Simulation to Address Gender-based Differences in Leadership, Teamwork, and Safety

Principal Investigator: Jeanne-Marie Guise, MD, MPH

Study Team Members: Matthew Hansen, MD; Tabria Harrod, MPH; Shammarie Mathis; Barbara Skarica, MPH, Amanda Schoonover, MPH

Organization: Oregon Health and Science University

Project Period: 09/30/2014-09/29/2017, NCE 09/30/2018

Federal Project Officer: Kerm Henriksen, PhD

Acknowledgement of Agency Support: This study was funded by the National Institutes of Health and Agency for Healthcare Research and Quality

Grant Number: 5R18HS023457

2. **Structured Abstract**

Purpose: Cultural norms can influence leadership and team interactions. This study characterizes residents' experiences and educational needs regarding leadership by developing and testing gender-sensitive training.

Scope: We developed and administered a National Leadership Survey (NLS) to characterize residents' perceptions of leadership and whether sex/gender or other differences affect leadership and teamwork. We engaged national experts to develop leadership training and tested it through a multicenter RCT.

Methods: We surveyed residents across the US; assembled the country's leaders in simulation, leadership, safety, and resident training to develop two curricula (concise TeamSTEPPS and individualized LEADS); and conducted a multicenter RCT.

Results: We received 2,549 resident responses from 326 programs nationwide. The NLS showed a multitude of biases reported by residents that impacted the effectiveness of their leadership, regardless of specialty. In total, 130 residents participated in the RCT, which showed that any leadership training paired with a structured debrief improved leadership and teamwork skills during simulations compared with no intervention. The mean leadership composite score was significantly higher for participants of the LEADS curriculum compared with the control group (p-value<0.05); this was not true for TeamSTEPPS. Although women showed significant improvement in leadership performance post intervention for both curricula, men only showed significant improvement after LEADS.

Key Words: Leadership, Teamwork, OB/GYN, EM, Unconscious Bias, Simulation, Curriculum

3. LEADS FINAL PROGRESS REPORT

3.1 Purpose

The purpose of this study was to characterize EM and OB/GYN residents' perceptions of leadership and teamwork (Aim 1); to engage national leaders in simulation, leadership, teamwork, and medical training; to develop simulation-based expert model leadership training (Aim 2); and to conduct a multicenter, randomized, controlled trial to evaluate the effectiveness of tailored leadership training (Aim 3). We hypothesized that societal expectations of sex may inhibit the leadership and followership actions needed for effective clinical teamwork and that a normatively adapted approach to leadership and team training that accounts for societal biases and individual experiences would increase the effectiveness of leadership and clinical teamwork.

3.2 Scope

According to The Joint Commission, inadequate leadership is a major contributor to half of sentinel healthcare events.[1] A recent review found that leadership training is crucial to enhance team performance and that, although leadership is a learnable skill, it is not often taught in medical school or residency.[2] Responsibilities of leaders include directing and coordinating activities; assessing team performance; assigning roles and coordinating tasks; setting the tone for interactions; fostering the self-efficacy of team members; seeking and evaluating information; facilitating problem solving and error recognition; and facilitating feedback.[2,3] Teams that do not have instantly recognizable leadership are less effective and have poor teamwork; therefore, the ability of physicians to act as team leaders is pivotal to effective teamwork in the emergency setting.[4]

The role of sex in healthcare teams has been the subject of psychology research for decades.[5-8] Female residents tend to choose less assertive behaviors in clinical scenarios compared with male residents.[9] In a study of Emergency Medicine (EM), surgery, internal medicine, and obstetrics and gynecology (OB/GYN) residents and nurses, sex was identified by both nurses and physicians as playing a major role in communication issues and leadership effectiveness, with female physicians reporting that their decisions were challenged more frequently than those of male physicians.[10] Although the US has focused on the rights of women in the workforce, sex still plays a role in the communication of healthcare teams. Programs that address the differing needs of men and women as leaders could play an important role in improving patient safety.

Leadership and clinical decision-making skills are critical in the practice of safe medicine, particularly during emergencies, when failures and lapses can mean the difference between life and death in a matter of seconds. Cultural norms, beliefs, and assumptions inherent in the environment influence team dynamics and pose a challenge to leadership and teamwork. Studies suggest that one's sex may affect a person's confidence to assume leadership and perceived competence in clinical decision making. Conceptual models of leadership and teamwork consider these norms, beliefs, and assumptions as inputs that affect team outputs. In today's healthcare system, teams are transitory, and mixed-sex teams are commonplace. In this situation, when leaders may not know their teammates yet, they need to be effective in acute emergency situations. Having skills to navigate unconscious beliefs could improve residents' confidence in becoming leaders and improve their effectiveness in acting as leaders during emergencies. Although interventions to improve communication and teamwork have been studied in a variety of medical settings, there remains a need to refine and explore strategies to adapt training to norms that have been long recognized as barriers and/or challenges to effective communication, leadership, and teamwork.

Participants
Aim 1: Emergency Medicine (EM and Obstetrics & Gynecology (OB/GYN residents in teaching hospitals in the United States
Aim 2: National panel of experts in leadership, education, simulation, unconscious bias, research, social science, and residency training
Aim 3: EM and OB/GYN residents from Oregon Health and Science University (OHSU, University of Massachusetts (U Mass, University of Pennsylvania (U Penn, Indiana University (IA, and University of Arizona (U Az

3.3 Methods
Aim 1- National Leadership Survey
We developed a survey intended to assess the perceived confidence and importance of various leadership attributes among US OB/GYN and EM residents. Additionally, the study asked for stories of experiences and evaluated whether any individual characteristics (e.g., physical stature, gender, race, ethnicity, etc. were perceived to affect residents' effectiveness in leading health teams. The survey development involved several phases. We first met with leaders of national residency organizations for OB/GYN and EM from the Council on Resident Education in Obstetrics and Gynecology (CREOG) and the Council of Emergency Medicine Residency Directors (CORD) to coordinate the survey with other activities, understand accepted survey methods for these groups, and request their support in promoting the survey and the larger multicenter trial. There were differences in regulations between the two residencies. EM residencies would provide emails for all residents and allow direct survey methods, whereas OB/GYN residency organizations do not allow direct contact of residents. In order to standardize the approach and responses, we approached both OB/GYN and EM residents through residency directors (the approach required for OB/GYN).

The format of the National Leadership Survey (NLS) was adapted from the well-tested and highly successful International Center for Executive Leadership in Academics' Leadership Learning and Career Development Survey (LLCD). Additional items were added to the survey to capture tasks and performance that are considered critical in leadership of acute care teams.[11,12] The LLCD is conducted annually as part of the Executive Leaders in Academic Medicine (ELAM) program, which is designed to prepare senior academic faculty for leadership positions in academic medicine. The LLCD format asks respondents to evaluate the importance of a given skill alongside their personal confidence in conducting that skill. Additional items were added to the survey to capture tasks and performance that are considered critical in leading acute care teams.

We emailed a request to all OB/GYN and EM residency directors, asking them to distribute the National Leadership Survey to their residents between April 2015 and February 2016. Prior to our initial contact with the programs, we sent a CORD- and CREOG-endorsed pre-letter signed by CREOG and CORD leadership to all residencies, describing the survey intent. Within 3 to 7 days of the pre-letter, an initial email was sent to residency directors (165 EM and 229 OB/GYN) that described the project, contained a link to the electronic survey, and asked directors to forward the survey to their residents. A reminder email was sent 7 to 10 days after, and another reminder was sent 7 to 14 days after the first reminder. Reminder emails were sent to both program directors and coordinators an additional five times throughout Fall 2015/Winter 2016 before the survey was closed in February 2016. The National Leadership Survey was administered using a secure web link that took participants to a customized website created specifically for this project.

This study was carried out as a needs assessment for a leadership intervention trial. The survey was conducted among all residencies, but the trial was conducted among five specific clinical sites. Among these sites, there was additional focus on obtaining responses to the survey.

The survey started by collecting demographic information (e.g., sex, type of residency program, and level of training and leadership skills). Next, the survey included a seven-point Likert-type questions on perceived self-confidence in, and importance of, 17 specific leadership skills (with a score of 0 indicating “not important” or “not confident,” respectively, and a score of 7 indicating “extremely important” or “extremely confident,” respectively).

Analysis

The questions from the survey were grouped into domains for analytical purposes: core skills, adaptability, skill improvement, team player, debriefing, and leadership recognition. SAS was used for all statistical analyses (Version 9.4, SAS Institute, Inc., Cary, NC).

We first used descriptive statistics to summarize the mean responses to the importance and confidence questions for each sex. We then used univariable linear regression models to directly compare the differences in mean importance and confidence scores between sex. Next, we conducted multivariable linear regression analysis to assess the association between post-graduate training year on the perceived importance and confidence while controlling for sex. We conducted a subgroup analysis of participants in the clinical trial when the survey response rate was higher by adding interaction terms for being a trial site participant, sex, and PGY variables in the multivariable regression model.

We then analyzed the short answer responses using content analysis and grounded theory, a methodology that identifies emergent themes within a text to develop frameworks that specifically address the research problem. One analyst independently conducted a line-by-line analysis of the responses to identify overarching themes and subthemes. After a preliminary review of the data, a second analyst reviewed a sample of the dataset to ensure reliability and accuracy of the coding scheme. Next, the two analysts developed a set list of codes that were modified as the initial review progressed. Any discrepancies between coding were resolved by consensus reached through discussion. Unstructured narrative text from survey responses were imported into NVivo 11 software (QSR International, Cambridge, MA) for qualitative analysis.

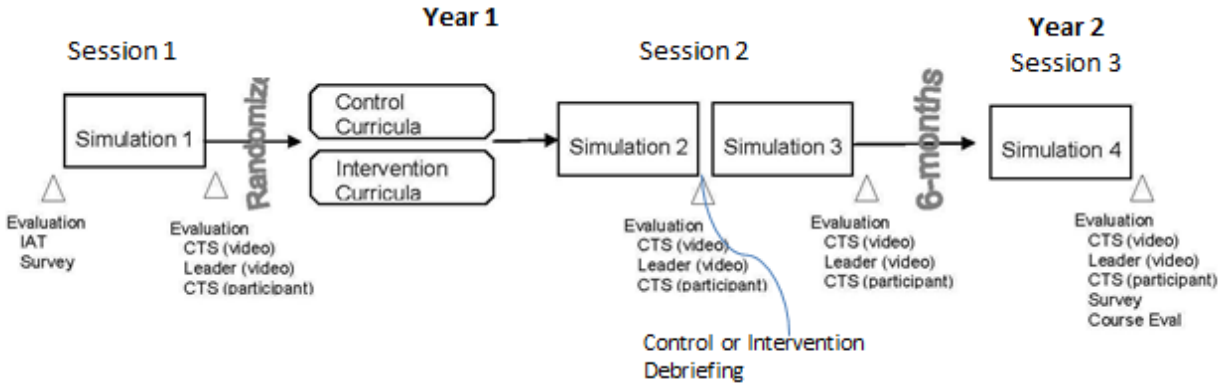
Aim 2- Curriculum Development

The Leadership Education Advanced During Simulation (LEADS) curriculum was developed in partnership with the Council on Resident Education in Obstetrics and Gynecology (CREOG) and the Council of Emergency Medicine Residency Directors (CORD). The objective of LEADS is to develop leadership skills in all physicians so they are prepared to effectively lead in the framework of an increasingly diverse and dynamic interprofessional work environment.

LEADS development was guided by a national panel of experts in leadership, education, simulation, unconscious bias, research, social science, and residency training. These experts reviewed results from the National Leadership Survey to help identify curricular content and develop learning objectives. Individual calls (in addition to group calls) were held with expert panel members to develop specific curricular sections, and the panel reviewed the curriculum to create the final version. We worked closely with CREOG and CORD representatives throughout the process, from needs assessment through implementation, disseminating messages, obtaining feedback, and coordinating with residency communities.

Aim 3- Randomized Controlled Trial

The overall process and evaluation plan is shown in the Figure below.



The grant proposed a head-to-head RCT of two curricula, TeamSTEPPS versus LEADS interventions. We brought campus research leaders together early in the grant process to evaluate the proposed work. The committee felt that it was extremely important to have an arm that did not have any additional curriculum in case the major effect was getting any teamwork training over none. Realizing that we had a lean budget and did not budget for additional sites, they recommended getting anyone who would be willing to conduct the sims alone, even if only at one site and one residency (e.g., only EM or OB/GYN). Therefore, we conducted a three-arm, multicenter RCT (stratified by sex) of simulation-based teamwork training using the TeamSTEPPS curriculum for arm 1, the tailored LEADS curriculum for arm 2, and no additional training (control) for arm 3 among EM and OB/GYN residents. All participating residents from OHSU, University of Massachusetts, University of Pennsylvania, Indiana University, and University of Arizona took part in a baseline simulation in which we assessed leadership behaviors to understand potential sex differences and allowed for measurement of change after the intervention. Initially the plan was for all nonsenior residents to participate (nonsenior only, because the trial would be going for a year and we needed the residents to stay throughout the study period for prospective evaluation). Discussions with residency representatives revealed that it would not be physically possible to get all eligible residents to complete the entire study. This was for several reasons: 1) family leave, 2) holidays, 3) off-service rotations, 4) space needs, and 5) amount of time required for teachers and simulation actors. Thus, the numbers differed in intended participation for the grant and prior to launch. The validated Clinical Teamwork Scale (CTS™)[13] and Detailed Leadership Evaluation (DLE) were used to objectively measure clinical teamwork and leadership behaviors of participants. The CTS™ contains 14 items, each ranging from 0 (unacceptable) to 10 (perfect). The DLE was adopted from the Leadership Behavior Description Questionnaire (LBDQ) and used a five-point Likert-scale questions to measure and assess leadership.[14]

The sites that were randomly assigned to either TeamSTEPPS or LEADS curricula included OHSU, University of Massachusetts, University of Pennsylvania, and Indiana University. After the initial baseline simulation, residents at the intervention sites were randomly assigned to one of two online curricula (TeamSTEPPS or LEADS) to begin their seven-part educational intervention for teamwork and leadership training. Upon logging onto the website, the computer randomly assigned them to either TeamSTEPPS or intervention (LEADS) curriculum and assigned them a unique ID. Site faculty leaders of the simulations and debriefings at each site were given a structured form to debrief on detailed sex-based leadership and teamwork skills in the tailored LEADS intervention group or general teamwork in the TeamSTEPPS intervention group.

Within 1 month after completion of the online modules, residents participated in two different simulations, the first designed to allow initial practice of newly learned skills (which was considered part of the curriculum, and the second for evaluation of skills learned. Six months later, all sites conducted an additional (4th simulation with each resident in order to evaluate for potential decay in leadership skills over time and the durability of the intervention effect. Each of the four simulations represented a novel clinical topic that both OB/GYN and EM residents would see in their practice. Simulations were videotaped to allow detailed evaluations from reviewers blinded to the intervention, study hypothesis, participant identity, residency program, and timing of simulation in the study sequence.

In addition to the four sites from which residents participated in the LEADS or TeamSTEPPS leadership curricula, the University of Arizona, served as a "true" control site for the study. The residents at this site were not exposed to either the LEADS or TeamSTEPPS curricula or tailored debriefs. They received the same four simulations that were used by the two intervention sites, and their simulation video was recorded for review. This provided a comparison to residents undergoing no intervention to clarify the magnitude of the effect.

Statistical Analysis

Data analysis and management were performed using SAS software, version 9.4, of the SAS System for Microsoft Windows (copyright © 2013 SAS Institute, Inc.). SAS and all other SAS Institute, Inc., product or service names are registered trademarks or trademarks of the SAS Institute, Inc., Cary, NC, USA. The main outcome of interest for the CTS™ was overall teamwork score (0-10), with 0 being unacceptable and 10 being perfect. The overall CTS™ score was left skewed and analyzed with nonparametric methods. The CTS™ scores for both simulation 1 (cardiac arrest) and simulation 3 (eclampsia) were described using median and range. A chi-squared test of independence was used to assess associations between sets of independent variables and between the independent variables and CTS™ score variables. Spearman correlation was used to determine correlations between overall CTS™ score and DLE variables of interest.

The difference in overall teamwork CTS™ scores was analyzed using the Wilcoxon rank-sum test for binary independent variables sex and department and the Kruskal-Wallis test for categorical independent variables (PGY, institution, and intervention arm). The Dwass-Steel-Critchlow-Flinger method was used for multiple comparisons of differences in CTS™ score.

We tested for differences in CTS™ score between the intervention arms and within independent variables for both simulation 1 and simulation 3. Additionally, the change in overall CTS™ score between simulation 3 (post intervention) and simulation 1 (pre intervention) was calculated, and the Wilcoxon sign-rank test was used to test the null hypothesis that the median change in CTS™ score is equal to zero within each intervention arm. To test for sex differences, the Wilcoxon sign-rank test was used to test if the change in CTS™ score was significantly different from zero within each intervention arm and stratified by sex. Differences in average leadership performance CTS™ score between simulation 1 (baseline) and simulation 3 (post intervention) were analyzed using the parametric equivalent: paired t-tests by treatment and stratified by sex to test for differences between men and women. Additionally, t-tests for significance in difference in change of leadership performance CTS™ score between men and women were performed.

The DLE was assessed for association to the intervention arm using Fisher's exact test for small cell count due to small sample size. When a p-value was not able to be calculated, Monte Carlo

simulation with 50,000 samples was used to estimate the Fisher's exact significance. A composite score was created for ease of interpretability of the DLE variables to reflect leadership skills of the participants using nine of the 14 DLE variables that were deemed most indicative of leadership. The mean was calculated and used as a composite measure for leadership skill. Differences in mean composite score were analyzed using analysis of variance (ANOVA) tests. T-tests were used to compare composite scores between men and women within each treatment group. Additionally, the intervention arm that had the highest proportion of individuals in the highest rank for the nine selected DLE variables was calculated. All p-values are two tailed unless noted otherwise, using a significance level of 0.05.

3.4 Results

Aim 1- National Leadership Survey

All training-level residents from 172 EM and 229 OB/GYN residencies across the country received the leadership survey.

Quantitative Analysis

We received at least one response from residents in 326 programs (81%), with 2,549 residents responding out of an estimated 14,234 eligible residents (16.4%). A total of 2,214 residents (15.5%) completed the portion of the survey that included Likert questions about importance and perceived self-confidence in leadership skills (Table 1). In the subset of five clinical sites participating in the RCT, 100.0% (n=287) completed the survey through the leadership skill Likert questions, with 92.0% (n=264) completing all Likert scale questions. The responses were generally balanced across the two specialties. Generally, the OB/GYN programs in our sample consisted of 4-year programs, and the EM programs consisted of 3-year programs. The study participants from the five RCT sites were generally similar to the overall group of survey respondents. For both female and male respondents, the domain of questions that was felt to be most important was the "core leadership skills," and the least important was "leader recognition." The highest scored individual "importance" question for female residents was "maintain good working relationships throughout the hierarchy of staff," with a mean score of 6.57. For male residents, the highest scored individual "importance" question was "being a highly effective team leader," with a mean score of 6.48. Female residents rated all categories (adaptability, leadership development, team player, and debriefing) as relatively more important compared with their male colleagues, except for leader recognition, whereas men rated "being a team leader whom everyone recognizes as in charge" as more important than women did. The leadership importance scores were statistically significantly higher among PGY1 residents compared with other training years for all questions but did not have any significant difference by specialty.

Overall, residents' scores for the confidence questions were lower than their perceived importance. The highest rated domain of confidence was the "personal growth" domain, which included questions on soliciting feedback from team members and managing personal stress. Male residents expressed higher mean confidence scores for all individual questions except for "showing compassion for team members," which was the highest rated individual confidence question for women, with a mean score of 5.92. For male residents, the highest mean confidence score was 5.73 for both "showing compassion for team members" and "maintaining good working relationships throughout the hierarchy of the staff."

Table 1. Univariable Analysis of Importance and Confidence Scores on Individual Leadership Questions by Sex, Respondents from All Sites (n=2,214)

Leadership Skills	**Importance Scores by Sex** Mean (SD)		**Confidence Scores by Sex** Mean (SD)	
	Female	Male	Female	Male
Core Skills	***6.43 (.62)***	***6.35 (.60)***	**4.88 (0.92)**	**5.10 (0.88)**
Being a highly effective team leader	6.54 (.72)	6.48 (.74)	4.74 (1.05)	5.08 (1.03)***
Prioritize and coordinate tasks	6.40 (.78)	6.30 (.78)**	5.00 (1.12)	5.04 (1.07)
Speak aloud new plans	6.51 (.76)	6.43 (.75)*	4.87 (1.31)	5.17 (1.21)***
Monitor environment and situational awareness	6.26 (.86)	6.18 (.89)*	4.93 (1.16)	5.13 (1.13)***
Adaptability	***6.15 (.86)***	***6.00 (.90)***	**4.61 (1.08)**	**4.90 (1.07)**
Adapt leadership styles to the needs of particular situations	6.26 (.87)	6.15 (.91)**	4.70 (1.12)	5.00 (1.12)***
Adapt leadership styles to the needs of particular team members	6.04 (.98)	5.86 (1.04)***	4.51 (1.16)	4.81 (1.16)***
Personal Growth	***6.09 (.81)***	***5.93 (.84)***	**4.54 (1.07)**	**4.78 (1.04)**
Solicit feedback from supervisors regarding leadership	5.95 (1.00)	5.78 (1.12)***	4.52 (1.35)	4.74 (1.28)***
Solicit feedback from team members regarding leadership approach or style	5.89 (1.04)	5.72 (1.11)***	4.48 (1.28)	4.65 (1.26)**
Manage personal work-related stress	6.43 (.83)	6.30 (.88)***	4.63 (1.31)	4.95 (1.34)***
Team Player	***6.22 (.90)***	***6.12 (.67)***	**5.32 (0.82)**	**5.41 (0.80)**
When appropriate, use collaborative decision-making processes	6.42 (.80)	6.27 (.84)***	5.26 (1.07)	5.48 (1.04)***
Create consensus and compromise in situations when there is conflict	6.36 (.80)	6.15 (.90)***	4.94 (1.13)	5.22 (1.12)***
Maintain good working relationships throughout the hierarchy of staff	6.57 (.73)	6.45 (.82)**	5.51 (1.14)	5.73 (1.06)***
Share leadership among team members	5.85 (1.05)	5.54 (1.15)***	4.97 (1.13)	4.99 (1.12)
Show compassion for team members	6.46 (.79)	6.18 (.95)***	5.92 (1.00)	5.73 (1.10)***
Encourage team members to speak up	6.25 (.86)	6.13 (.89)**	5.31 (1.14)	5.33 (1.14)
Debriefing	***6.22 (.90)***	***5.96 (1.01)***	**4.71 (1.23)**	**4.71 (1.24)**
Incorporate reflection and debriefing into evaluation of cases	6.22 (.90)	5.96 (1.01)***	4.71 (1.23)	4.71 (1.24)
Recognition	***5.79 (1.11)***	***5.92 (1.07)***	**4.47 (1.19)**	**4.83 (1.16)**
Being a team leader whom everyone recognizes as in charge	5.79 (1.11)	5.92 (1.07)**	4.47 (1.19)	4.83 (1.16)***

*p-value<0.05, **p-value≤0.01, ***p-value≤0.001

Confidence in possessing leadership skills showed statistically significant increases (p>0.01) with increasing year in residency across all skills assessed in the survey across all domains. Between PGY1 and PGY3, confidence increased the most in speaking aloud new plans (mean difference: 0.88), being a team leader whom everyone recognizes as in charge (mean difference: 0.86), and prioritizing and coordinating tasks (mean difference: 0.78).

There were no differences by specialty to the confidence questions.

In multivariable analysis, mean confidence scores in being a highly effective team leader remained higher among men than women (5.08, SD 1.031 vs. 4.74, SD 1.049; p-value<0.0001) and higher among more senior residents (PGY1 4.45, SD 1.102; PGY3 5.2, SD .9555; PGY4 5.42, SD .88; p-value<0.0001. Confidence scores in prioritizing and coordinating tasks were similar across sexes (5.00, SD 1.119; vs. 5.04, SD 1.069; p-value<0.0001 but higher among more senior residents (PGY1 4.57, SD 1.16; PGY3 5.35, SD .957; PGY4 5.63, SD .895; p-value<0.0001). **Confidence scores in speaking aloud new plans were higher among men (5.17, SD 1.206) than women (4.87, SD 1.306; p-value<0.0001) and again rose with seniority (PGY1 4.49, SD 1.334; PGY3 5.37, SD 1.097; PGY4 5.59, SD 1.067; p-value<0.0001).** Confidence in encouraging members to speak up did not statistically differ between sexes but did increase with seniority (PGY1 5.2, SD 1.219; PGY3 5.47, SD 1.092; PGY4 5.56, SD .978; p-value<0.0001). Confidence in monitoring environment and situational awareness was statistically higher among men (5.13, SD 1.126; vs. 4.93, SD 1.161; p-value<0.0001) and among more senior residents (PGY1 4.72, SD 1.264; PGY3 5.23, SD 1.028; PGY4 5.36, SD 1.014; p-value<0.0001). In our subgroup analysis of sites included in the clinical trial with a higher survey response rate, the interaction terms, including being a clinical trial site, were not significant, indicating that there was no effect modification of the results based on clinical trial site status.

Qualitative Analysis

The most common reported factors to influence perceived leadership effectiveness were experience, age, and sex (Table 2). Residents reported biases that influenced how they were perceived as leaders. Physical biases that imparted a positive influence on leadership were appearing older and being male, white, tall, loud, and English proficient. **Physical biases that imparted a negative influence on leadership were appearing young; being female, non-white, short/petite, or overweight; speaking with an accent; having a first language other than English; and speaking a language that did not match that of the patient's.**

When residents were asked to report the most common situation in which poor leadership impacted patient care, communication problems were the most consistently reported theme followed by ambiguity in identifying the leader, resistance to female leadership, and lack of personal confidence in leadership. Among communication problems, both male and female residents reported that failing to speak with authority and ambiguous leadership were problems they commonly observed. They reported that this resulted in "chaos" being present while attempting to care for patients, which caused disruptions and disorganization, leading to poor outcomes.

Age/Experience: Residents were also asked how sex/gender, race/ethnicity, language, age, and other personal characteristics impacted how others responded to them as team leaders. For this question, residents indicated that age and experience were the most influential personal characteristics on leadership. Age and experience are obviously correlated, and residents reported both younger age and decreased experience had a negative impact on their ability to lead. This included noting that team members and staff were less likely to follow their orders, checking in with a more experienced provider, as well as residents themselves having less confidence in their leadership due to lack of lived experience as physicians.

Table 2: Factors Believed to Influence Leadership	
Experience and Age	*"I personally feel that any sort of leadership I take I am undermined by my lack of experience and age. People constantly ask me if I am old enough to be a doctor. My father even says that he only wants physicians who have some wrinkles and grey hair."*
Sex	*Female Resident: "I have had instances when female nurses disregarded me because I am a young female physician. They do not take me seriously and I have to work really hard to gain their respect to prove to them that I think they are a valuable part of the team. I see the female nurses give respect to my male colleagues easier."* *Male Resident: "I honestly think being a male in OB/GYN is helpful; the nurses tend to challenge the female residents more than me. The older nurses tend to take on a motherly role and are more willing to help me, I think, than my female co-residents."*
Race/Ethnicity	*"Yes, as an African American physician, your decision-making processes are questioned more by ancillary staff initially; however, over time when the rapport is built, they generally tend to trust your judgment more. It's more difficult to earn respect as a female African American than my white male counterparts."* *"I am Caucasian and unfortunately this has benefitted me. There is a large degree of racism (subtle and overt) in my workplace. Though I benefit from it, the fact that it exists is upsetting and negative for me."*
Language	*"I don't think it plays much of a role. I speak Spanish fluently, which is a huge asset in patient care in my area, and many times I'm one of the few people in the room who can communicate fluently with our Spanish-speaking patients."* *"Me being able to speak the same language as the patient better improves patient care because we can respond more appropriately to their needs."*
Physical Appearance	*"I am a tall, white, male. I am not naive enough to think this doesn't affect how people perceive me. And for the most part this has been in a positive manner. I have frequently seen my female colleagues dismissed as nurses and non-white colleagues questioned unnecessarily on their decisions.* *"I'm overweight. I notice a definitive different in the treatment of me versus other residents who are not; they are listened to more and seem more respected."*

Sex/Gender: After age/experience, the next most commonly reported theme that impacted how team members responded to the leader was sex. Female residents commonly reported experiencing difficulties in leadership, with team members refusing to carry out orders or circumventing their leadership by finding other staff members: "*I think gender definitely influences how team members respond to me. Being a young female, support staff do not take me as seriously and it takes time for them to respect me and I think that mostly has to do with my gender*." Some female residents reported experiencing "toxic" female-female interactions, primarily with nursing staff as well as needing to act extra assertive to be effective while being labeled as "bitchy" as a result: "*As a woman, I have to learn to use my big girl voice in an emergency situation. I feel like I need to become more like a man - in my posture, my voice, my words - in order to be taken seriously."* Conversely, male residents with similar behavior did not experience the same negative labels. In addition, female residents reported that they were frequently confused for or addressed as nurses, students, or other ancillary staff by patients, while males were recognized and addressed as doctors regardless of training or experience level: *"People often assume I am a nurse because I am a female. This has happened with other members of the medical team as well as with patients, sometimes after I have introduced myself as 'Doctor.' I need people to know my place in the team for the team to operate effectively."* These experiences were more frequently reported among female EM residents compared with OB/GYN residents, with some female OB/GYN residents indicating that female sex was not a

disadvantage in their field because the majority of their staff and all of their patients were female. Further, female residents reported that not being recognized as leaders adversely affected team dynamics. We noted the majority of comments from men regarding sex and leadership related to the positive effect of male sex on perceived leadership efficacy.

Physical Appearance: Next, residents reported that physical appearance had an impact on their leadership. Male residents reported that being tall and white had a positive impact on others' perceptions of them as a team leader: *"I'm a tall confident dude. People called me doctor when I was a first-year medical student and wanted my opinion, I have the look, and I look people in the eye. I'm intense and passionate about what I do; people see that. I also have a physical presence to match my attitude. It's every day of my life, it's how people look at me and how they respond or listen when I have something to say. When I spoke up in a resuscitation, people got quiet and looked at me, even though the attendings knew I was only an intern."* Conversely, female residents reported that being short, appearing young, or being from a race that they perceived made them appear young were disadvantages and resulted in team members not perceiving them as leaders. We noted the majority of comments from male residents regarding physical appearance related to the positive effects of certain aspects of their appearance on perceived leadership efficacy.

Race/Ethnicity: Residents also experienced leadership challenges based on race and ethnicity, though the vast majority of the sample identified themselves as white. Most experiences related to non-white residents, especially non-white female residents, who reported they were not recognized as physicians or as leaders. Some residents described instances in which they were confused for ancillary staff and often found it difficult to gain respect from staff members and/or patients: *"As a black woman I think it can feel as though you have to prove yourself more. It is noticeable in the 'looks' that you get from other team members and families when they know you are the decision maker."* Most of the respondents in this category identified themselves as Asian and African American/Black. Asian residents indicated a sense of stereotypes prevailing, in which Asians are viewed as not dominant. Finally, one resident described not feeling valued as a provider due to being African American. Race/ethnicity seemed to interact with sex with many quotes related to negative experiences of non-white female residents.

Aim 2- Curriculum Development

Needs Assessment

The LEADS curriculum was designed around residents' needs as identified through the National Leadership Survey (results above). The overarching goal of the LEADS curriculum is to help residents develop leadership strategies and personal skills that enable them to effectively lead diverse and dynamic healthcare teams in their future practice.

Development

The curriculum has well-defined goals and objectives covering knowledge, attitudes, behaviors, and skills. We used adult learning theory, which asserts that adults need to feel activated and internally motivated to learn, that they bring life experiences to learning, and that adults favor practical applications in learning. The curriculum included several specific leadership tools that were designed to be deployed in clinical practice. The leadership theories that provided the foundation for LEADS were taken from patient safety (TeamSTEPPS) and academic leadership (ELAM), focusing on knowledge, skills, and professional development/character. We instructed residents to practice with the content and tools they developed in the simulation scenarios that

followed the online curriculum. One of the innovative aspects of our curriculum was the development of an interactive platform that allows residents to integrate their experiences and needs with the evidence to build a personalized toolkit.

LEADS was developed as seven web-based modules, each lasting 10 minutes or less, that were intended maximize resident convenience and leverage interactive technologies that enable residents to build their own personalized leadership toolkit – something they can adjust throughout their careers. Residents were able to access modules through a personal portal, where they could enter and exit the curriculum at their own pace and across computer and mobile platforms. The modules provided educational content as well as interactive opportunities. The 10-minute maximum module duration was chosen in accordance with best practices thought to promote maximal effectiveness in online learning.[15] LEADS also provides simulation scenarios and faculty debriefing guides that provide opportunities for residents to practice their skills and participate in reflection and feedback. These simulation scenarios and structured debriefing guides were developed as supplementary materials to allow residency programs to offer residents opportunities to practice their skills and participate in faculty-guided reflection and feedback. The LEADS web-based curriculum can be used alone or combined with these simulations.

Quantitative Data

After the final module of both curricula, four Likert questions were included with a five-point scale, from extremely unlikely to extremely likely, assessing how likely the participants were to use this curriculum again. The four questions were 1. How likely is this curriculum to be helpful to you compared to previous leadership training you have participated in? 2. How likely are you to change your leadership practice and/or style based on this curriculum? 3. How likely are you to refer to this curriculum to reinforce these skills in the future? 4. How likely are you to recommend this curriculum to other residents in your field who have not taken it yet? Unfortunately, due to a computer programming issue, data from these Likert scale questions were not captured, so we are unable to include these data in an analysis and comparison of the two curricula. However, although these Likert responses were lost, the majority of participants in both curricula provided narrative responses about what they liked and disliked about each curriculum (Tables 3 and 4).

Qualitative Data

After the final module of both curricula, two short answer questions were included to assess what the participants liked and disliked about the curricula.

In total, 83% (n=45) of residents in the LEADS arm and 92% (n=46) of residents in the TeamSTEPPS arm completed the short answer section of their respective curricula. Overall, participants in LEADS offered more complimentary comments about the curriculum. When asked what they liked most about the curriculum, “Videos” were identified as a favorite among most LEADS participants, whereas TeamSTEPPS participants identified “Tips & Tools – Communication Tools.” LEADS participants described the videos as being effective in demonstrating awareness and appropriate body posture while in leadership roles. Conversely, TeamSTEPPS participants said they most enjoyed the SBAR and CUS models presented. In addition to “Videos,” LEADS participants said they enjoyed “Reviewing Concepts,” in which they were presented with the opportunity to review or reinforce concepts learned in prior training (e.g., closed loop communication, leadership principles, etc.). Among TeamSTEPPS

participants, “Videos” were identified as another favorite attribute from the curriculum, as they “made points clear” and the content itself was found to be particularly helpful (Table 3).

Table 3. What Residents Enjoyed About the Curriculum

Topic	LEADS Residents	TeamSTEPPS Residents
Videos	#1 Topic for LEADS residents • Videos demonstrating awareness, posture • Great to illustrate point	#2 Topic for TeamSTEPPS residents • Video examples made points clear • Videos were helpful
Reviewing Concepts	#2 Topic for LEADS residents • Reviewing/reinforcing concepts learned in prior training • Refreshing ideas of closed-loop comm., leadership principles, unconscious bias	#3 Topic for TeamSTEPPS residents • Shared mental model • Overview of leadership skills • Characteristics of high-performing teams
Tips & Tools	#3 Topic for LEADS residents • Methods for improving leadership skills • Realistic examples of good and bad leadership • Self-reflection practices • Recognizing triggers	#1 Topic for TeamSTEPPS residents • SBAR • CUS model
Examples	#4 Topic for LEADS residents • Specific examples were helpful • Examples of good and bad leadership	#4 Topic for TeamSTEPPS residents • Examples given
Phrases	#5 Topic for LEADS residents • Phrases to help with leadership • Phrases for good communication	• Topic not mentioned by TeamSTEPPS residents
Format	**Interactive Modules:** Tied for #4 topic for LEADS residents • Maintained engagement • Forcing interaction with material	**Illustrations:** #5 Topic for TeamSTEPPS residents • Graphics • Diagrams

Participants from both arms were also asked to describe their least favorite attribute about their respective curriculums (Table 4).

Table 4. What Residents Disliked About the Curriculum

Topic	LEADS Residents	TeamSTEPPS Residents
Exercises	#1 Topic for LEADS residents • Drag & drop exercises • Relevancy of material • Redundancy of tasks • Too many fill-ins	• Topic not identified by TeamSTEPPS residents
Videos	#2 Topic for LEADS residents • Videos	#1 Topic for TeamSTEPPS residents • Length of videos – too long • Video of teamwork failure – explanation didn’t make sense • Outdated videos • Videos do not work on tablet interface
Concepts Presented	#3 Topic for LEADS residents • Material was superficial, would be nice to do a more in-depth review	#2 Topic for TeamSTEPPS residents • Introduction describing why teamwork is important • Too much information without a clear takeaway • Information presented was redundant • Seemed redundant –concepts previously learned (SBAR)
Format	Tied for #2 Topic for LEADS residents **Length:** • Hard to complete in one setting • Too long • Information could have been summarized in fewer modules	#3 Topic for TeamSTEPPS residents **Audio:** • Audio and video did not sync up sometimes • Outdated audio • Verbose audio **Visual Appearance:** • Outdated visual style • Visual display on screen is too small • Format of module is “clunky”

Most LEADS participants said “Exercises” were among their least favorite, specifically identifying the “Drag & Drop” exercises. Many TeamSTEPPS participants identified “Videos” as their least favorite attribute. Videos were described as being too long, being outdated, or not making sense of the material presented. Along with “Videos,” TeamSTEPPS participants identified “Concepts Presented” as another least favorite curriculum attribute. Responses indicated that participants found the information presented was redundant, as most of the concepts presented were learned in prior training (e.g., SBAR).

Aim 3- Randomized Controlled Trial

At the start of the study, 155 residents agreed to participate and completed simulation 1. Of these participants, 138 (89% of the original 155) continued with the intervention portion of the study and were randomly assigned into either the TeamSTEPPS (n=58), LEADS (n=59), or control (n=21). Twenty-four individuals were not included in analysis due to scheduling conflicts, withdrawal, or loss of video or audio function. Only individuals with data for both simulation 1 and simulation 3 were included in analysis (n=114), as seen in Table 5 below.

Overall, 50 people went through the TeamSTEPPS curriculum, 53 went through LEADS, and 20 people were in the control and did not receive any intervention. However, due to equipment malfunction, nine videos from simulation 3 were not usable, all of which were people from OHSU, EM, and the LEADS arm. Thus, 44 individuals were used for video analysis for LEADS. Just over 90% of OB/GYN participants were women, compared with 39.8% of EM participants. The difference in proportion of women in the two departments is, as expected, statistically significant, with women more likely to be in OB/GYN than in EM (chi-square=36.5, p-value<0.0001). This is expected, given the proportion of women that comprised OB/GYN and EM residents and fellows nationally in 2017 (82.9% and 35.5%, respectively).

The characteristics of participants within each intervention arm were explored (Table 5).

Table 5. Baseline Characteristics by Treatment Arm of Participants Included in Video Analysis

	TeamSTEPPS *(n = 50)*		**LEADS** *(n = 44)*		**Control** *(n = 20)*		**Total** *(n = 114)*	
	n	%	n	%	n	%	n	%
Sex								
Female	36	72.0	24	54.6	9	45.0	69	60.5
Male	14	28.0	20	45.5	11	55.0	45	39.5
Institution								
OHSU	10	20.0	10	22.7	0	0.0	20	17.5
Penn	11	22.0	10	22.7	0	0.0	21	18.4
U Mass	18	36.0	14	31.8	0	0.0	32	28.1
IU	11	22.0	10	22.7	0	0.0	21	18.4
AZ	0	0.0	0	0.0	20	100.0	20	17.5
PGY								
1	30	60.0	29	65.9	13	65.0	72	63.2
2	12	24.0	10	22.7	7	35.0	29	25.4
3	8	16.0	5	11.4	0	0.0	13	11.4
Department								
OB/GYN	24	48.0	21	47.7	0	0.0	45	39.5
EM	26	52.0	23	52.3	20	100.0	69	60.5

There was no significant difference in the proportion of women in the three intervention arms (chi-square=2.11, p-value=0.35. Additionally, there was no significant difference in PGY in the three treatment arms (chi-square=5.35, p-value=0.25. As expected, there was a significant difference in the proportion of EM residents among the treatment arms, with EM participants most likely to be in the control arm (chi-square=15.8, p-value=0.0004, and significant differences in the proportion of participants from each institution in the three treatment arms (chi-square=147.1, p-value<0.0001. This is expected, given that all the controls were EM residents who were from the University of Arizona.

Clinical Teamwork Scale (CTS™:

All CTS™ variables for both simulations were significantly left skewed according to the Kolmogorov-Smirnov test for normality, apart from the binary target fixation variable. Thus, nonparametric methods were used for analyses on raw CTS™ variables. However, the created Change in Leadership Performance variable was normally distributed (p-value=0.09, and parametric equivalents were used.

Overall Teamwork:

Simulation 1 was the pre-intervention baseline scenario. The overall CTS™ score median was 5 for TeamSTEPPS, 6.5 for LEADS, and 6 for controls. There was not enough evidence to demonstrate a significant difference among these medians (chi-square=2.1, p-value=0.35. There was a significant difference in baseline CTS™ scores between men and women (Z=2.4, p-value=0.007, with the male median of 7 being significantly higher than the female median of 5.

Simulation 3 was the post-curriculum intervention scenario. The overall CTS™ score median was 8 for both TeamSTEPPS and LEADS, whereas the control group had a median CTS™ score of 7. According to the Kruskal-Wallis test, at least one of the medians was significantly different. Upon further investigation using the DSCF multiple-comparison test, we found that the median CTS™ score for LEADS participants was significantly higher than that of control participants (DSCF=3.3, p-value=0.049). There was insufficient evidence to show significant differences in overall CTS™ score between men and women, EM and OB/GYN residents, and PGY.

The median change in overall CTS™ score was +2 for both TeamSTEPPS and LEADS and was +1 for controls. According to the Wilcoxon sign-rank test, the changes in CTS™ score for both TeamSTEPPS (p-value<0.0001) and LEADS (p-value=0.0007) groups were significantly different from zero, whereas the change in CTS™ score for controls was not found to be significantly different from zero (p-value=0.08). This shows that any curriculum is better than no curriculum in terms of increasing overall teamwork according to the change in overall CTS™ score between simulation 1 and simulation 3. For all participants, men had significantly less change in the CTS™ score than women (Z=-1.66, p-value=0.05). When stratified by treatment group, we found that men had significantly less improvement in CTS™ score compared with women in the TeamSTEPPS arm (Z=-2.75, p-value=0.003); however, this finding did not persist in the LEADS or control arms. The change in CTS™ score was significantly greater for EM residents compared with OB/GYN residents (Z=2.05, p-value=0.02).

Sex Differences in Overall Teamwork:

Difference in CTS™ score between men and women within each intervention arm were assessed for simulation 1 and simulation 3 individually as well as for the change in CTS™ score (Table 6). For simulation 1, men had significantly higher overall CTS™ scores compared with women

within the TeamSTEPPS arm (Z=3.65, p-value<0.001). This difference was not found within the LEADS or control arms. There was insufficient evidence to show significant differences in overall CTS™ score between men and women for any of the intervention arms for simulation 3. Within the TeamSTEPPS arm, the median change in CTS™ score was not significantly different from zero for men (p-value=0.88); however, the difference was significant for women (p-value<0.0001). This can be explained by the fact that men in TeamSTEPPS at baseline (simulation 1) had significantly higher scores than women; thus, women had greater room for improvement. Within the LEADS arm, median change in CTS™ score for women was marginally significantly different from zero (p-value=0.055) and the median change for men was significantly different from zero (p-value=0.003).

Table 6. Difference in Median Change in Overall Teamwork CTS Score Within Each Intervention Arm Stratified by Sex

	Male			Female		
	median	S	p-value	median	S	p-value
Treatment						
TeamSTEPPS	0	2	0.88	3	194.5	<0.0001*
LEADS	2	65	0.003*	2	54.5	0.05*
Control	1	10.5	0.23	0	5.5	0.25

*significant at α=0.05

Last, for the control arm, there was not enough evidence to show a significant change in CTS™ score between simulation 1 and simulation 3. Again, these results would suggest that any intervention improves clinical teamwork skills as opposed to no intervention. However, these improvements are seemingly driven by men of the TeamSTEPPS arm and by both men and women of the LEADS arm.

Leader Performance:

Change in leadership performance between simulation 1 and simulation 3 is significant for both TeamSTEPPS (p-value<0.0001) and LEADS (p-value<0.0001. On average, individuals in the TeamSTEPPS intervention arm improved by 2.8 points, whereas individuals from the LEADS arm improved by 2.0 points. The change in leadership performance was not significantly different from zero for the control group (p-value=0.11.

Sex Differences in Leader Performance:

Women showed significant improvement in leadership performance between simulation 1 and simulation 3 (t=6.9, p-value<0.0001) (Table 7). On average, women improved by 2.7 points. Men also showed significant improvement (p-value=0.0009); however, on average, men improved by 1.5 points from simulation 1 to simulation 3. This shows that **women's leadership performance improved, on average, more than men's did.** In the TeamSTEPPS arm, the mean change in Leadership Performance CTS score was significantly different between men and women, with women on average improving by 2.5 points more than men (t=2.6, p-value=0.01).

Table 7. Difference in Mean Change in Leadership Performance CTS Scores Between Men and Women

	Mean difference		p-value
Treatment			
TeamSTEPPS	2.5	2.6	0.014*
LEADS	0.32	0.37	0.71
Control	-1	-0.86	0.39

*significant at α=0.05

The difference in change in Leadership Performance CTS™ score between men and women was not significant in the LEADS or control groups.

Men in TeamSTEPPS did not show significant improvement from simulation 1 to simulation 3 (t=1.3, p-value=0.23); however, women

in TeamSTEPPS showed significant improvement (t=6.5, p-value<0.0001), averaging a 3.5-point increase from simulation 1 to simulation 3 (Table 8). On average, women in LEADS improved in leadership performance by 2.2 points (t=3.5, p-value=0.002), and men in LEADS improved on average by 1.9 points (t=3.11, p-value=0.006). Neither women (t=0.6, p-value=0.55) nor men (t=1.6, p-value=0.15) in the control group showed significant change in leadership performance between simulation 1 and simulation 3. In summation, **both men and women showed significant improvement in leadership performance post-intervention when they took part in the LEADS curriculum**. This is in contrast to the TeamSTEPPS curriculum, in which only women showed significant improvement in leadership performance. These findings imply that LEADS may be a preferred curriculum for enhancing leadership performance for individuals in EM and OB/GYN residencies around the country.

Table 8. Difference in Mean Change in Leadership Performance CTS Score Within Each Intervention Arm Stratified by Sex

	Male			Female		
	mean	t	p-value	mean	t	p-value
Treatment						
TeamSTEPPS	1.0	1.3	0.23	3.5	6.5	<0.0001*
LEADS	1.9	3.1	0.006*	2.2	3.5	0.002*
Control	1.5	1.6	0.15	0.44	0.6	0.55

*significant at α=0.05

Detailed Leadership Evaluation:

Simulation 3:

For simulation 3, multiple leader-specific DLE variables were found to be associated with intervention arm. **Degree of stress (p-value=0.05), leader planned work to be done (p-value=0.02), timeliness of communication (p-value=0.02), accurate communication (p-value=0.04), and shared knowledge communication (p-value=0.02) were all found to be significantly associated with the intervention arm.**

Nine of the 14 detailed leadership evaluation variables were decided to be the most important measures of leadership performance: degree of leadership, leader decided what should be done, leader assigned tasks, frequent communication, timeliness of communication, accurate communication, problem-solving communication, shared knowledge communication, and mutual respect communication. **For 89% of these measures, the LEADS arm had a larger proportion of individuals in the highest category**, indicating the best performance. The control group had the lowest proportion of individuals in the highest category for every DLE measure. This may indicate that, although power to detect statistical significance is low, a pattern may exist in the leadership skills between the three arms of the study.

For simulation 1, the only leader-specific DLE variable that was associated with the intervention arm was degree of stress (p-value=0.03), indicating that the distribution of stress level differed for individuals within each intervention arm. Of the individuals who appeared “not stressed at all,” 50% were in the TeamSTEPPS curriculum group.

DLE Composite Score:

The mean DLE composite score was significantly higher for participants of the LEADS curriculum compared with the control group (p-value<0.05). There were no significant differences in mean of the composite DLE when comparing LEADS vs. TeamSTEPPS, nor TeamSTEPPS vs. controls. T-tests comparing means of the composite score by sex within each intervention arm revealed that there were no significant differences between men and women (Table 9).

Table 9. Difference in Mean DLE Composite Between Men and Women Within Each Intervention Arm								
	Male			Female				
	mean	95% CI	SD	mean	95% CI	SD	t	p-value
Treatment								
TeamSTEPPS	4.2	(3.7,4.7)	0.88	4.2	(3.9,4.4)	0.66	0.13	0.89
LEADS	4.4	(4.0,4.7)	0.76	4.4	(4.1,4.6)	0.67	0.00	0.99
Control	3.7	(3.2,4.1)	0.69	4.0	(3.5,4.5)	0.75	1.12	0.28

List of Publications and Products

Given that this study included a randomized control trial, we are not able to publish any data during the study period to prevent the risk of contamination of study participants. Therefore, all data collection activities needed to be completed prior to writing manuscripts.

We currently have four papers drafted based on the results from Aims 1 and 2, and we are drafting three more papers based on results from Aim 3.

Aim 1 and 2 papers:

1. Guise JM, Hansen M, Ross H, et al. LEADS: Development of a curriculum & toolkit for residents in leadership and addressing unconscious bias. (to be submitted) (descriptive paper)

2. Hansen M, Kinsey E, Guise JM, et al. Perceptions of leadership and attainment of clinical leadership skills: results from the national leadership survey of EM and OB/GYN residents. (to be submitted) (quant paper)

3. Hansen M, Skarica B, Guise JM, et al. Implicit gender bias among US resident physicians. (to be submitted) (IAT)

4. Guise, JM, Harrod T, Hansen M, et al. Residents' Reported Experiences Leading Healthcare Teams & Their Perceived Educational Needs: Responses from a National Leadership Survey OB/GYN and Emergency Medicine Programs. (to be submitted) (bystander/qual paper)

REFERENCES

1. Joint Commission on Accreditation of Healthcare Organizations. *Update: Sentinel event statistics.* 2006.
2. Salas E, Burke C, Stagl K. Developing teams and team leaders: Strategies and principles. *Leader development for transforming organizations: Growing leaders for tomorrow*. New York, NY: Routledge; 2004:325-355.
3. Cannon-Bowers J, Tannenbaum S, Salas E, Volpe C. Defining competencies and establishing team training requirements In: R. Guzzo, E. Salas, & Associates (Eds.), Team effectiveness and decision making in organizations, pp. 333-80. San Francisco, CA; 1995.
4. Borrill C, Carletta J, Carter A, et al. *The effectiveness of health care teams in the National Health Service.* Birmingham, AL: University of Aston; 2000.
5. Stein L. The doctor-nurse game *Arch Gen Psych.* 1967;16(6):699-703.
6. Haddad A. The nurse/physician relatonship and ethical decision making. *AORN J.* 1991;53(1):151-154.
7. Stegman C. *The effects of social and environmental factors on informal teaching-learning transactions between female nurses and female physicians* Fort Lauderdale, FL, Nova University; 1987.
8. Pringle R. *Sex and medicine: Gender, power and authority in the medical profession.* Cambridge, UK Cambridge University Press; 1998.
9. Bartels C, Goetz S, Ward E, Carnes M. Internal medicine residents' perceived ability to direct patient care: impact of gender and experience. *J Womens Health (Larchmt).* 2008;17(10):1615-1621.
10. Wear D, Keck-McNulty C. Attitudes of female nurses and female residents toward each other: A qualitative study in one U.S. teaching hospital. *Acad Med.* 2004;79(4):291-301.
11. Cooper S, Wakelam A. Leadership of resuscitation teams: "Lighthouse leadership." *Resuscitation.* 1999;42(1):27-45.
12. Cooper S. Developing leaders for advanced life support: evaluation of a training programme. *Resuscitation.* 2001;49(1):33-38.
13. Guise J, Deering S, Kanki B, et al. Validation of a tool to measure and promote clinical teamwork. *Simul Healthc.* 2008;3(4):217-223.
14. Hemphill J, Coons A. Leadership behavior description questionnaire (LBDQ). Columbus, OH: Ohio State University; 1950.
15. Nalaka E, Edirisinghe S, Rajulu S, Rajulu S. Micro-learning: a way to enhance learning pathways.