Academic Medicine

DOI: 10.1097/ACM.0000000000004573

The Effects of Leadership Curricula With and Without Implicit Bias Training on Graduate Medical Education: A Multicenter Randomized Trial

Matt Hansen, MD, MCR, Tabria Harrod, MPH, Nathan Bahr, PhD, Amanda Schoonover, MPH, Karen Adams, MD, Josh Kornegay, MD, Amy Stenson, MD, Vivienne Ng, MD, MPH, Jennifer Plitt, MD, Dylan Cooper, MD, Nicole Scott, MD, Sneha Chinai, MD, Julia Johnson, MD, Lauren Weinberger Conlon, MD, Catherine Salva, MD, Holly Caretta-Weyer, MD, MHPE(c), Trang Huynh, MD, David Jones, MD, MBS, MCR, Katherine Jorda, MD, Jamie Lo, MD, Ryanne Mayersak, MD, MS, Emmanuelle Paré, MD, MSCE, Kate Hughes, DO, Rami Ahmed, DO, MHPE, Soha Patel, MD, MSPH, Suzana Tsao, DO, Eileen Wang, MD, Tony Ogburn, MD, and Jeanne-Marie Guise, MD, MPH

M. Hansen is associate professor of emergency medicine and pediatrics, Oregon Health & Science University School of Medicine, Portland, Oregon.

T. Harrod is senior research associate, Department of Obstetrics and Gynecology, Oregon Health & Science University School of Medicine, Portland, Oregon.

N. Bahr is senior research associate, Department of Obstetrics and Gynecology, Oregon Health & Science University School of Medicine, Portland, Oregon.

A. Schoonover is senior research assistant, Department of Obstetrics and Gynecology, Oregon Health & Science University School of Medicine, Portland, Oregon.

K. Adams is professor of obstetrics and gynecology, Oregon Health & Science University School of Medicine, Portland, Oregon.

J. Kornegay is associate professor of emergency medicine, Oregon Health & Science University School of Medicine, Portland, Oregon.

A. Stenson is associate professor of obstetrics and gynecology, Oregon Health & Science University School of Medicine, Portland, Oregon.

V. Ng is assistant professor of emergency medicine, The University of Arizona College of Medicine, Tucson, Arizona.

J. Plitt is assistant clinical professor of emergency medicine, University of Arizona College of Medicine, Tucson, Arizona.

D. Cooper is professor of clinical emergency medicine, Indiana University School of Medicine, Bloomington, Indiana.

N. Scott is assistant professor of clinical obstetrics and gynecology, Indiana University School of Medicine, Bloomington, Indiana.

S. Chinai is assistant professor of emergency medicine, University of Massachusetts Chan Medical School, Boston, Massachusetts.

J. Johnson is professor of obstetrics and gynecology, University of Massachusetts Chan Medical School, Boston, Massachusetts.

L.W. Conlon is assistant professor of emergency medicine, Perelman School of Medicine at the University of Pennsylvania,

C. Salva is associate professor of clinical obstetrics and gynecology, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, Pennsylvania.

H. Caretta-Weyer is assistant director, Emergency Medicine Residency Program, Stanford University School of Medicine, Stanford, California.

T. Huynh is associate professor of pediatrics, Oregon Health & Science University School of Medicine, Portland, Oregon.

D. Jones is associate professor of emergency medicine, Oregon Health & Science University School of Medicine, Portland, Oregon.

K. Jorda is assistant professor of obstetrics and gynecology, Oregon Health & Science University School of Medicine, Portland, Oregon.

J. Lo is assistant professor of obstetrics and gynecology, Oregon Health & Science University School of Medicine, Portland, Oregon.

R. Mayersak is assistant professor of emergency medicine, Oregon Health & Science University School of Medicine, Portland, Oregon.

E. Paré is associate professor of obstetrics and gynecology, Oregon Health & Science University School of Medicine, Portland, Oregon.

K. Hughes is assistant professor of emergency medicine, The University of Arizona College of Medicine, Tucson, Arizona.

R. Ahmed is professor of emergency medicine, Indiana University School of Medicine, Bloomington, Indiana.

S. Patel is assistant professor of obstetrics and gynecology, Vanderbilt University School of Medicine, Nashville, Tennessee.

S. Tsao is associate professor of emergency medicine, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, Pennsylvania.

E. Wang is professor of obstetrics and gynecology, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, Pennsylvania.

T. Ogburn is professor and chair of obstetrics and gynecology, University of Texas Rio Grande Valley School of Medicine, Edinburgh, Texas.

J.M. Guise is professor of obstetrics and gynecology, Oregon Health & Science University School of Medicine, Portland, Oregon.

Correspondence should be addressed to Matt Hansen, CR 114, 3181 SW Sam Jackson Pk Rd, Portland, OR 97239; email: hansemat@ohsu.edu; Twitter: @MattHansenMD.

Supplemental digital content for this article is available at http://links.lww.com/ACADMED/B220.

Acknowledgments: The authors wish to thank the following people for helping in the development of the curriculum and facilitating simulations and other activities to ensure the success of this study: David Biffar, Ameer F. Ibrahim, Barbara Kanki, Lee Learman, Diane Magrane, Maria Manriquez, James McNulty, Joan Nagel, Meg O'Reilly, and Howard Ross.

Funding/Support: Funding support for this study was provided by Agency for Healthcare Quality and Research R18 HS23457. The funding body did not have a role in study design, data collection, analysis, or interpretation.

Other disclosures: None reported.

Ethical approval: This study was approved by the Institutional Review Board of Oregon Health & Science University, IRB00011053. Participants provided consent prior to participation.

Abstract

Purpose

To determine whether a brief leadership curriculum including high-fidelity simulation can improve leadership skills among resident physicians.

Method

This was a double-blind randomized controlled trial among obstetrics and gynecology (OB/GYN) and emergency medicine (EM) residents across 5 academic medical centers from different geographic areas of the United States, 2015–2017. Participants were assigned to 1 of 3 study arms: the LEADS (Leadership Education Advanced During Simulation) curriculum, a shortened TeamSTEPPS (Team Strategies and Tools to Enhance Performance and Patient Safety) curriculum, or as active controls (no leadership curriculum). Active controls were recruited from a separate site and not randomized in order to limit any unintentional introduction of materials from the leadership curricula. The LEADS curriculum was developed in partnership with the Council on Resident Education in Obstetrics and Gynecology and Council of Residency Directors in Emergency Medicine as a novel way to provide a leadership toolkit. Both LEADS and the abbreviated TeamSTEPPS were designed as six 10-minute interactive web-based modules.

The primary outcome of interest was the leadership performance score from the validated Clinical Teamwork Scale instrument measured during standardized high-fidelity simulation scenarios. Secondary outcomes were 9 key components of leadership from the detailed leadership evaluation measured on 5-point Likert scales. Both outcomes were rated by a blinded clinical video reviewer.

Results

One hundred and ten OB/GYN and EM residents participated in this 2-year trial. Participants in both LEADS and TeamSTEPPS had statistically significant improvement in leadership scores from "average" to "good" ranges both immediately and at the 6-month follow-up, while controls remained unchanged in the "average" category throughout the study. There were no differences between the LEADS and TeamSTEPPS curricula with respect to the primary outcome.

Conclusions

Residents who participated in a brief structured leadership training intervention had improved leadership skills that were maintained at 6-month follow-up.

Leadership skills are critical to the safe and effective practice of medicine. The Joint Commission reports that inadequate leadership is associated with half of the serious injuries and unanticipated deaths in the United States health care system.[1] There is wide recognition that leadership training is critical in graduate medical education.[2–5] Graduating resident physicians are expected to competently lead multidisciplinary health care teams during stressful situations where life and limb are on the line. Despite the critical importance of physician leadership skills, leadership training is often unstructured or not included in graduate medical education.[6] There is little evidence to guide medical educators in identifying the most effective leadership training. Also, many studies of leadership training have utilized self-reported outcomes or analyzed participant satisfaction.[7]

There is also increasing recognition about the pervasiveness of implicit bias in health care. Specifically, the role of gender in health care teams has been the subject of psychology research for decades.[8–11] One study found that female residents tend to choose less assertive behaviors in clinical scenarios compared with their male counterparts.[12] The same study found that 88% of female and 49% of male internal medicine residents rated gender as one of the top 3 disadvantages in directing patient care. In a previously published component of this study, emergency medicine (EM) and obstetrics–gynecology (OB/GYN) residents identified gender as playing a major role in leadership effectiveness, with 55% of female residents reporting their own gender to be harmful in being perceived as a leader.[13]

In this randomized controlled trial (RCT), our primary aim was to evaluate the efficacy of 2 targeted leadership training curricula focused on acute care of a critical patient. The curricula focused on team leadership skills such as directing communication, assigning roles, developing a shared mental model, prioritizing tasks, and providing summaries and anticipatory guidance.

Each curriculum included an asynchronous/online didactic module with an in-person simulation, in comparison to a control group with in-person simulation only. One of the 2 curricula was a shortened version of the popular TeamSTEPPS (Team Strategies and Tools to Enhance Performance and Patient Safety) course, focused on leadership training.[14] The other was a customized leadership curriculum that included implicit bias training and allowed the development of an individualized leadership toolkit created by the study team in partnership with an expert advisory panel called LEADS (Leadership Education Advanced During Simulation). We hypothesized that the personalized LEADS curriculum would result in improved leadership assessed during a standardized high-fidelity simulation scenario.

Method

Study design

This was a prospective 3-arm double-blinded RCT. Participants, evaluators, and analysts were blinded to the assignment arm. We randomized participants from 4 study sites (Oregon Health & Science University, University of Indiana School of Medicine, University of Massachusetts Chan Medical School, Perelman School of Medicine at the University of Pennsylvania) to participate in 1 of 2 leadership training curricula and a fifth site (The University of Arizona College of Medicine) served as an active control. Outcomes were evaluated by blinded expert video review from standardized high-fidelity simulations where participants led care for critically ill simulated patients. Expert reviewers were drawn from the study sites (K.A., J.K., A.S., V.N., J.P., D.C., N.S., S.C., L.W.C., C.S., H.C.W., T.H., D.J., K.J., J.L., R.M., E.P., K.H., R.A., S.P., S.T., E.W.) and were blinded to the scenario and the specialty of the participant. The trial was conducted between August 2015 and September 2017 and was approved by the institutional review board of Oregon Health & Science University (IRB00011053), as well as all other participating sites.

Study participants

Participants in this trial were EM and OB/GYN resident physicians from 5 geographically diverse academic residency programs. We selected these 2 specialties as OB/GYN is traditionally female-dominated while EM is traditionally male-dominated.[15] In addition, resident physicians from both specialties are expected to provide leadership in initial stabilization of an overlapping set of acute medical emergencies.

We randomized participants from 4 study sites into 1 of 2 leadership curricula while residents from the fifth site served as an active control (simulation without leadership curricula). We recruited this fifth site as the active control site to limit any unintentional introduction of materials from the leadership curricula, simulation interventions, and structured debriefing tools the intervention sites used with participants. Participants were excluded from the study if they were in the last year of their training program, as the study spanned 2 academic years.

Intervention

The 2 intervention arms of the study consisted of 2 different leadership curricula, each with an asynchronous/online didactic training program of 60 minutes duration and an associated standardized simulation scenario and structured debriefing, designed to reinforce specific learning objectives from the online training. The active control arm consisted of the same simulation scenarios as the intervention arms, with a clinically focused unstructured debriefing without specific leadership training conducted according to usual practice at the site. Although confederates were involved in the simulations, they did not participate in the debriefings. One intervention arm included a modified TeamSTEPPS essentials program in which the curriculum was shortened to elements most relevant for leadership presented as a 1-hour interactive training module. The second intervention arm consisted of the LEADS curriculum specifically created for this study. In addition to the incorporation of leadership and teamwork principles also found in TeamSTEPPS, LEADS was specifically designed to address implicit bias regarding sex, race, and other factors considered to be barriers to effective leadership. A previous component of this study confirmed that implicit bias against female EM and OB/GYN residents as team leaders existed; this bias, as well as others, are barriers to effective leadership.[12] Curricular development for LEADS was guided by a national panel of experts in leadership, education, simulation, implicit bias, and graduate medical education. The LEADS curriculum consisted of six 10-minute interactive training modules. Both LEADS and TeamSTEPPS curricula involved reinforcement of leadership skills through a standardized simulation scenario followed by a structured debriefing designed to address the elements in the toolkit.

Trial design

Each participant led 4 high-fidelity simulation scenarios that occurred across 3 simulation sessions, with the second session including 2 scenarios (Figure 1). All simulation scenarios involved a high-fidelity mannequin and 2 embedded participants (confederates) as nursing staff, with EM or OB/GYN faculty observing and facilitating the structured debriefings. Participants established their baseline leadership performance in the first simulation scenario (simulation 1), a cardiac arrest in a non-pregnant female patient. Following the completion of this scenario, participants (excluding those in the control site) were individually randomized into one of the 2 intervention study arms and were asked to complete the assigned online curriculum at their convenience. Within 6 weeks of completing the curriculum, subjects participated in a second simulation session. During this session, each participants took part in 2 simulation scenarios: the first involved a neonatal resuscitation following a precipitous birth and the second required management of a pregnant patient with eclampsia. Both simulation scenarios (simulations 2 and

3) concluded with structured debriefings. The purpose of the neonatal resuscitation scenario was to reinforce skills highlighted in the online training modules. Our primary outcome, leadership performance after curriculum exposure, was measured during the second scenario of the second session (simulation 3), in which participants managed eclampsia. Finally, 3 to 6 months later, participants returned for a third simulation session involving one scenario (simulation 4), a female patient with pyelonephritis and sepsis. All scenarios were designed by a collaborative group of EM and OB/GYN physicians including residency leadership from both specialties to ensure the scenarios would apply to all residents involved and focused on the initial stabilization of the patient. Our secondary outcome was leadership performance assessed 3 to 6 months after curriculum exposure. This was intended to evaluate potential treatment effect over time.

Study outcomes

The primary outcome was the leadership performance score on the Clinical Teamwork Scale (CTS), which is a validated tool commonly used in simulation, education, and clinical arenas.[16–22] The leadership performance score has a range of 0–10, with 0 being "unacceptable," 1–3 being "poor," 4–6 being "average," 7–9 being "good," and 10 being "perfect" leadership. We also conducted a detailed leadership evaluation (DLE), which was adapted for this study from the previously validated Leadership Behavior Description Questionnaire (LBDQ) and included 9 specific leadership components scored on a 5-point Likert-type scale. Faculty physicians with subject matter expertise performed blinded video reviews of the simulation scenarios to score participant leadership performance using the aforementioned tools. We blinded video reviewers to the study arm of the participants they were scoring and the specialty of participants and did not review participant videos from their own site. The video reviewers were trained by the study team on the CTS and DLE instruments. After initial training, they completed the two tools (15 elements for CTS and 24 for DLE) for 3 “gold standard” videos with scores established by consensus of the research team. If any single reviewer was more than 3 points different than the gold standard, there was additional feedback, training, and testing until this standard was met (this process was only required for 1 individual reviewer).

Randomization

The leadership curriculum (LEADS or TeamSTEPPS) was allocated to each participant using a random number generator in Microsoft Excel blocked on institution and specialty.

Statistical analysis

We employed descriptive statistics to depict demographic characteristics by intervention arm. Pearson’s chi-squared was used to test for associations between intervention arm and demographic variables: sex, postgraduate year, specialty, and institution. Unadjusted scores for the primary outcomes of interest—leadership performance CTS and composite DLE—for the baseline simulation session 1 (cardiac arrest), post-intervention simulation session 2 (eclampsia), and post follow-up simulation session 3 (sepsis) were described using mean and standard deviation. We included only individuals with data for all simulations in analysis.

In this 3-arm RCT, we used linear mixed-effects models to compare leadership performance between participants in the LEADS vs TeamSTEPPS curricula, and between participants in both intervention curricula (LEADS or TeamSTEPPS) compared with active controls (no curriculum) at the initial assessment simulation (session 2, simulation 3) and 3- to 6- month follow-up (session 3, simulation 4). The treatment effect is the mean difference in leadership scores between intervention curricula and controls at simulations 3 and 4, controlling for the baseline leadership score obtained during simulation 1. The regression models included a random effect for institution to account for clustering. Likelihood ratio tests (LRTs) were used to obtain *P* values, and all *P* values less than .05 were considered statistically significant. We conducted all analyses using STATA statistical software, version 15 (StataCorp TM, College Station, TX). Similarly, we used baseline-adjusted linear mixed-effects models to analyze the 9 key components of leadership from the DLE. These models included a random effect for institution to account for clustering.

Last, we conducted post-hoc analyses repeating the same mixed-effects models exclusively among EM residents, since they were the only specialty in the control group.

Sample size calculation

Prior to starting the study, our sample size calculation found that a sample size of 120 residents would provide 90% power to detect a difference of 1.5 points (standard deviation = 1.7) in the CTS leadership performance between the intervention arms.

Results

In total, 155 residents agreed to participate and completed the baseline simulation. One hundred and thirty-eight (89%) residents were randomized to either LEADS (n = 59) or TeamSTEPPS (n = 58) and participated in the asynchronous leadership curricula. Seventeen residents did not return for second session and were therefore not randomized. All residents at 1 site were assigned to serve as the active control group (n=20 EM residents). A total of 110 (80% of those randomized, LEADS n = 42, TeamSTEPPS n = 48, control n = 20) residents from all arms completed all 4 simulations and were included in the final analysis. Twenty-eight individuals were not included in the final analysis due to not participating in all simulation scenarios, voluntary study withdrawal after randomization, or poor video or audio quality making it impossible to assess performance using video review. Ninety percent (n = 37/41) of OB/GYN participants were female compared with 40% of EM participants. Table 1 displays the demographic characteristics of study participants in each study arm that completed the study.

Leadership performance

Unadjusted means and standard deviations for both leadership performance CTS and for composite DLE are shown in Table 2. After exposure to either LEADS or TeamSTEPPS curricula, participants experienced an improvement in mean CTS from the "average" to "good" range, while participants in the active control remained in the "average" CTS range throughout all phases of assessment (Figure 2). In linear mixed-effects regression modeling, both LEADS (*P* value = .043) and TeamSTEPPS (*P* value = .047) arms were associated with statistically significant improvement in leadership scores between simulation 1 (baseline) and simulation 3 (post-intervention) compared to controls. These effects persisted after 3- to 6-month follow-up for both LEADS (B = 1.09; 95% confidence interval [CI] = 0.11, 2.08; *P* value = .03) and TeamSTEPPS (B = 1.63; 95% CI = 0.66, 2.59; *P* value = .001) (Table 3).

When comparing LEADS and TeamSTEPPS participants directly, there were no significant differences in CTS scores at either simulation 3 (B = 0.06; 95% CI = 0.72, 0.85; *P* value = .87) or at simulation 4 (B = -0.52; 95% CI = -1.24, 0.21; *P* value = .16) after controlling for baseline scores and with TeamSTEPPS as the reference group.

Detailed leadership evaluation

There were no significant differences in the DLE components between participants in the various arms at baseline. Unadjusted means for DLE components at baseline, simulation 3, and simulation 4 are found in Supplemental Digital Appendix 1, at http://links.lww.com/ACADMED/B220. At post-intervention simulation 3, residents in the LEADS curriculum showed significantly greater improvement in degree of leadership (*P* value = .003), problem solving (*P* value = .02), and shared knowledge (*P* value = .001) compared with the control arm. After the follow-up period, differences persisted for both problem-solving (*P* value = .04) and shared knowledge (*P* value = .03), however, degree of leadership remained only marginally significant (*P* value = .06). LEADS participants also showed significantly greater improvement in deciding what should be done (*P* value = .02), and timely communication (*P* value = .03) post-follow-up compared with the control arm.

After simulation 3, TeamSTEPPS participants showed significantly greater improvement in shared knowledge (*P* value = .007) compared with the control arm, which persisted at follow-up (*P* value = .03). After simulation 4, TeamSTEPPS participants also showed significantly greater improvement in degree of leadership (*P* value = .01), deciding what should be done (*P* value = .003), assigning group members to particular tasks (*P* value = .004), frequent communication (*P* value = .02), timely communication (*P* value = .02), and problem solving (*P* value < .001) compared with the control arm (Table 3).

Subgroup analyses among EM residents

Among EM residents, unadjusted CTS mean leadership scores improved in both the LEADS and TeamSTEPPS arms, improving from the "average" to "good" range, while the active control group remained in the "average" range. This is consistent with the results from the full sample analysis.

Among EM residents at simulation 3, leadership performance was significantly improved for TeamSTEPPS participants (*P* value = .01) and improved for LEADS participants, though this did not reach statistical significance (*P* value = .06) compared with the control arm after adjusting for baseline. This differs from the results including all participants, where both LEADS and TeamSTEPPS participants showed significant improvement. Improvement in leadership performance was retained after follow-up (simulation 4) for both LEADS (*P* value = .03) and

TeamSTEPPS (*P* value=0.02) participants (Table 3). The results of the DLE assessment were similar for the primary and for the EM-resident only analysis.

Discussion

In this 3-arm, double blinded, randomized controlled trial, we found that that a 60-minute asynchronous leadership training course, reinforced by a brief simulation scenario with structured debriefing, has a significant and lasting impact on leadership behaviors in an acute care setting among EM and OB/GYN residents. We did not find any significant difference between the LEADS curriculum and the modified TeamSTEPPS curriculum. The incorporation of implicit bias training in the LEADS curriculum did not lead to a significant difference in leadership scores compared to TeamSTEPPS. This could be due to the highly scripted nature of the scenarios in the study, where the confederates were trained in how to respond to the leader. TeamSTEPPS could also be effective at assisting in overcoming implicit bias, or the LEADS curriculum may not have provided effective tools at addressing implicit bias. Further research is needed in this area.

To our knowledge, there are few randomized control trials evaluating the effectiveness of leadership education and training among residents. One previous study found that brief leadership training improved leadership communication and resuscitation performance among medical students in simulated cardiac arrest.[23] Another study conducted at a Level I trauma center found that a 4-hour simulation training curriculum improved leadership skills during video-recorded trauma resuscitations.[24] Ninety minutes of crew-resource-management training was also found to improve care during simulated cardiac arrest scenarios.[25] Our study adds to this literature by finding sustained impact of brief leadership training 3 to 6 months after the intervention, compared with an active control without specific leadership training. Together, these studies indicate that structured leadership training, with or without the addition of implicit bias training, is effective in improving leadership skills. Future research should evaluate the optimal format and duration of training to maximize the impact on trainees while using resources efficiently.

Though leadership in the acute care scenarios improved in general with both curricula, we found LEADS and TeamSTEPPS each affected leadership somewhat differently, measured by the components of the DLE. It is unknown which aspects of leadership are most important for patient care. Also, it is possible that there is variability in leadership learning needs between learners that are important to address. The subgroup analysis of EM residents found both curricula resulted in significant long-term improvement in leadership, though short-term improvement was only significant for TeamSTEPPS. We believe these results should be interpreted cautiously because of the small sample size of this sub-analysis.

The Accreditation Council for Graduate Medical Education emphasizes the importance of interpersonal and communication skills as a core competency among all specialties; however, there is no clear statement regarding leadership training and expectations.[26] Given the results of this and other studies, we recommend all residency programs have a structured leadership training curriculum that incorporates simulation. Given the variable resources and expertise available among U.S. residency programs, specialty societies or other academic organizations are ideally positioned to create and provide this content. Our study found that relatively brief training can have a significant impact on leadership performance, indicating that such training should be feasible for all residency programs.

Limitations

We encountered several specific challenges in this study that should be considered when planning future studies. First, this study required busy residents to consistently participate in repeated simulation scenarios spread across 2 academic years, leading to fewer than expected enrollments. Next, each scenario required embedded participants to be “directed” by the team leader. In pilot scenarios, we noted that the embedded participants could significantly confound the observed leadership, dependent on their own subject matter expertise or familiarity with simulation-based medical education. We carefully trained embedded participants and tightly scripted their responses in the scenarios as much as possible to limit this potential source of bias. The control group was an entirely separate site, made up exclusively of EM residents, and the subjects were not randomized to the control group. Our study had relatively large confidence intervals, suggesting that a significantly larger sample would be needed to detect the relatively small expected differences between leadership curricula in future studies. We may not have had sufficient power to detect differences between the 2 intervention groups. In addition, our leadership intervention was relatively brief in duration, and we did not verify how long study participants spent engaged in the curriculum. The impact of a more extended or intense curriculum is unknown. Future studies should investigate the “dose–response” relationship between leadership training and outcomes and/or evaluate the performance of the web-based curriculum alone. Also, future studies could evaluate and control for potential bias against female participants among the leadership assessors. We felt this was necessary to avoid unintended exposure to the training materials among the controls. All participants at this site were EM residents, which is a potential confounder.

Conclusions

We found that residents who participated in a brief structured leadership training intervention, with or without the addition of implicit bias training, had improved leadership skills that were maintained at 3-to 6-month follow-up.

References

1. Agency for Healthcare Quality and Research. Improving America's Hospitals: The Joint Commission's Annual Report on Quality and Safety 2007. https://psnet.ahrq.gov/issue/improving-americas-hospitals-joint-commissions-annual-report-quality-and-safety-2007. Accessed December 21, 2021.
2. Angood P. Cultivating physician leadership is necessary for lasting and meaningful change. Acad Med. 2015;90:1290.
3. Blumenthal DM, Bernard K, Bohnen J, Bohmer R. Addressing the leadership gap in medicine: Residents' need for systematic leadership development training. Acad Med. 2012;87:513–522.
4. Jardine D, Correa R, Schultz H, et al. The need for a leadership curriculum for residents. J Grad Med Educ. 2015;7:307–309.
5. Saravo B, Netzel J, Kiesewetter J. The need for strong clinical leaders: Transformational and transactional leadership as a framework for resident leadership training. PLoS ONE. 2017;12:e0183019.
6. Hartzell JD, Yu CE, Cohee BM, Nelson MR, Wilson RL. Moving beyond accidental leadership: A graduate medical education leadership curriculum needs assessment. Military Med. 2017;182:e1815–1822.
7. Frich JC, Brewster AL, Cherlin EJ, Bradley EH. Leadership development programs for physicians: A systematic review. J Gen Intern Med. 2015;30:656–674.
8. Stein LI. The doctor–nurse game. Arch Gen Psychiatry. 1967;16:699–703.

9. Stegman CB. The effects of social and environmental factors on informal teaching: Learning transactions between female nurses and female physicians [dissertation]. Ft. Lauderdale, FL: Nova Southeastern University; 1987.

10. Loudon R. Sex and medicine: Gender, power and authority in the medical profession. Fam Pract. 1999;16:322–322.

11. Haddad AM. The nurse/physician relationship and ethical decision making. AORN J. 1991;53:151–154, 156.

12. Bartels C, Goetz S, Ward E, Carnes M. Internal medicine residents' perceived ability to direct patient care: Impact of gender and experience. J Women's Health (Larchmt). 2008;17:1615–1621.

13. Hansen M, Schoonover A, Skarica B, Harrod T, Bahr N, Guise J-M. Implicit gender bias among US resident physicians. BMC Med Educ. 2019;19:396.

14. Agency for Healthcare Research and Quality. TeamSTEPPS. https://www.ahrq.gov/teamstepps/index.html. Accessed December 6, 2021.

15. Accreditation Council for Graduate Medical Education. ACGME Data Resource Book. https://www.acgme.org/About-Us/Publications-and-Resources/Graduate-Medical-Education-Data-Resource-Book. Accessed December 6, 2021.

16. Guise J-M, Deering SH, Kanki BG, et al. Validation of a tool to measure and promote clinical teamwork. Simulation in Healthcare. 2008;3:217–223.

17. Herzberg S, Hansen M, Schoonover A, et al. Association between measured teamwork and medical errors: an observational study of prehospital care in the USA. BMJ Open. 2019;9:e025314.

18. Parush A, Mastoras G, Bhandari A, et al. Can teamwork and situational awareness (SA) in ED resuscitations be improved with a technological cognitive aid? Design and a pilot study of a team situation display. J Biomed Informatics. 2017;76:154–161.

19. Innocenti F, Angeli E, Alesi A, Scorpiniti M, Pini R. Teamwork evaluation during emergency medicine residents' high-fidelity simulation. BMJ Simul Technol Enhanc Learn. 2016;2:12–18.

20. Fransen AF, de Boer L, Kienhorst D, Truijens SE, van Runnard Heimel PJ, Oei SG. Assessing teamwork performance in obstetrics: A systematic search and review of validated tools. Eur J Obstet Gynecol Reprod Biol. 2017;216:184–191.

21. Miller D, Crandall C, Washington C, McLaughlin S. Improving teamwork and communication in trauma care through in situ simulations. Acad Emerg Med. 2012;19:608–612.

22. Gilfoyle E, Koot DA, Annear JC, et al. Improved clinical performance and teamwork of pediatric interprofessional resuscitation teams with a simulation-based educational intervention. Pediatr Crit Care Med. 2017;18:e62–e69.

23. Hunziker S, Bühlmann C, Tschan F, et al. Brief leadership instructions improve cardiopulmonary resuscitation in a high-fidelity simulation: A randomized controlled trial. Crit Care Med. 2010;38:1086–1091.

24. Fernandez R, Rosenman ED, Olenick J, et al. Simulation-based team leadership training improves team leadership during actual trauma resuscitations: A randomized controlled Trial. Crit Care Med. 2020;48:73–82.

25. Fernandez Castelao E, Boos M, Ringer C, Eich C, Russo SG. Effect of CRM team leader training on team performance and leadership behavior in simulated cardiac arrest

scenarios: A prospective, randomized, controlled study. BMC Med Educ. 2015;15:116.

26. Accreditation Council for Graduate Medical Education. ACGME Common Program Requirements (Residency). July 1, 2020. https://www.acgme.org/Portals/0/PFAssets/ProgramRequirements/CPRResidency2020.pdf. Accessed December 6, 2021.

Figure Legends

Figure 1

Overview of study design and timeline of study interventions and data collection among EM and OB/GYN residents, from a multi-institutional, randomized, controlled study of leadership curricula with and without implicit bias training, 2015–2017. Participants were individually randomized to the LEADS or TeamSTEPPS hybrid online/simulation clinical leadership curricula and had evaluations of leadership measured by blinded video assessment of the CTS during high-fidelity simulation scenarios during the various phases of the study.

Abbreviations: EM, emergency medicine; OB/GYN, obstetrics–gynecology; LEADS, Leadership Education Advanced During Simulation; TeamSTEPPS, Team Strategies and Tools to Enhance Performance and Patient Safety; CTS, Clinical Teamwork Scale.

Figure 2

Leadership performance subscore of the CTS tool assessed by blinded video reviewers during the baseline, immediate post-intervention, and post follow-up (after 3-6 months) data collection periods, from a multi-institutional, randomized, controlled study of leadership curricula with and without implicit bias training. EM and OB/GYN residents were randomized to the TeamSTEPPS or LEADS clinical team leadership curricula or to an active control group.

Abbreviations: CTS, Clinical Teamwork Scale; EM, emergency medicine; TeamSTEPPS, Team Strategies and Tools to Enhance Performance and Patient Safety; LEADS, Leadership Education Advanced During Simulation.

Table 1

Baseline Characteristics of Participants by Intervention Arm Among EM and OB/GYN Residents, From a Multi-Institutional, Randomized, Controlled Study of Leadership Curricula With and Without Implicit Bias Training , 2015–2017

Demographic	LEADS, no. (%) (n = 42)	TeamSTEPPS, no. (%) (n = 48)	Control, no. (%) (n = 20)	Total, no. (%) (n = 110)	χ^2	*P* value
Sex					4.7	.095
Female	19 (45)	14 (29)	11 (55)	44 (40)		
Male	23 (55)	34 (71)	9 (45)	66 (60)		
PGY					5.11	.28
1	29 (69)	28 (58)	13 (65)	70 (64)		
2	9 (21)	12 (25)	7 (35)	28 (25)		
3	4 (10)	8 (7)	0 (0)	12 (11)		
Department					14.53	.001[a]
OB/GYN	19 (45)	22 (46)	0 (0)	41 (37)		
EM	23 (55)	26 (54)	20 (100)	69 (63)		

Abbreviations: OB/GYN, obstetrics and gynecology; EM, Emergency Medicine; LEADS, Leadership Education Advanced During Simulation; TeamSTEPPS, Team Strategies and Tools to Enhance Performance and Patient Safety; PGY, postgraduate year.

[a]Significant at $P < 0.05$.

Table 2

Results From Unadjusted Analyses of Mean Leadership Performance Scores on the Clinical Teamwork Scale, by Intervention Among Emergency Medicine and Obstetrics–Gynecology Residents, From a Multi-Institutional, Randomized, Controlled Study of Leadership Curricula With and Without Implicit Bias Training, 2015–2017

Intervention arm	Baseline, mean (SD)	Post-intervention, mean (SD)	Post-follow-up, mean (SD)
LEADS	5.88 (2.21)	7.86 (2.01)	6.95 (1.71)
TeamSTEPPS	5.1 (2.53)	7.77 (1.93)	7.29 (1.92)
Controls	5.25 (2.77)	6.25 (2.05)	5.7 (2.54)

Abbreviations: LEADS, Leadership Education Advanced During Simulation; TeamSTEPPS, Team Strategies and Tools to Enhance Performance and Patient Safety.

Table 3

Results From Linear Mixed-Effects Regression Models of Differences in Outcome Variables Immediately After the Intervention and at 3–6 Month Follow Up Among Emergency Medicine and Obstetrics–Gynecology Residents, From a Multi-Institutional, Randomized, Controlled Study of Leadership Curricula With and Without Implicit Bias Training, 2015–2017

	Post intervention simulation				Post follow-up simulation			
	LEADS		TeamSTEPPS		LEADS		TeamSTEPPS	
Component and outcome variable	B (95% CI)	*P* value	B (95% CI)	*P* value	B (95% CI)	*P* value	B (95% CI)	*P* value
Clinical Teamwork Scale								
Leadership performance	1.58 (0.05, 3.10)	.04[a]	1.53 (0.02, 3.05)	.05[a]	1.09 (0.11, 2.08)	.03[a]	1.63 (0.66, 2.59)	.001[a]
Detailed leadership assessment								
Degree of leadership[b]	0.65 (0.22, 1.1)	.003[a]	0.32 (-0.1, 0.74)	.14	0.46 (-0.01, 0.94)	.06	0.59 (0.13, 1.05)	.01[a]
Decided what should be done[c]	0.43 (-0.23, 1.08)	.2	0.09 (-0.57, 0.74)	.79	0.49 (0.09, 0.89)	.02[a]	0.6 (0.2, 0.99)	.003[a]
Assigned group members to particular tasks[c]	0.39 (-0.39, 1.17)	.33	-0.05 (-0.82, 0.73)	.91	0.19 (-0.25, 0.64)	.39	0.64 (0.20, 1.1)	.004[a]
Frequent communication[d]	0.30 (-0.09, 0.69)	.14	0.17 (-0.22, 0.56)	.34	0.40 (-0.03, 0.84)	.07	0.53 (0.10, 0.95)	.02[a]
Timely communication[d]	0.52 (-0.05, 1.1)	.08	0.33 (-0.23, 0.89)	.25	0.45 (0.03, 0.86)	.03[a]	0.49 (0.09, 0.89)	.02[a]
Accurate communication[d]	0.53 (-0.09, 1.2)	.09	0.24 (-0.38, 0.85)	.45	0.28 (-0.10, 0.66)	.15	0.37 (-0.005, 0.73)	.053
Problem-solving[d]	0.73 (0.12, 1.3)	.02[a]	0.53 (-0.07, 1.12)	.08	0.47 (0.03, 0.91)	.04[a]	0.79 (0.36, 1.22)	<.001[a]
Shared knowledge[d]	0.89 (0.38, 1.42)	.001[a]	0.69 (0.19, 1.2)	.007[a]	0.56 (0.07, 1.1)	.03[a]	0.55 (0.06, 1.03)	.03[a]
Mutual respect[d]	0.43 (-0.28,1.14)	.23	0.48 (-0.23,1.2)	.18	-0.14 (-0.56, 0.27)	.49	-0.25 (-0.65, 0.16)	.23

Abbreviations: LEADS, Leadership Education Advanced During Simulation; TeamSTEPPS, Team Strategies and Tools to Enhance Performance and Patient Safety; CI, confidence interval.

[a]Significant at *P* value < .05.

[b]From 1 = "Does not assume leadership" to 5 = "Firmly assumes leadership."

[c]From 1 = "Never" to 5 = "Always."

[d] From 1 = "Not at all" to 5 = "Completely."

Figure 1

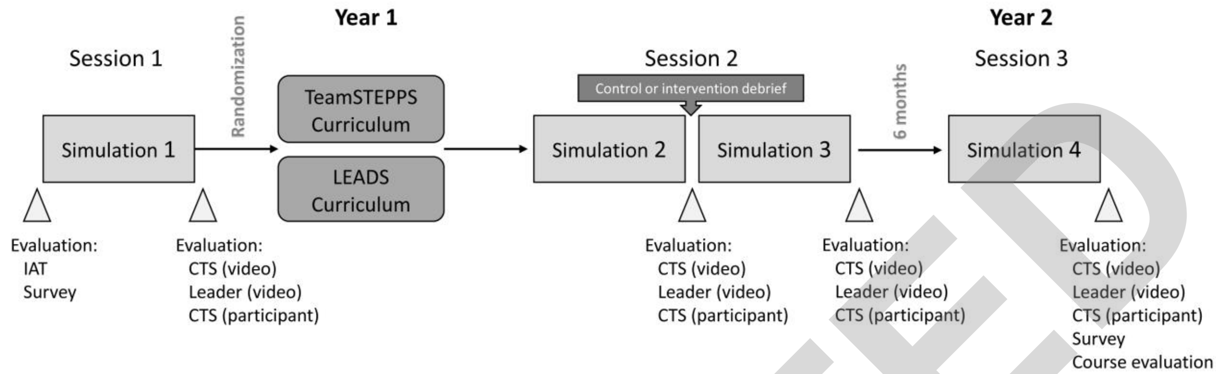
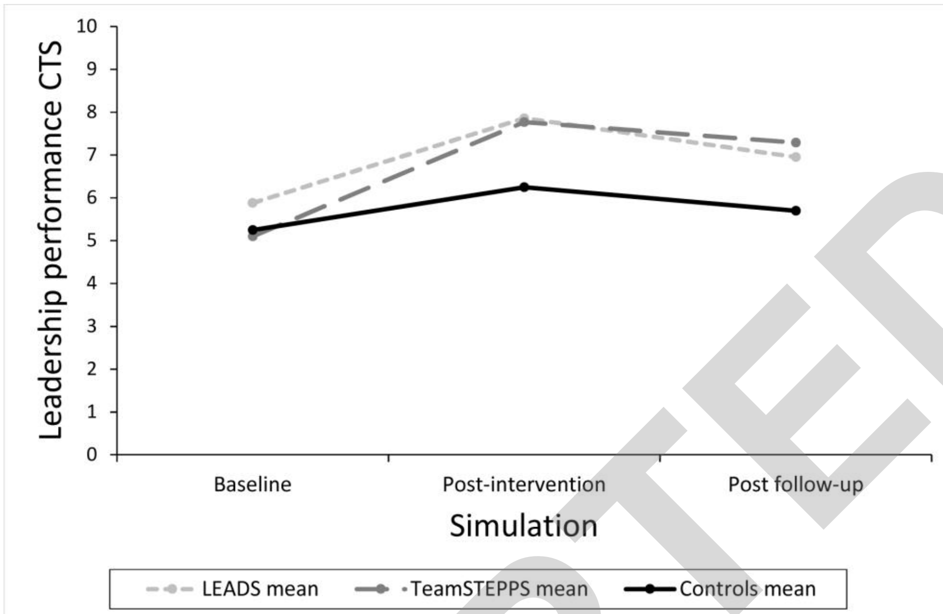

Year 1
Year 2
Session 1
Randomization
Session 2
Control or intervention debrief
6 months
Session 3
TeamSTEPPS Curriculum
LEADS Curriculum
Simulation 1
Simulation 2
Simulation 3
Simulation 4
Evaluation:
IAT
Survey
Evaluation:
CTS (video)
Leader (video)
CTS (participant)
Evaluation:
CTS (video)
Leader (video)
CTS (participant)
Evaluation:
CTS (video)
Leader (video)
CTS (participant)
Evaluation:
CTS (video)
Leader (video)
CTS (participant)
Survey
Course evaluation

Figure 2



10
9
8
7
6
5
4
3
2
1
0
Leadership performance CTS
Baseline
Post-intervention
Post follow-up
Simulation
LEADS mean
TeamSTEPPS mean
Controls mean