: consultation with novel methods and simulation for UME longitudinal training. *West J Emerg Med.* 2015;16(6):845–850. doi:10.5811/westjem.2015.9.25520
8. Maraccini AM, Houmanfar RA, Kemmelmeier M, Piasecki M, Slonim AD. An inter-professional approach to train and evaluate communication accuracy and completeness during the delivery of nurse-physician student handoffs. *J Interprof Educ Pract.* 2018;12:65–72. doi:10.1016/j.xjep.2018.06.003
9. O'Toole JK, Stevenson AT, Good BP, et al; I-PASS Study Group. Closing the gap: a needs assessment of medical students and handoff training. *J Pediatr.* 2013;162(5):887–888. doi:10.1016/j.jpeds.2013.01.045
10. Kazmi M, Wong S, Conklin P, Cohen D, Wackett A, Lu WH. Preparation for residency: effect of formalized patient handover instruction for fourth-year medical students.. *Int Med Educ.* 2025;4(2):10. doi:10.3390/ime4020010
11. Higgins Joyce A. Team-based simulation for medical student handoff education. *MedEdPORTAL.* 2016;12:10486. doi:10.15766/mep_2374-8265.10486
12. Fessell D, McKean E, Wagenschutz H, et al. Medical improvisation training for all medical students: 3-year experience. *Med Sci Educ.* 2020;30(1):87–90. doi:10.1007/s40670-019-00885-0

13. Starmer AJ, Spector ND, Srivastava R, et al; I-PASS Study Group. Changes in medical errors after implementation of a handoff program. *N Engl J Med*. 2014;371(19):1803–1812. doi:10.1056/NEJMsa1405556
14. Sheth S, McCarthy E, Kipps AK, et al. Changes in efficiency and safety culture after integration of an I-PASS-supported handoff process. *Pediatrics*. 2016;137(2):e20150166. doi:10.1542/peds.2015-0166
15. Sonoda K, Nakaishi L, Salter C. Standardizing sign-out with I-PASS handoff in family medicine residency. *PRiMER*. 2021;5:8. doi:10.22454/PRiMER.2021.678175
16. Klamen DL, Reynolds KL, Yale B, Aiello M. Students learning handovers in a simulated in-patient unit. *Med Educ*. 2009;43(11):1097–1098. doi:10.1111/j.1365-2923.2009.03488.x
17. Kessler CS, Tadisina KK, Saks M, et al. The 5Cs of consultation: training medical students to communicate effectively in the emergency department. *J Emerg Med*. 2015;49(5):713–721. doi:10.1016/j.jemermed.2015.05.012
18. Kooken J, Welsh ME, McCoach DB, et al. Test order in teacher-rated behavior assessments: is counterbalancing necessary? *Psychol Assess*. 2017;29(1):98–109. doi:10.1037/pas0000314
19. Starmer AJ, Landrigan C, Srivastava R, et al. I-PASS Education Executive Committee. I-PASS handoff curriculum: faculty observation tools. *MedEdPORTAL*. 2013;9:9570. doi:10.15766/mep_2374-8265.9570
20. Revelle W. Psych: Procedures for Psychological, Psychometric, and Personality Research. *Personality Project*. 2007;2. doi:10.32614/CRAN.package.psych

Copyright © 2026 by the Society of Teachers of Family Medicine

Table 1. Ratings by Two Independent Raters for Each I-PASS Component Based on Modified I-PASS Tool

Component	Rater 1 N = 12 n (%)	Rater 2 N = 12 n (%)
I—Illness severity: Accurate categorization of patient as stable, watcher, or unstable		
Accurate categorization	9 (75)	8 (67)
Inaccurate categorization	3 (25)	4 (33)
P—Patient summary: Provides a history of events leading up to admission		
Not at all	0	0
To a small extent	0	0
To a moderate extent	12 (100)	5 (42)
To a great extent	0	7 (58)
P—Patient summary: Provides a pertinent past medical history of patient		
Not at all	9 (75)	11 (92)
To a small extent	1 (8.3)	1 (8.3)
To a moderate extent	2 (17)	0
To a great extent	0	0
P—Patient summary: Provides a history of hospital course		
Not at all	0	0
To a small extent	0	0
To a moderate extent	11 (92)	6 (50)
To a great extent	1 (8.3)	6 (50)
P—Patient summary: Provides an ongoing assessment and relevant plan		
Not at all	0	4 (33)
To a small extent	0	4 (33)
To a moderate extent	11 (92)	3 (25)
To a great extent	1 (8.3)	1 (8.3)
A—Action list: Provides a to-do list		
Not at all	0	0
To a small extent	0	0
To a moderate extent	10 (83)	5 (42)
To a great extent	2 (17)	7 (58)
A—Action list: Provides appropriate ownership and timeline for to-do list		

(Continued)

**Table 1. Ratings by Two Independent Raters for Each I-PASS Component Based on Modified I-PASS Tool
(continued)**

Component	Rater 1 N = 12 n (%)	Rater 2 N = 12 n (%)
Not at all	6 (50)	3 (25)
To a small extent	3 (25)	2 (17)
To a moderate extent	3 (25)	4 (33)
To a great extent	0	3 (25)
S—Situation awareness and contingency planning: Provides appropriate insight into potential problems		
Not at all	0	0
To a small extent	1 (8.3)	2 (17)
To a moderate extent	10 (83)	6 (50)
To a great extent	1 (8.3)	4 (33)
S—Situation awareness and contingency planning: Offers suggested plans for each problem		
Not at all	0	0
To a small extent	1 (8.3)	5 (42)
To a moderate extent	11 (92)	4 (33)
To a great extent	0	3 (25)
S—Synthesis by receiver: Summarizes what was heard including restating key to-do list items		
Not at all	0	0
To a small extent	0	0
To a moderate extent	6 (50)	8 (67)
To a great extent	6 (50)	4 (33)
S—Synthesis by receiver: Asks questions to better clarify the patient's condition		
Not at all	1 (8.3)	4 (33)
To a small extent	1 (8.3)	0
To a moderate extent	6 (50)	2 (17)
To a great extent	4 (33)	6 (50)

Table 2. Absolute Agreement Score Calculated for Each I-PASS Component

Component	N = 12 n (%)
I—Illness severity: Accurate categorization of patient as stable, watcher, or unstable	11 (92)
P—Patient summary: Provides a history of events leading up to admission	5 (42)
P—Patient summary: Provides a pertinent past medical history of patient	8 (67)
P—Patient summary: Provides a history of hospital course	5 (42)
P—Patient summary: Provides an ongoing assessment and relevant plan	3 (25)
A—Action list: Provides a to-do list	7 (58)
A—Action list: Provides appropriate ownership and timeline for to-do list	1 (8.3)
S—Situation awareness and contingency planning: Provides appropriate insight into potential problems	6 (50)
S—Situation awareness and contingency planning: Offers suggested plans for each problem	5 (42)
S—Synthesis by receiver: Summarizes what was heard including restating key to-do list items	8 (67)
S—Synthesis by receiver: Asks questions to better clarify the patient's condition	7 (58)

Table 3. Near Agreement Score Calculated for Each I-PASS Component

Component	N = 12 n (%)
I—Illness severity: Accurate categorization of patient as stable, watcher, or unstable	11 (92)
P—Patient summary: Provides a history of events leading up to admission	5 (42)
P—Patient summary: Provides a pertinent past medical history of patient	8 (67)
P—Patient summary: Provides a history of hospital course	5 (42)
P—Patient summary: Provides an ongoing assessment and relevant plan	3 (25)
A—Action list: Provides a to-do list	7 (58)
A—Action list: Provides appropriate ownership and timeline for to-do list	1 (8.3)
S—Situation awareness and contingency planning: Provides appropriate insight into potential problems	6 (50)
S—Situation awareness and contingency planning: Offers suggested plans for each problem	5 (42)
S—Synthesis by receiver: Summarizes what was heard including restating key to-do list items	8 (67)
S—Synthesis by receiver: Asks questions to better clarify the patient's condition	7 (58)

Table 4. Kappa Scores and Rating for Each Component of I-PASS

Component	κ score (confidence interval low–high)	κ rating
I—Illness severity: Accurate categorization of patient as stable, watcher, or unstable	0.800 (0.432–1.000)	Strong
P—Patient summary: Provides a history of events leading up to admission	0	None
P—Patient summary: Provides a pertinent past medical history of patient	–0.091 (–0.258–0.076)	None
P—Patient summary: Provides a history of hospital course	–0.167 (–0.475–0.142)	None
P—Patient summary: Provides an ongoing assessment and relevant plan	–0.091 (–0.232–0.050)	None
A—Action list: Provides a to-do list	0.250 (–0.082–0.582)	Minimal
A—Action list: Provides appropriate ownership and timeline for to-do list	–0.079 (–0.499–0.340)	None
S—Situation awareness and contingency planning: Provides appropriate insight into potential problems	0.250 (–0.085–0.585)	Minimal
S—Situation awareness and contingency planning: Offers suggested plans for each problem	0.192 (–0.119–0.504)	None
S—Synthesis by receiver: Summarizes what was heard including restating key to-do list items	0.333 (–0.170–0.836)	Minimal
S—Synthesis by receiver: Asks questions to better clarify the patient's condition	0.457 (0.055–0.860)	Weak

Note: Rating was determined using Cohen's interpretation of κ . Illness severity was calculated separately as a binary rating. Kappa scores for I-PASS components were calculated as a Likert rating and with additional quadratic weighting.

Table 5. Interrater Dichotomized Scores and Rating for I-PASS Components

Component	Interrater agreement N = 12 n (%)
P—Patient summary: Provides a history of events leading up to admission	12 (100)
P—Patient summary: Provides a pertinent past medical history of patient	10 (83)
P—Patient summary: Provides a history of hospital course	12 (100)
P—Patient summary: Provides an ongoing assessment and relevant plan	8 (67)
A—Action list: Provides a to-do list	12 (100)
A—Action list: Provides appropriate ownership and timeline for to-do list	6 (50)
S—Situation awareness and contingency planning: Provides appropriate insight into potential problems	12 (100)
S—Situation awareness and contingency planning: Offers suggested plans for each problem	12 (100)
S—Synthesis by receiver: Summarizes what was heard including restating key to-do list items	12 (100)
S—Synthesis by receiver: Asks questions to better clarify the patient's condition	8 (67)

Note: Interrater agreement expressed after dichotomizing the modified I-PASS tool. Kappa scores and rating for each component of I-PASS. For the purposes of dichotomization, "not at all" and "to a small extent" scores were consolidated into a "low" category; "to a moderate extent" and "to a great extent" scores were consolidated into a "high" category. Interrater agreement is expressed as a percentage, noting how many times the raters agreed on a group I-PASS performance being in the low or high category. Illness severity was excluded because it already was rated in a binary setting.