The American Journal of Surgery 221 (2021) 285–290

Contents lists available at ScienceDirect

The American Journal of Surgery

journal homepage: www.americanjournalofsurgery.com

Multi-Disciplinary Trauma Evaluation and Management Simulation (MD-TEAMS) training for emergency medicine and general surgery residents

Katharine E. Caldwell [a, *], Al Lulla [b], Collyn T. Murray [b], Rahul R. Handa [a], Ernesto J. Romo [b], Jason W. Wagner [b], Paul E. Wise [a], Jennifer M. Leonard [a], Michael M. Awad [a]

[a] Department of Surgery, Washington University in St. Louis, St. Louis, MO, United States
[b] Division of Emergency Medicine, Washington University in St. Louis, St. Louis, MO, United States

ARTICLE INFO

Article history:
Received 8 June 2020
Received in revised form
11 August 2020
Accepted 7 September 2020

Keywords:
Surgical education
Simulation
Trauma
Communication
Multi-disciplinary

ABSTRACT

Background: Successful trauma resuscitation relies on multi-disciplinary collaboration. In most academic programs, general surgery (GS) and emergency medicine (EM) residents rarely train together before functioning as a team.

Methods: In our Multi-Disciplinary Trauma Evaluation and Management Simulation (MD-TEAMS), EM and GS residents completed manikin-based trauma scenarios and were evaluated on resuscitation and communication skills. Residents were surveyed on confidence surrounding training objectives.

Results: Residents showed improved confidence running trauma scenarios in multi-disciplinary teams. Residents received lower communication scores from same-discipline vs cross-discipline faculty. EM residents scored higher in evaluation and planning domains; GS residents scored higher in action processes; groups scored equally in team management. Strong correlation existed between team leader communication and resuscitative skill completion.

Conclusion: MD-TEAMS demonstrated correlation between communication and resuscitation checklist item completion and communication differences by resident specialty. In the future, we plan to evaluate training-related resident behavior changes and specialty-specific communication differences by residents.

Introduction

Trauma remains the leading cause of death for individuals under the age of 44.[1] Patients who present with traumatic injuries require the coordinated care of a resuscitative team to provide precise, life-saving intervention. The success of the resuscitative team relies on clear communication and effective leadership.[2] Non-technical errors during the initial resuscitation are responsible for up to one third of preventable errors leading to in-hospital, trauma-related death.[3] Team-based training of trauma teams to improve non-technical and communication skills is essential to improving overall trauma outcomes[4,5] and many metrics of trauma patient care, including reduction of time to critical operation[6] or intervention.[7] These findings underscore the need for communication-based education to provide effective resuscitation and to prevent trauma-related deaths.

Simulation-based training with facilitated feedback is an effective method for teaching and assessing communication,[8] trauma management,[9,10] and leadership skills.[7,11] High-fidelity, manikin-based simulations are conducive to trauma training due to the ability to create changes in physiology, practice performance of resuscitative procedures,[12] and offers benefits in competency-based assessment.[13]

While simulation-based training is well-described, many of these interventions target only one group of residents[8–10,14] or inter-professional teams.[7,15–17] Few studies have targeted specific trauma team training to multi-disciplinary teams of emergency medicine (EM) and general surgery (GS) residents.[18] Additionally, while many instruments have been developed to evaluate the resuscitative team performance,[19–21] no specific instrument has been developed for the evaluation of the resuscitation team leader with items which could be used for resident evaluation and

* Corresponding author.
E-mail address: kecaldwell@wustl.edu (K.E. Caldwell).

https://doi.org/10.1016/j.amjsurg.2020.09.013
0002-9610/

milestone assessment. Here, we describe the implementation and results of our Multi-Disciplinary Trauma Evaluation and Management Simulation (MD-TEAMS) training for EM and GS residents and describe novel tools for evaluation of simulation skills and communication of team leaders.

Methods

Simulation development and evaluation

At Washington University in St. Louis, EM and GS residents alternate trauma resuscitation leadership roles in the trauma bay. Despite significant, independent resident trauma training ongoing within the Department of Surgery[22] and the Division of Emergency Medicine, no previous efforts existed at our institution to implement a combined trauma-management simulation. This was identified as an area in need of improved education by both groups to advance trauma patient care.

Trauma simulation cases were developed internally based on pre-identified goals for evaluation by a consensus panel of expert faculty and resident members. Cases were divided into essential steps for appropriate medical resuscitation. Each item was noted by a checkbox and represented a single point that resident teams could achieve by completion of the associated item. The total items for each resuscitation varied by scenario, ranging from 40 to 52 unique items which had to be completed by each team. Description of cases can be reviewed in Table 1, and full cases are found in Supplemental Materials 1. To account for differences in rater scores, error rates were calculated as 0.5 points deducted for each rater's indication that an event did not occur or was performed incorrectly.

Leadership assessment development and evaluation

For our leadership assessment tool (Supplementary Materials 2), team leaders were evaluated in three domains as has been previously described by Rosenman et al.[11]: evaluation and planning, action processes, and interpersonal skills/team management. We reviewed checklist items included in previously validated trauma team checklists (Nontechnical Skills for Trauma (T-NOTCHES),[19] Trauma Emergency Assessment Measure (TEAM),[20] and Trauma Team Communication Assessment (TTCA-24)[21]) and adapted items that directly applied to team leader performance, which mapped to ACGME competencies.[23] A total of thirteen items were selected. Faculty evaluators rated each resident on a Likert scale of one to five, with one representing ineffective and five representing effective communication in that domain. A maximum score was 65. Evaluators were not aware which items corresponded to each domain as to prevent bias or false congruency.

Survey development

Before and after curricular participation residents were asked to rate their comfort leading a trauma resuscitation and working on a multi-disciplinary team on a Likert scale of one to five, with one indicating no comfort and five indicating total comfort. After the curriculum, residents were additionally surveyed on their opinion of the simulation training on a five-point Likert scale, with one representing not useful and five representing very valuable.

Simulation

This was a single-institution study of 18 residents, four PGY2 and five PGY4 GS residents and nine PGY3 EM residents. At our institution, GS residents complete five years of clinical training, while EM residents complete four. All simulations were conducted on a single day at Washington University in St. Louis. Residents were divided into teams of six members with three EM residents and three GS residents present on each team.

Two EM nursing staff were present to assist and served the role of the nurse in each room. The nurses involved in this study were both ER nurses, each with greater than five years of experience working in trauma resuscitation. Nurses were instructed to not perform tasks independently during the resuscitation, but like other team members, were allowed to provide suggestions to the team leader on next steps, or to prompt the team leader for noted missing tasks.

Residents received a brief, 30-min didactic session before simulations reminding them of the appropriate steps for ATLS trauma-resuscitation and to discuss common pitfalls of communication within the trauma resuscitation bay. Resident teams rotated through three separate scenarios with 15 minutes to complete the scenario, and 5 minutes to debrief the scenario events with faculty observers.

MetiMen mannequins were used in all scenarios. Trained simulation operators were stationed in the control room and were responsible for monitoring the scenario and appropriately changing vital signs according to the scenario and the resident activities. Two faculty members with expertise in trauma resuscitation, one EM and one GS, monitored each scenario from the control room. Residents had access to all necessary equipment in the room to assess patients or perform indicated procedures. Teams selected their own unique resuscitation leader for each scenario. Teams were assessed both for their ability to successfully perform medical tasks and on the leadership skills of the team leader as described above.

Statistical analysis

Graph pad statistical software (version 8.2.0) and SPSS (version 26) were used for data analysis and figure development. Descriptive statistics were evaluated, and data are expressed as raw scores or as mean and standard deviation, as noted in each figure. Inter-rater reliability was determined by evaluation of Cohen's kappa statistic with a kappa statistic >0.40 considered to be moderate and >0.60 considered to be substantial inter-rater reliability, per standard analysis.[24] Differences in communication scores between

Table 1
Summary of trauma scenarios.

Scenario 1
Young M patient brought in after multi-GSW. Resident teams must identify and prioritize management of multiple injuries including management of hemothorax requiring chest tube placement and loss of airway requiring intubation.
Scenario 2
Patient involved in MVC due to stroke. Residents must manage multiple fractures and need to identify and manage atrial fibrillation and concern for stroke.
Scenario 3
Elderly patient s/p fall with significant LE deformity and brain bleed. Patient develops herniation and resident teams must prioritize management of herniation over management of LE fracture.

groups were evaluated by Mann Whitney U testing for non-parametric, non-paired, independent groups. Pearson's r was used to evaluate correlation between communication scores and total percent of checklist items completed after linearity of the relationship was confirmed. A p value < 0.05 was considered statistically significant for all analyses.

Results

A total of eighteen residents, nine GS and nine EM residents, participated in MD-TEAMS on a single day. Total error rates varied by scenario and faculty rater, however in all groups, completion of the secondary survey and completion of appropriate adjuncts demonstrated significant errors across multiple teams (Supplemental Fig. 1). Common errors included failure to order trauma labs, place a Foley catheter, and order appropriate studies.

We assessed the inter-rater reliability of the simulation scoring instruments and communication instrument. Moderate to substantial inter-rater reliability, as assessed by kappa statistic, existed in all three scenarios: kappa 0.45, 0.86, 0.48. Moderate inter-rater reliability also existed in the use of the communication scoring tool across all scenarios as measured by a weighted kappa statistic to account for a categorical scale: weighted kappa 0.41.

No significant differences were noted between the overall resuscitation checklist or communication scores of EM or GS team leaders. Additionally, no significant differences existed in the overall ratings of faculty in EM or GS disciplines in either general resuscitation scores or communication ratings. Despite overall similarities in total scores, we performed further subgroup analysis and determined that significant differences did exist between the communication ratings of matched attending/resident pairs and unmatched attending/resident pairs (p 0.05). (Fig. 1A). When GS residents were scored by GS faculty they were scored lower (mean 53.6) than when scored by EM faculty (mean 58.4). Likewise, EM residents scored lower when scored by EM faculty (mean 54.8) as compared to the scores given by GS faculty (mean 58) (Fig. 1B).

Differences in team leader communication scores were also discovered on sub-group analysis of communication skill subsets. *A priori* design of the communication tool divided the rating metrics into a total of three major subcategories of leadership skills (Supplemental Fig. 2): evaluation and planning, action processes, and interpersonal skills/team management. GS resident team leaders demonstrated higher scores in communication skills centered around action planning (mean GS 4.60 v EM 3.96, p < 0.01), while EM residents demonstrated higher scores in skills in the category of evaluation and planning (mean EM 4.54 v GS 3.98, p < 0.01) (Fig. 2). There was no significant difference in the ratings of team management of either group (mean GS 4.44 v EM 4.27, p 0.3).

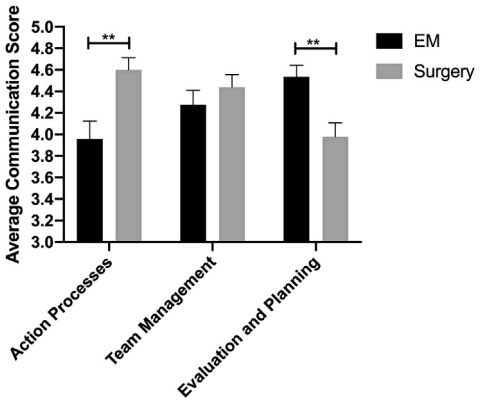

Fig. 2. Differences in communication scores exist between surgical and EM residents. GS residents (n = 10) demonstrated higher scores in communication skills centered around action processes mean 4.60 v 3.96 (p < 0.01) while EM residents (n = 8) demonstrated greater scores in communication ratings based on evaluation and planning mean 4.54 v 3.98 (p < 0.01). There was no significant difference between the two groups in team management-based skills mean 4.27 v 4.44 (p 0.32).

Finally, we found significant correlation between the performance of teams on essential medical items on the scenario checklist to total communication scores. Teams headed by leaders who received higher communication scores resulted in higher percentages of completed resuscitation checklist items, with the correlation varying by scenario (Fig. 3). Scenario 1 (Fig. 3A) and 3 (Fig. 3C) showed the strongest correlations with Pearson's r of 0.92 (p < 0.01) and 0.95 (p < 0.01) respectively. Scenario 2 (Fig. 3B) demonstrated weaker correlation, which did not reach statistical significance (Pearson's r 0.29, p 0.58).

Improvements were seen in overall resident comfort running trauma resuscitations and in working with residents among other disciplines (Fig. 4A). Residents were also asked about their satisfaction with the curriculum as a method of learning trauma resuscitation and team management skills. Overall on a Likert scale from one (not useful) to five (very valuable) the curriculum was given an average rating of 4.44 with 94.4% of residents scoring the

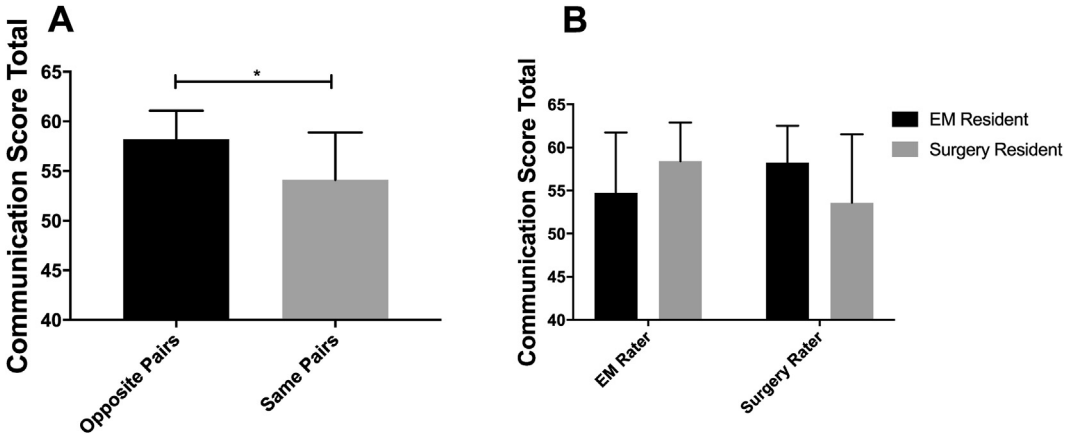

Fig. 1. Differences exist between matched attending/resident ratings and unmatched resident/attending ratings. A) Matched pairs of residents and attendings (n = 9) resulted in lower communication scores (mean 54.2) than unmatched pairs (n = 9) (mean 58.3) (p 0.05). B) Breakdown of matched pairs by specialty: EM faculty with EM resident (n = 4), mean 54.8, EM faculty with GS resident (n = 5) mean 58.4, GS faculty with EM resident (n = 4) mean 58, GS faculty with GS resident (n = 5) mean 53.6.

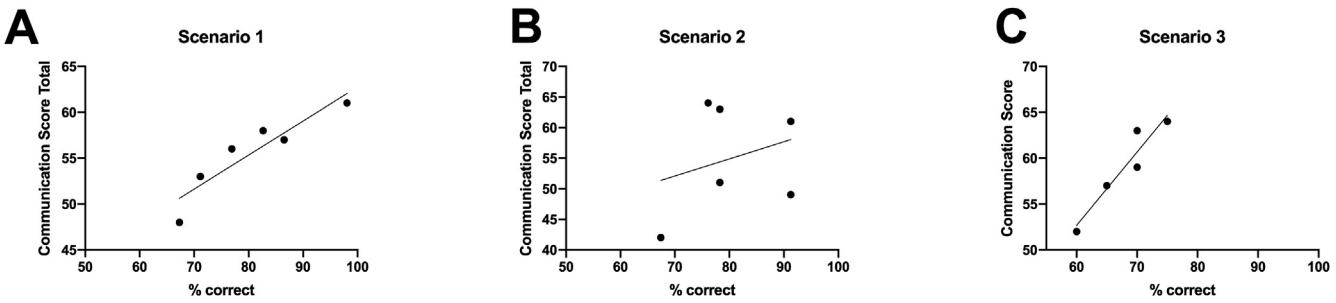

Fig. 3. Communication scores predict total percentage of resuscitation checklist items completed.
A), B), C) Correlation between the communication scores of a team leader and the total percentage of checklist items which were completed by the team: scenario 1 (A) Pearson's r 0.92 $p < 0.01$, scenario 2 (B) Pearson's r 0.29 p 0.6, and scenario 3 (C) Pearson's r 0.95 $p < 0.01$.

curriculum as valuable or very valuable (Fig. 4B). Specific subgroup breakdown of changes in comfort level between specialty and PGY years can be found in Supplemental Fig. 2 demonstrating greater differences in the comfort of PGY3 EM and PGY4 GS residents in leading trauma simulations.

Discussion

EM and GS residents are regularly expected to interact in high-pressure scenarios in the trauma resuscitation bay, however these residents do not often have the ability to train together before it is necessary that they work together. Our intention in the design of MD-TEAMS was to provide multi-disciplinary training in order to improve resident comfort working on resuscitative teams, as well as to develop a tool for measurement of specific leadership skills and weaknesses of resident resuscitation team leaders.

From our evaluation of trends in scoring, we noted three interesting findings. First, same-discipline pairs of evaluators and participants resulted in lower scores for that participant when compared to opposite-pair scores. This was noted, despite overall lack of differences in total scores in both simulation and communication evaluations. We hypothesize this effect is due to increased standards held by faculty members for residents within their own specialties when compared to residents in other specialties. In psychology research, evaluations of poor performing in-group members has been demonstrated to be judged more harshly by in-group members than poor performance of an out-group member; a phenomena known as the "black sheep effect".[25,26] We could find no previous studies which have described this phenomena in the evaluation of residents. Though this would be impossible to complete in a single-center institution where faculty are familiar with residents within their own discipline, use of outside, expert evaluators could be used to blind faculty as to the specialty of each resident, removing this evaluation effect for larger studies of multi-disciplinary team training.

Secondly, it was noted that while EM and GS residents were scored equally well in overall communication abilities, they demonstrated differences in leadership skills. EM residents outscored GS residents in evaluation and planning skills, while GS residents performed higher in communication skills related to action processes. Both groups scored equally in their ability to manage a team. Evaluation and planning skills involve skills focused on strategy formulation, sharing of knowledge to facilitate understanding, and reflection. Action processes involved domains which track a team's progress toward the goal such as coordinating the activities of other team members and monitoring the performance of team members. As the communication metric used in this study has been internally developed, we were unable to compare it to prior studies, however we anticipate that the difference in skill sets is likely attributable to the types of training residents in different disciplines receive or to personalities typically drawn to different specialties. We were unable to find any studies that directly compared differences in communication styles in the trauma bay between GS and EM residents and recommend this as an area of further research.

Finally, we were able to demonstrate a direct link between the

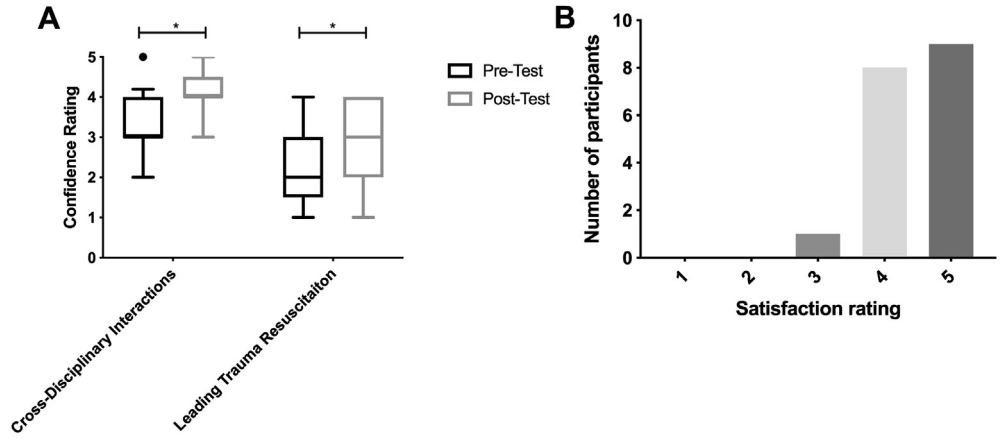

Fig. 4. Trends toward increased comfort managing trauma resuscitations and working with other specialty residents existed. A) Residents (n = 18) confidence in cross-disciplinary interactions between pre- and post-simulation self-assessments increased from median 3 to 4 (p 0.03) and in leading trauma resuscitations median 2 to 3 (p 0.03). B) Average resident rating of simulation training on 5-point Likert scale from 1 (not useful) to 5 (very valuable).

communication skills of team leaders and the ability of teams to successfully complete necessary steps in a trauma resuscitation. The ability to study the link between communication and essential task completion in a simulation setting is an attractive method to assess communication-based curricula and creates important data for the ability to assess the link between communication and patient safety.

Several communication tools have been developed to score the communication skills of residents in patient care settings.[27,28] Many of these address resident communication with a patient and not within a team, and thus, cannot be used for trauma scenarios. We found that none of these fit our goal of providing specific feedback to the team leader in the domains we were hoping to assess. Further, traditional trauma resuscitation communication evaluations, such as the TEAM, T-NOTECHS, and TTCA-24, often fail to assess the behavior of the resuscitation team leader in isolation, preventing specific resident assessment and feedback. For this reason, we sought to create our own instrument which used domains and anchors similar to several prior described in the TEAM, T-NOTECHS, and TTCA-24, but focused on specific, actionable feedback which could be provided to a resident to improve their performance or used to assess resident acquisition of milestones. We believe this instrument demonstrates significant promise as a method for resident leadership evaluation in the trauma bay and could be used to provide measurable feedback to residents on their performance for resident evaluation and training.

The strength of our intervention is in demonstrating benefits of multi-disciplinary training for trauma resuscitation and management. With this curriculum, we were able to demonstrate significant resident satisfaction (Kirkpatrick level 1[30]) with the learning intervention. We did not assess pre- and post-simulation training changes in resident behavior in the trauma bay or patient outcomes after the study, but we were able to identify interesting trends in error rates. This training allows for simulation learning interactions between EM and GS residents, which are critical as multi-disciplinary training has been proven to be important in developing effective team-dynamics.[16,29] Additionally, we were able to describe a new metric for assessment of team-leader communication in trauma simulation, which correlates with team performance on essential task completion and demonstrates inter-relater reliability within this small, single-institution sample. Further multi-institutional use of the metric might establish validity.

This study exhibits some limitations common of single-institution studies. Our study included a limited sample size and a single day intervention. Due to sample size, participation of only PGY3 EM residents, and team leaders being senior PGY4 residents in the GS cohort, we were unable to perform any subgroup analysis to compare the performance of junior and senior residents in the leadership position. PGY3 EM and PGY4 GS residents demonstrated the greatest changes in comfort leading trauma scenarios as is expected from their greater participation in this role. In future iterations of this curriculum, expanding the cohort of residents to include more junior residents would allow for comparison of curricular effects between junior and senior resident. While moderate kappa agreement existed between raters on communication scores, there were variations by the groups of raters and the team evaluated, with some excellent and some poor agreement which did not cluster toward high or low performing teams. This indicates that additional studies would be needed to validate this tool. Further, while significant agreement did exist between scenario raters, some disagreement existed on the completion of key tasks during the simulation. Asking faculty to score teams in real time is challenging as there is significant activity occurring simultaneously in a trauma resuscitation scenario and the speed of events can be very rapid. Further testing with review of videos of the team scenarios would likely allow for more accurate scoring, as faculty would be able to watch videos more than once or more closely review especially difficult areas in order to obtain more accurate assessments. Of importance, is the fact that this study did not address the role of interdisciplinary interactions in trauma resuscitation. ER nurses did participate in two of the three scenarios, however we chose to not fully assess the role of nursing staff due to this staffing limitation. Nursing staff plays an essential role in live trauma resuscitation, often acting independently to complete checklist items or reminding team leaders of incomplete items. We did not assess the number of completed checklist items that required prompting from nursing staff or other residents on the team. We believe this would be an important and interesting direction for additional iterations in this curriculum and would like to use this curriculum in the future to better assess the interactions of interdisciplinary and not just multidisciplinary teams. Finally, some residents may have already had significant experiences in running traumas or working in the emergency room. Though only a few weeks into the academic year, the resident's experience likely varied significantly based on their previous rotations, which may explain the variation in comfort and skill level which existed between residents. Further studies with larger groups of residents or longitudinal data over several years would be better able to validate this training and could demonstrate more benefits. Additional studies, including assessing changes in live trauma resuscitation behavior, after the simulation is completed, would be necessary to demonstrate higher Kirkpatrick level changes from such a training.

MD-TEAMS training allowed us to discover several interesting uses for multi-disciplinary training and findings regarding faculty evaluation and subspecialty resident skills. As we have described here, multi-disciplinary team training can be used as a method of common error identification to determine if recurrent errors exist. Though not attempted based on this study, such information could be used to identify areas in need of quality improvement. Additionally, error identification in simulation scenarios could be compared to patient outcome data or real-time trauma resuscitation video to determine if interventions might be required. Larger interventions on the patient care side to reduce errors could also be further tested in the simulation lab. We suggest these be future goals of evaluation of multi-disciplinary team training research.

Conclusion

MD-TEAMS is an effective way to increase EM and GS resident comfort working with residents of different teams and in running trauma scenarios. Additionally, it offers interesting insight into differences in resident scoring by same or cross-discipline faculty, and differences in communication styles between resident specialties. Team-training can be used to identify individual resident strengths and weaknesses in leadership, as well as common or repeated errors by resident teams, both of which may be used as a needs assessment to develop focused curricula in the future.

Declaration of competing interest

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. Salary support for KEC was provided by the Washington University School of Medicine Surgical Oncology Basic Science and Translational Research Training Program grant **T32CA009621**, from the National Cancer Institute (NCI). All other authors declare no relevant conflicts of interest.

Acknowledgements

Thanks to Julie Woodhouse, simulation operators, and other staff members of the Wood Simulation Center for their assistance in setting up and providing technical expertise for simulation management.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.amjsurg.2020.09.013.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. Salary support for KEC was provided by the Washington University School of Medicine Surgical Oncology Basic Science and Translational Research Training Program grant **T32CA009621**, from the National Cancer Institute (NCI)

References

1. WISQARS (Web-based injury statistics query and reporting System) Injury Center CDC. Published September 18 https://www.cdc.gov/injury/wisqars/index.html; 2019. Accessed November 20, 2019.
2. Cooper S, Wakelam A. Leadership of resuscitation teams: "lighthouse leadership". *Resuscitation*. 1999;42(1):27–45. https://doi.org/10.1016/s0300-9572(99)00080-5.
3. MacLeod JBA, Cohn SM, Johnson EW, McKenney MG. Trauma deaths in the first hour: are they all unsalvageable injuries? *Am J Surg*. 2007;193(2):195–199. https://doi.org/10.1016/j.amjsurg.2006.09.010.
4. Vioque SM, Kim PK, McMaster J, et al. Classifying errors in preventable and potentially preventable trauma deaths: a 9-year review using the Joint Commission's standardized methodology. *Am J Surg*. 2014;208(2):187–194. https://doi.org/10.1016/j.amjsurg.2014.02.006.
5. Pucher PH, Aggarwal R, Batrick N, Jenkins M, Darzi A. Nontechnical skills performance and care processes in the management of the acute trauma patient. *Surgery*. 2014;155(5):902–909. https://doi.org/10.1016/j.surg.2013.12.029.
6. Murphy M, Curtis K, Lam MK, Palmer CS, Hsu J, McCloughen A. Simulation-based multidisciplinary team training decreases time to critical operations for trauma patients. *Injury*. 2018;49(5):953–958. https://doi.org/10.1016/j.injury.2018.01.009.
7. Capella J, Smith S, Philp A, et al. Teamwork training improves the clinical care of trauma patients. *J Surg Educ*. 2010;67(6):439–443. https://doi.org/10.1016/j.jsurg.2010.06.006.
8. Knudson MM, Khaw L, Bullard MK, et al. Trauma training in simulation: translating skills from SIM time to real time. *J Trauma*. 2008;64(2):255–263. https://doi.org/10.1097/TA.0b013e31816275b0. discussion 263-264.
9. Miyasaka KW, Martin ND, Pascual JL, Buchholz J, Aggarwal R. A simulation curriculum for management of trauma and surgical critical care patients. *J Surg Educ*. 2015;72(5):803–810. https://doi.org/10.1016/j.jsurg.2015.03.001.
10. Lee SK, Pardo M, Gaba D, et al. Trauma assessment training with a patient simulator; a prospective, randomized study. *J Trauma*. 2003;55(4):651–657. https://doi.org/10.1097/01.TA.0000035092.83759.29.
11. Rosenman ED, Shandro JR, Ilgen JS, Harper AL, Fernandez R. Leadership training in health care action taems: a systematic review. *Acad Med*. 2014;89(9):1295–1306. https://doi.org/10.1097/ACM.0000000000000413.
12. Berkenstadt H, Ben-Menachem E, Simon D, Ziv A. Training in trauma management: the role of simulation-based medical education. *Anesthesiol Clin*. 2013;31(1):167–177. https://doi.org/10.1016/j.anclin.2012.11.003.
13. Ziv A, Rubin O, Sidi A, Berkenstadt H. Credentialing and certifying with simulation. *Anesthesiol Clin*. 2007;25(2):261–269. https://doi.org/10.1016/j.anclin.2007.03.002.
14. Moorman ML, Capizzani TR, Feliciano MA, French JC. A competency-based simulation curriculum for surgical resident trauma resuscitation skills. *Int J Crit Illness Inj Sci*. 2017;7(4):241–247. https://doi.org/10.4103/IJCIIS.IJCIIS_12_17.
15. Holcomb JB, Dumire RD, Crommett JW, et al. Evaluation of trauma team performance using an advanced human patient simulator for resuscitation training. *J Trauma*. 2002;52(6):1078–1086. https://doi.org/10.1097/00005373-200206000-00009.
16. Roberts NK, Williams RG, Schwind CJ, et al. The impact of brief team communication, leadership and team behavior training on ad hoc team performance in trauma care settings. *Am J Surg*. 2014;207(2):170–178. https://doi.org/10.1016/j.amjsurg.2013.06.016.
17. Shapiro MJ, Morey JC, Small SD, et al. Simulation based teamwork training for emergency department staff: does it improve clinical team performance when added to an existing didactic teamwork curriculum? *BMJ Qual Saf*. 2004;13(6):417–421. https://doi.org/10.1136/qshc.2003.005447.
18. Sullivan S, Campbell K, Ross JC, et al. Identifying nontechnical skill deficits in trainees through interdisciplinary trauma simulation. *J Surg Educ*. 2018;75(4):978–983. https://doi.org/10.1016/j.jsurg.2017.10.007.
19. Steinemann S, Berg B, DiTullio A, et al. Assessing teamwork in the trauma bay: introduction of a modified "NOTECHS" scale for trauma. *Am J Surg*. 2012;203(1):69–75. https://doi.org/10.1016/j.amjsurg.2011.08.004.
20. Cooper S, Cant R, Porter J, et al. Rating medical emergency teamwork performance: development of the Team Emergency Assessment Measure (TEAM). *Resuscitation*. 2010;81(4):446–452. https://doi.org/10.1016/j.resuscitation.2009.11.027.
21. DeMoor S, Abdel-Rehim S, Olmsted R, Myers JG, Parker-Raley J. Evaluating trauma team performance in a Level I trauma center: validation of the trauma team communication assessment (TTCA-24). *J Trauma Acute Care Surg*. 2017;83(1):159–164. https://doi.org/10.1097/TA.0000000000001526.
22. Onufer EJ, Cullinan DR, Wise PE, Punch LJ. Trauma technical skill and management exposure for junior surgical residents - the "SAVE lab 1.0. *J Surg Educ*. 2019;76(3):824–831. https://doi.org/10.1016/j.jsurg.2018.12.003.
23. Nasca TJ, Philibert I, Brigham T, Flynn TC. The next GME accreditation system–rationale and benefits. *N Engl J Med*. 2012;366(11):1051–1056. https://doi.org/10.1056/NEJMsr1200117.
24. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33(1):159–174.
25. Steain A, Stanton CJ, Stevens CJ. The black sheep effect: the case of the deviant ingroup robot. *PLoS One*. 2019;14(10), e0222975. https://doi.org/10.1371/journal.pone.0222975.
26. Abrams D, Palmer SB, Rutland A, Cameron L, Van de Vyver J. Evaluations of and reasoning about normative and deviant ingroup and outgroup members: development of the black sheep effect. *Dev Psychol*. 2014;50(1):258–270. https://doi.org/10.1037/a0032461.
27. Trickey AW, Newcomb AB, Porrey M, et al. Assessment of surgery residents' interpersonal communication skills: validation evidence for the communication assessment tool in a simulation environment. *J Surg Educ*. 2016;73(6):e19–e27. https://doi.org/10.1016/j.jsurg.2016.04.016.
28. Yule S, Flin R, Paterson-Brown S, Maran N, Rowley D. Development of a rating system for surgeons' non-technical skills. *Med Educ*. 2006;40(11):1098–1104. https://doi.org/10.1111/j.1365-2929.2006.02610.x.
29. Steinemann S, Berg B, Skinner A, et al. In situ, multidisciplinary, simulation-based teamwork training improves early trauma care. *J Surg Educ*. 2011;68(6):472–477. https://doi.org/10.1016/j.jsurg.2011.05.009.
30. Kirkpatrick D, Kirkpatrick J. *Evaluating Training Programs: The Four Levels*. Berrett-Koehler Publishers; 2006.