Received: 25 October 2020 | Revised: 24 November 2020 | Accepted: 11 December 2020

DOI: 10.1002/emp2.12348

JACEP OPEN | WILEY

ORIGINAL RESEARCH

The Practice of Emergency Medicine

Does team leader gender matter? A Bayesian reconciliation of leadership and patient care during trauma resuscitations

Elizabeth D. Rosenman MD[1] | Anthony Misisco MA[2] | Jeffrey Olenick PhD[2,5] | Sarah M. Brolliar[1] | Anne K. Chipman MD, MS[1] | Marie C. Vrablik MD, MCR[1] | Georgia T. Chao PhD[2,6] | Steve W. J. Kozlowski PhD[2,6] | James A. Grand PhD[3] | Rosemarie Fernandez MD[4]

[1] Department of Emergency Medicine, University of Washington, Seattle, Washington, USA

[2] Department of Psychology, Michigan State University, East Lansing, Michigan, USA

[3] Department of Psychology, University of Maryland, College Park, Maryland, USA

[4] Department of Emergency Medicine, University of Florida, Gainesville, Florida

[5] Department of Psychology, Old Dominion University, Norfolk, Virginia

[6] Department of Psychology, University of South Florida, Tampa, Florida

Correspondence
Rosemarie Fernandez, MD, University of Florida, Department of Emergency Medicine, Gainesville, Florida.
Email: fernandez.r@ufl.edu

Funding and support: Funding and support for this project were provided by the Agency for Healthcare Research and Quality (1R18HS022458-01A1) and the Department of Defense (W81XWH-18-1-0089). The funding sources had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, approval, or decision to submit the manuscript.

Abstract

Objective: Team leadership facilitates teamwork and is important to patient care. It is unknown whether physician gender-based differences in team leadership exist. The objective of this study was to assess and compare team leadership and patient care in trauma resuscitations led by male and female physicians.

Methods: We performed a secondary analysis of data from a larger randomized controlled trial using video recordings of emergency department trauma resuscitations at a Level 1 trauma center from April 2016 to December 2017. Subjects included emergency medicine and surgery residents functioning as trauma team leaders. Eligible resuscitations included adult patients meeting institutional trauma activation criteria. Two video-recorded observations for each participant were coded for team leadership quality and patient care by 2 sets of raters. Raters were balanced with regard to gender and were blinded to study hypotheses. We used Bayesian regression to determine whether our data supported gender-based advantages in team leadership.

Results: A total of 60 participants and 120 video recorded observations were included. The modal relationship between gender and team leadership ($\beta = 0.94$, 95% highest density interval [HDI], -.68 to 2.52) and gender and patient care ($\beta = 2.42$, 95% HDI, -2.03 to 6.78) revealed a weak positive effect for female leaders on both outcomes. Gender-based advantages to team leadership and clinical care were not conclusively

Supervising Editor: Marna Greenberg, DO, MPH

supported or refuted, with the exception of rejecting a strong male advantage to team leadership.

Conclusions: We prospectively measured team leadership and clinical care during patient care. Our findings do not support differences in trauma resuscitation team leadership or clinical care based on the gender of the team leader.

KEYWORDS
Bayesian analysis, gender, leadership, resuscitation, teamwork, trauma

1 | INTRODUCTION

1.1 | Background

There is a growing body of research evaluating physician gender-based differences in health care delivery,[1,2] including 2 studies specifically evaluating the relationship between team leader gender and resuscitation performance.[3,4] Resuscitation team leadership is a critical skill for physicians from a variety of health care specialties.

Effective team leadership is linked to better patient care,[5] and failures in team leadership can lead to adverse events and present a threat to patient safety.[6,7] Significant gender-based differences in team leadership could have major implications for physician training, team dynamics, workplace culture, and patient preferences.

Research demonstrates there are gender-based differences in medical education assessment and feedback practices.[8–11] Specific to team leadership, Ju et al demonstrated that physician trainees scored female team leaders less favorably than male team leaders acting in a standardized simulation-based resuscitation.[12] In other words, male and female actors portraying the same, scripted, team leader role were evaluated differently, with male actors scoring more favorably.

1.2 | Importance

Understanding whether there are inherent gender-based differences in team leadership skills, separate from any superimposed bias in assessment practices, is important. A true difference in leadership skills should prompt modifications to current leadership training efforts within medical education to close this gap. However, if differences stem primarily from gender bias in assessment practices, efforts should focus on the assessment tools themselves, as well as training medical educators and others who use them. Two existing studies explore the relationship between team leader gender and patient care during resuscitations; however, these studies have conflicting results and neither provide a prospective evaluation of team leadership in the clinical setting.[3,4]

1.3 | Goals of this investigation

We prospectively evaluated both team leadership and patient care in emergency department (ED) trauma teams led by male and female physicians during actual trauma resuscitations. We used Bayesian methodology to expand on previous efforts[3,4] evaluating the relationship between gender and both team leadership and patient care to more accurately determine the likelihood of gender-based advantages to team leadership.

2 | METHODS

2.1 | Study design

We performed a secondary analysis of data from a larger randomized controlled trial to examine the relationship between team leader gender and the quality of (1) team leadership and (2) patient care.[5] In the original study we evaluated the impact of a simulation-based team leadership training on these same outcomes using video recordings of actual ED-based trauma resuscitations (ClinicalTrials.gov Identifier: NCT03155490). The current study uses the preintervention data collected from April 2016 to December 2017. The University of Washington Institutional Review Board approved this study.

2.2 | Participants and setting

Participants included 60 second- and third-year emergency medicine and general surgery residents assigned to the trauma resuscitation team leader role at Harborview Medical Center, an urban, Level 1 trauma center within the University of Washington that serves a 5-state region (AK, WA, WY, MT, ID), and has over 5000 trauma-related admissions per year. Per institution protocol, all participants were certified in Advanced Trauma Life Support before functioning as trauma team leaders. Study participation was voluntary and a research coordinator obtained written consent before video recording. Participants completed a demographic survey at the time of enrollment.

2.3 | Data collection

Adult trauma resuscitations were included if (1) they were led by an enrolled study participant and (2) they met trauma team activation criteria as outlined in the Harborview Medical Center Trauma Registry.[13] Resuscitations were excluded if the patient was (1) pregnant, (2) pronounced dead or left the ED within 5 minutes of arrival, (3) under do not resuscitate or comfort care orders, (4) in police custody, or (5) found to have non-traumatic mechanisms or isolated burns as a primary diagnosis. Videos were blurred to obscure patient identity. Two video-recorded observations were included for each participant, resulting in 120 observations. Patient characteristics and patient-related covariates (eg, injury severity score [ISS])[14] were extracted from the Harborview Medical Center trauma registry for each patient. We included ISS as a covariate in this team leadership study because ISS reflects the scope of tasks required by the team and team leader.

2.4 | Outcome measures and coding

The team leadership and patient care measures were developed through an iterative process that provided evidence of validity as described by Cook et al.[15] The specific steps included a thorough review of the literature and input from subject experts (content validity), as well as rater training and duplicate coding of a portion of the observations to determine interrater reliability (internal structure) across a range of performance levels.[5,16] Additionally, scores from the team leadership measure were shown to correlate with patient care measure scores as predicted by conceptual models of functional team leadership (relationship to other variables).[5,17]

The leadership measure targeted important team leadership behaviors identified through 2 systematic reviews and subject matter expert input.[18,19] Example behaviors that were captured before and during the resuscitation included information sharing, stating a plan, seeking input, task assignment, and initiating a team huddle. The maximum composite team leadership score was 38. The measure, as well as additional information regarding the supporting validity evidence, was published with the original study.[5]

The patient care measure was based on existing trauma care guidelines and checklists as well as input from subject matter experts.[20–25] The patient care measure was flexible, containing some conditional items that were dependent on patient condition. Example items that were universal to all observations included assessing the airway, assessing mental status, and obtaining/confirming vascular access. Example items that were conditional (ie, dependent on patient condition) included transfusing blood products and performing a focused assessment with sonography for trauma. The maximum composite patient care score ranged from 20 to 38. Scores were normalized to a 100-point scale to allow comparison across observations. The measure, as well as additional information regarding the supporting validity evidence, was published with the original study.[16]

Two independent groups of trained raters coded the observations for team leadership (n = 4) and patient care (n = 2). Raters were balanced with regard to gender and blinded to the study hypotheses. For the team leadership measure 56% of the observations (n = 67) were coded in duplicate. Following recommendations by Byrt et al, we calculated the probability and bias adjusted kappa (PABAK) for each item to adjust for prevalence given the measure was targeting low base rate events.[26] The mean PABAK was 0.97 across all items. For the patient care measure 15% of the observations (n = 18) were coded in duplicate, with a mean Cohen's κ of 0.72 across all items. Additional details regarding rater training and the coding process are available in the original study.[5]

The Bottom Line

Prior studies suggest that female team leaders perform less favorably than male team leaders in simulation-based resuscitation. This study used updated assessment strategies, including Bayesian analysis, to appraise leader performance during live clinical trauma resuscitations. The study found no gender-based performance advantages in either leadership quality or clinical care.

2.5 | Analysis

Team leader demographics and resuscitation characteristics were compared between female and male team leaders. Categorical data were compared using Pearson chi-square test of independence. Interval data were compared using independent-samples *t* test.

We elected to examine the effects of gender on leadership and patient care outcomes in our data using a Bayesian analytic framework.[27,28] In contrast to null hypothesis significance testing—which conventionally involves choosing a *single* "null" hypothesis (eg, the difference between male and female leadership skills = 0) and examining the probability that one's data/findings would be observed if that null hypothesis were true (eg, *p*(data|hypothesis))—Bayesian approaches attempt to summarize the probability of *all* possible hypotheses given the observed data/findings and prior beliefs about the plausibility of those possible hypotheses (*p*(hypothesis|data)).[29] Bayesian inference is particularly useful for interpreting the questions raised in the present study because (1) there are conflicting results in the literature about the magnitude and direction of gender differences in leadership and patient care and (2) it allows us to evaluate the plausibility of different hypotheses about these effects based on our data.

Consistent with standard practices for conducting a Bayesian analysis, our analyses proceeded as follows.[29] First, the statistical model for examining the effects of gender on our outcomes of interest was defined. Our data involved multiple observations of leadership skill and patient care for each participant; consequently, we specified a 2-level random effects regression model to account for non-independence in these nested data:

$$\text{Level1 (observation/patient)}: DV_{ij} = \pi_{0i} + \pi_{1i}\left(ISS_{ij}\right) + e_{1i}$$

$$\text{Level2 (participant)}: \ \pi_{0i} = \beta_{00} + \beta_{01}\left(Gender_i\right) + e_{2i}$$

$$\pi 1i = \beta 10$$

where DV_{it} represents the team leadership behavior or patient care dependent variable for leader *i* on patient *j*, ISS_{ij} is a control variable for the injury severity of the patient, and leader gender is a dummy-coded variable (men = 0 and women = 1). β_{01} is the term of principal interest, specifying the average difference between genders on the dependent variable. Statistical analyses were conducted in RStan, an R interface for Stan (Stan Development Team, 2019).[30]

Second, we selected diffuse normal distributions centered at 0 to represent our prior beliefs for each of the modeled parameters (ie, β's). The selection of this prior distribution meant that we did not privilege any credible hypothesis regarding the magnitude or direction of gender effects on leadership or patient care scores as more plausible (ie, the believability of hypotheses within the normally expected range for these gender effects were essentially equal). We chose this non-committal prior given that the previous work by Amacher et al and Meier et al offered conflicting results regarding the magnitude and direction of gender differences and neither study specifically evaluated team leadership during actual patient care.[3,4]

We fit our regression model to the data to compute the posterior distribution for each of the modeled parameters. The posterior distribution in Bayesian statistics summarizes the plausibility of all possible values for each modeled parameter given the observed data/findings and the prior beliefs (eg, how plausible is $\beta_{01} = .5$? $\beta_{01} = 1$? $\beta_{01} = -2$?). A highest density credibility interval (HDI) can be computed for this posterior distribution to summarize the range of most credible/believable estimates from the analysis. For the present analyses, the modal posterior parameter estimates (ie, the β_{01} with the highest plausibility) and 95% HDI for the effects of gender on each outcome are reported.

In addition to interpreting the 95% HDI of the posterior distributions, we also sought to examine the extent to which the magnitude and direction of the gender differences in leadership and patient care observed in our study were consistent with those reported by Amacher et al and Meier et al, respectively.[3,4] To do so we first transformed the effects reported by these authors into standardized effect sizes so that they could be meaningfully compared against our results.[31] For team leadership, we used the odds ratios reported by Amacher et al for the primary outcome to calculate a standardized effect size of a Cohen's $d = 0.58$ in favor of men.[3] For patient care, we used the odds ratios reported by Meier et al for 2 primary outcomes: likelihood of return of spontaneous circulation ($d = 0.17$) and survival to discharge ($d = 23$).[4] We averaged these findings to calculate and overall standardized effect size of a Cohen's $d = 0.20$ in favor of women. These computed effect sizes for leadership and patient care were subsequently translated into β estimates so that they could be placed on the same scale as the regression coefficients computed in our analyses and used for comparison.

Rather than compare the plausibility of a single possible effect size/point estimate for a hypothesis (ie, $\beta = .5$), it is common in Bayesian statistics to evaluate the plausibility of a *range* of plausible values that are, for all intents and purposes, equivalent (ie, any value for β between .4 and .6 is functionally the same as $\beta = .5$).[29] This is most easily accomplished by evaluating the extent to which the 95% HDI of a posterior distribution overlaps with a "region of practical equivalence" (ROPE) specifying a range of values that are practically indistinguishable from one another.

For the present analyses, we established a ROPE around the effect sizes for team leadership and patient care computed from Amacher et al and Meier et al that could be compared against the posterior HDI computed from our findings.[3,4] To inform our choices for the size of the ROPEs, we relied on Cohen's guidelines to identify a range of parameter values around both effect sizes that would differ by less than a conventionally small effect size (ie, the difference between the lower and upper limit of the ROPE corresponds with a small effect).[32]

Although the specific purpose for comparing the posterior HDI and ROPEs in our analyses was to compare our conclusions to those reported in the literature on gender differences in leadership and patient care, a unique value of Bayesian statistics is the potential to evaluate the extent to which one's results are consistent with *alternative* conclusions as well. That is, although we sought to evaluate the extent to which our findings support Amacher et al's conclusion of a male advantage in team leadership,[3] we can also examine the extent to which our results are instead supportive of a *female advantage* or *no gender differences* in leadership. Similarly, although we sought to evaluate the extent to which our results were consistent with Meier et al's conclusion of a female advantage in patient care,[4] we can examine the extent to which our findings are instead suggestive of a *male advantage* or *no gender differences* in this outcome. This is accomplished by simply "moving" the ROPEs, which serve as the point of comparison, such that they are centered on parameter estimates consistent with a particular conclusion. Consequently, we compared the posterior HDI computed for gender differences in both the leadership and clinical care metrics against ROPEs reflecting a male advantage, female advantage, or no difference in these outcomes.

The extent to which our data supported these different conclusions was assessed by examining the overlap between the posterior HDI and each corresponding ROPE. By convention, if the 95% HDI shares no overlap with the ROPE, the target value is rejected as a credible estimate and if the 95% HDI is completely contained within the ROPE, the target value is accepted as a credible estimate.[29] If the 95% HDI and ROPE partially overlap, there are insufficient data to determine whether the target value is a credible estimate.

3 | RESULTS

Team leader demographics by team leader gender are provided in Table 1. Patient and resuscitation characteristics by team leader gender are provided in Table 2. There were no significant differences in the leader, patient, or resuscitation characteristics between male and female team leaders, with the exception of patient ethnicity. Team leadership and patient care scores were weakly correlated (r = 0.22, 95% confidence interval, 0.4–0.38).

Correlations between other variables, including gender and ISS, were not significant. Descriptive statistics and correlations for the study variables are provided in Table 3.

TABLE 1 Characteristics of participants by gender

Team leader characteristic	Male (n = 40)	Female (n = 20)
Age, year; mean (SD)	30 (2.6)	29 (1.4)
Race, % (n)		
American Indian or Alaskan Native	0	0
Black or African American	2.5 (1)	5 (1)
Native Hawaiian or Other Pacific Islander	0	0
Asian	20 (8)	5 (1)
White	70 (28)	85 (17)
Other	7.5 (3)	5 (1)
Ethnicity, % (n)		
Hispanic or Latino	7.5 (3)	0
Not Hispanic or Latino	92.5 (37)	100 (20)
Residency year, % (n)		
Postgraduate training year 2	52.5 (21)	60 (12)
Postgraduate training year 3	47.5 (19)	40 (8)
Specialty, % (n)		
General surgery	25 (10)	25 (5)
Emergency medicine	75 (30)	75 (15)

SD, standard deviation.

3.1 | Parameter estimation: team leadership

The posterior estimates for the relationship between gender and team leadership are presented in Figure 1. The modal relationship between gender and team leadership was β_{01} = 0.94 (95% HDI, -.68 to 2.52). The intercept value was 7.14 (5.74, 8.50) and the ISS estimate was 0.03 (-.02, .08). The extent to which the observed relationship between gender and team leadership supported the effect reported by Amacher et al[3] was assessed by comparing the overlap between the posterior HDI with ROPEs reflecting a male-advantage effect (β_{01} = -2.63, lower bound = -3.54, upper bound = -1.72), a null effect (β_{01} = 0, lower bound = -0.91, upper bound = 0.91), and a female-advantage effect (β_{01} = 2.63, lower bound = 1.72, upper bound = 3.54). The present data do not support a male-advantage effect consistent with the size observed by Amacher et al.[3] The extent to which the observed results were more consistent with a null or female-advantage leadership effect is inconclusive.

3.2 | Parameter estimation: patient care

The posterior estimates for the relationship between gender and patient care are presented in Figure 2. The modal relationship between gender and patient care was β_{01} = 2.42 (95% HDI, -2.03 to 6.78). The intercept value was 59.9 (55.4, 64.2) and the ISS estimate was 0.13 (-0.04, 0.29). The extent to which the observed relationship between gender and patient care supported the effect reported by Meier et al[4] was assessed by comparing the overlap between the posterior HDI with ROPEs reflecting a male-advantage effect (β_{01} = -2.52, lower bound = -5.04, upper bound = 0), a null effect (β_{01} = 0, lower bound = -2.52, upper bound = 2.52), and a female-advantage effect (β_{01} = 2.52, lower bound = 0, upper bound = 5.04). The posterior HDI partially overlapped the ROPEs for all 3 comparisons, indicating that the observed effects fail to offer conclusive support for a male, female, or null advantage.

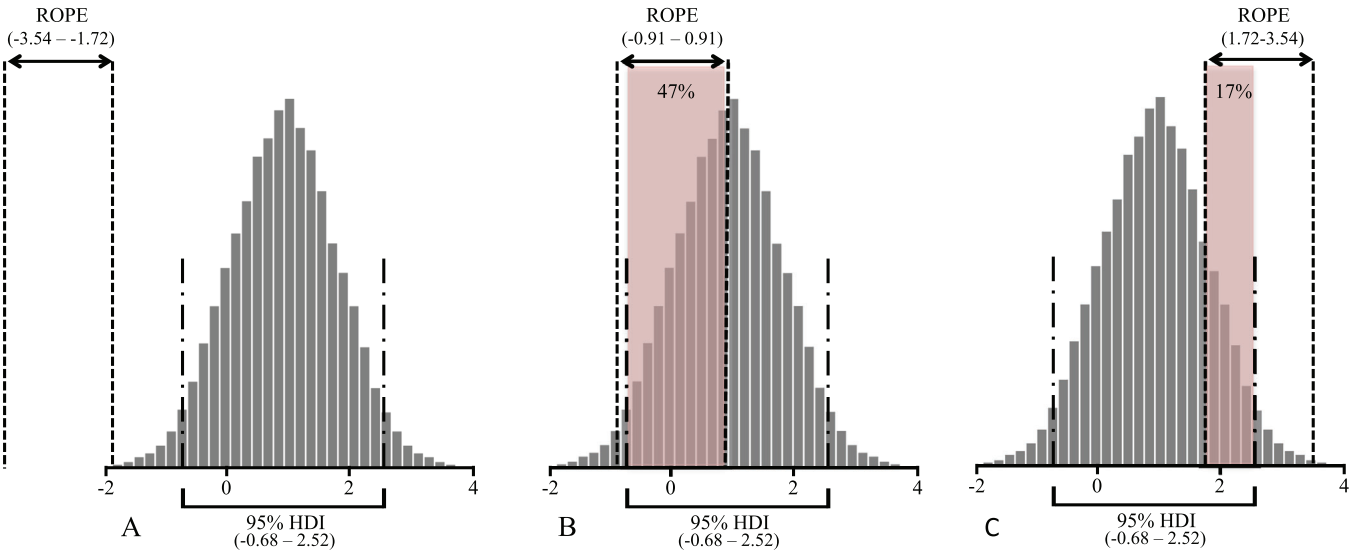

FIGURE 1 Posterior distribution of relationship between gender and team leadership behaviors comparing plausibility of male-advantage, null, and female-advantage effects. Each plot presents the same posterior distribution with a mean of 0.94 (95% HDI, -0.68 to 2.52). Plot (A) displays a ROPE centered on a male-advantage effect, plot (B) displays a ROPE centered on a null effect, and plot (C) displays a ROPE centered on a female-advantage effect. By convention, a model is rejected if 0% of the ROPE lies within the 95% HDI and is inconclusive if the ROPE partially overlaps the 95% HDI.[29] ROPE, region of practical equivalence; HDI, highest density interval

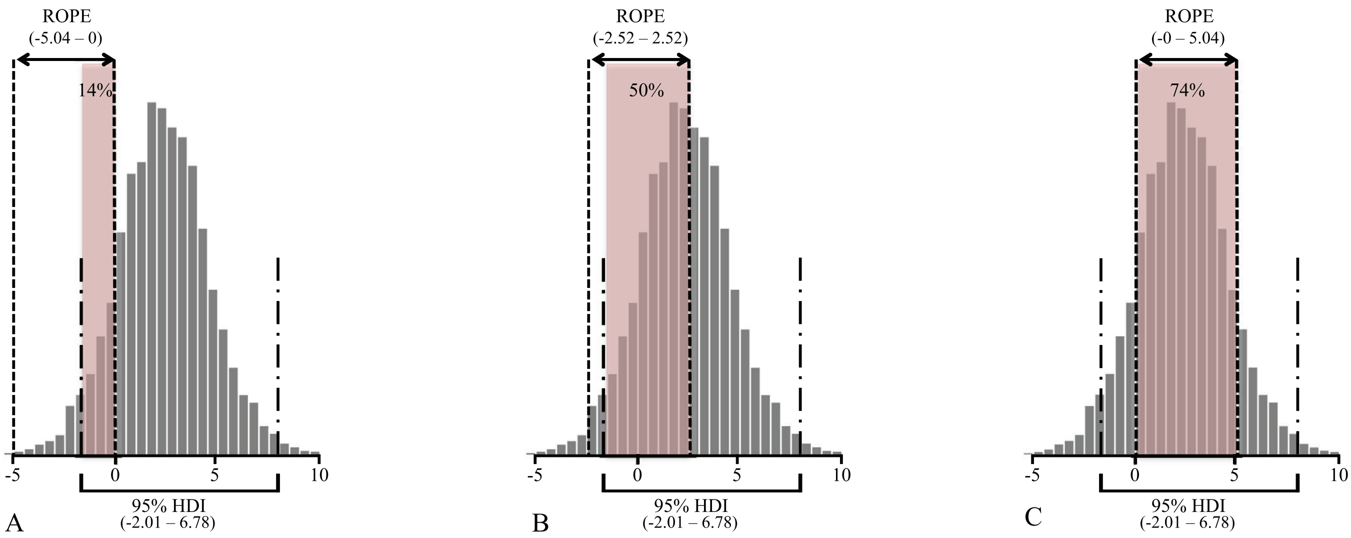

FIGURE 2 Posterior distribution of relationship between gender and patient care comparing male-advantage, null, and female-advantage effects. Each plot presents the same posterior distribution with a mean of 2.42 (95% HDI: -2.03 to 6.78). Plot (A) displays a ROPE centered on a male-advantage effect, plot (B) displays a ROPE centered on a null effect, and plot (C) displays a ROPE centered on a female-advantage effect. By convention, a model is inconclusive if the ROPE partially overlaps the 95% HDI.[29] ROPE, region of practical equivalence; HDI, highest density interval

The parameter estimates presented here are informed by previously published work; however, a summary table of potential credible parameter estimates for both sets of analyses is provided in Table S1. Parameter estimates for the 2 constructs are strongly convergent.

3.3 | Limitations

Our study has several limitations. First, team leaders were assessed in the clinical setting within actual health care teams. Although this increases the relevance of the findings, it also introduces team variability. The resuscitation teams varied in size and composition, factors we were unable to include in statistical models. Furthermore, team interdependence makes it difficult to isolate an individual's performance. We included 2 observations for each team leader to mitigate this threat; however, we cannot exclude team factors as a potential influence on the team leader. Another potential limitation is the use of novel team leadership and patient care assessment measures. The development and application of these measures provide supporting evidence of validity specific to this study setting; however. they have not been used in other contexts. Finally, this study was conducted at a single academic institution using resident team leaders in trauma resuscitations. As a result, these findings may not be generalizable to more experienced physicians, different practice sites, or different types of resuscitations.

There are also limitations specific to Bayesian analyses. For parameter estimation, using priors that inadequately represent the phenomenon under consideration can result in misleading inferences.[33] We used diffuse uninformed priors instead of potentially flawed priors from the extant literature. Doing so ensured that the posterior distribution for our parameter estimates were primarily informed by our observed data, rather than building on an assumption of a gender-based advantage to team leadership. Although this practice is reasonable when a study is the first of its kind, future research should use the present results to inform prior distributions involving gender effects on leadership and patient care behaviors. Finally, the sample size for the present study was relatively small. The HDI was wider than the ROPEs used in our parameter estimation, which precluded us from accepting a parameter estimate as reasonable, beyond rejecting a strong male-advantage effect.

4 | DISCUSSION

Prior work suggesting inherent gender-based differences in team leadership skills was controversial.[3,4] Amacher et al demonstrated that within medical student teams, male team leaders made more "secure" team leadership statements and their teams maintained more hands-on time during the first 3 minutes of cardiopulmonary resuscitation.[3] However, their work was conducted in a simulated setting with novice participants. In contrast, Meier et al reviewed 1082 in-hospital cardiac arrests and found that resuscitations led by females had higher rates of return of spontaneous circulation and survival to discharge.[4] The work by Meier et al found a difference in patient outcomes but not in the quality of the cardiopulmonary resuscitation delivered (eg, compression rate and depth), leaving the mechanism for the difference in clinical outcomes unknown.

In our study we address several limitations present in prior work. We prospectively assessed the performance of designated team leaders during 2 unique clinical events using metrics that focus on behaviors rather than on leadership style. Furthermore, we measured both performance (patient care) and process (team leadership)

JACEP OPEN
JOURNAL OF THE AMERICAN COLLEGE OF EMERGENCY PHYSICIANS OPEN
A GLOBAL JOURNAL OF EMERGENCY MEDICINE
Open Access
WILEY

TABLE 2 Patient and resuscitation characteristics by gender of team leader

Patient characteristics	Resuscitation with male team leader (n = 77)[a]	Resuscitation with female team leader (n = 41)[a]
Patient gender, % (n)		
Male	84.4 (65)	80.5 (33)
Female	15.6 (12)	19.5 (8)
Patient age, mean (SD), years	42 (17.4)	46 (18)
Patient race, % (n)		
White	81.8 (63)	68.3 (28)
Black	7.8 (6)	14.6 (6)
Asian	2.6 (2)	7.3 (3)
Pacific Islander/Native Hawaiian	1.3 (1)	0 (0)
Native American	1.3 (1)	9.8 (4)
Other or not identified	5.2 (4)	0 (0)
Patient ethnicity,[b] % (n)		
Hispanic	16.9 (13)	4.9 (2)
Non-Hispanic	77.9 (60)	95.1 (39)
Not reported	5.2 (4)	0 (0)
Injury severity score[c], mean (SD)	20.4 (14.2)	22.6 (15.8)
Trauma team activation level[d], % (n)		
Full	58.4 (45)	63.4 (26)
Modified	41.6 (32)	36.6 (15)
Cause of trauma, % (n)		
Blunt	72.7 (56)	75.6 (31)
Penetrating	27.3 (21)	24.4 (10)
Primary transport mode[e], % (n)		
Ground transport	56.6 (43)	65.9 (27)
Aeromedical transport	42.2 (32)	24.2 (14)
Self-presentation	1.3 (1)	0 (0)
Type of response[f], % (n)		
Transfer	37.3 (29)	36.6 (15)
Field	62.3 (48)	63.4 (26)

SD, standard deviation.
[a]N = 118, patient and resuscitation characteristic data missing for 2 observations.
[b]Significant difference between groups ($P = 0.047$).
[c]Baker et al.[14].
[d]As per trauma activation criteria[13].
[e]N = 117, 1 event had no arrival mode reported.
[f]Transfer patients arrived from another healthcare facility, whereas field responses did not receive care at another facility before arrival.

TABLE 3 Descriptive statistics and correlations of study variables, including patient care and leadership performance

Variable	*M*	*SD*	1	2	3
1. Gender	0.33	0.47			
2. ISS	21.01	14.71	.09		
			[-.10, .26]		
3. Patient care score	63.34	12.88	.10	.15	
			[-.08, .27]	[-.03, .33]	
4. Leadership score	8.10	3.98	.13	.12	.22*
			[-.05, .30]	[-.07, .29]	[.04, .38]

ISS, injury severity score; M, mean; SD, standard deviation.
Values in square brackets indicate the 95% confidence interval for each correlation.
*Indicates $P < 0.05$.

simultaneously as recommended in the team science literature.[34] Our results do not support previously reported gender-based advantage to team leadership.

Our work adds to the body of literature examining the role of gender in resuscitation team leadership and clinical performance; however, we would suggest that ongoing attempts to identify gender-based superiority or inferiority in team leadership are misplaced. Effective team leadership consists of a set of discrete behaviors that can be defined, trained, and assessed. Thus we agree with Meier et al that all appropriately trained physicians, regardless of gender, can provide high-quality resuscitation care.[4]

Studies in other domains report gender-based differences in leadership style that affect leadership emergence in various contexts, with male leaders favored in task-related events.[35] However, leadership style does not equate with leadership effectiveness[36] and current approaches to evaluating team leadership may be contributing to the reported gender gap in performance. Although society and perceptions have evolved over time, gender-based expectations still exist.[37] In health care, effective resuscitation team leadership is often described as "assertive," "dominant," or "directive," and residents from both emergency medicine and internal medicine have described challenges faced by female team leaders attempting to fit a prescribed leadership style.[38,39]

Team leadership is a complex construct and effective team leaders do more than just delegate and command. It is crucial that training initiatives and assessments account for the numerous ways in which a team leader can promote team performance. This includes supporting the attitudes, behaviors, and cognition of the team.[40] A rigorous evaluation of trauma teams characterized leadership as (1) contingent, depending on the needs of team; (2) functional, with the team leader picking up tasks or roles to ensure the job is done; (3) flexible, adapting to changing team and patient conditions; and (4) shared, allowing for emergence of leadership behaviors from other team members.[41] Although this work was specific to trauma teams, it may be applicable in other resuscitation teams that face similar challenges, such as variability in team composition. Regardless of team type, leadership assessments that rely on leadership style can contaminate outcomes when evaluating gender-based performance, such as the focus on directive leadership leading to the conclusion of "inferior female leadership" in work by Amacher et al.[3] To advance resuscitation team leadership practices we must focus on training and assessing effective leadership skills in an equitable manner.

In conclusion, our prospective evaluation of team leadership and patient care during trauma resuscitations does not support a strong gender-based advantage in resuscitation team leadership. Team leadership is a complex construct, and it is crucial that training initiatives and assessments account for the numerous ways in which a team leader can promote team performance. Focusing on the behaviors, not inherent traits,[42] that constitute effective team leadership will allow resuscitation leaders to optimize their individual and team performance, irrespective of gender.

ACKNOWLEDGMENTS

The authors thank Fred Oswald, PhD for his review of the Bayes analyses and Jessica Webb, PhD for her assistance with video coding and rater training. Also Joseph Shuluk for his assistance with data management; Joshua Stanfield, Sayward Nelson, Brian Le, and Senita Lavulavu for assistance with video processing; and our video coders (Athena Golematis, Kelsey Phillips, Katherine Bouma, Michael Hinnawi, Brandt Gruizinga, Jack Brodeur, Nicole Freyholtz, and Callum McCulloch). Finally, the authors thank the Harborview Medical Center Emergency Department, the Trauma Surgery service, and the numerous other clinical services involved in trauma care, for facilitating this research.

CONFLICTS OF INTEREST

EDR reports grant funding from the Agency for Healthcare Research and Quality (1R18HS022458-01A1) and the Department of Defense (W81XWH-18-1-0089). GTC reports grant funding from the Army Research Institute for the Behavioral and Social Sciences (W911NF-14-1-0026) and the National Science Foundation. SWJK reports grant funding from the Army Research Institute for the Behavioral and Social Sciences (W911NF-14-1-0026), the National Aeronautics and Space Administration (NNX16AR52G), the National Science Foundation (1533947), and honoraria from the American Psychological Association; Organizational Psychology; Organizational Science, Translation, and Application. JAG reports grant funding from the Department of Defense (W81XWH-18-1-0089) and the Army Research Institute for the Behavioral and Social Sciences (W911NF-14-1-0026). RF reports grant funding from the Agency for Healthcare Research and Quality and the Department of Defense (W81XWH-18-1-0089). AM, JO, AKC, MCV, and SB have no conflicts of interest to report.

AUTHOR CONTRIBUTIONS

EDR, JAG, and RF conceived of the study concept and design. RF obtained research funding. EDR, AM, JO, SMB, AKC, MCV, and RF acquired the data, with AM, GTC, and SWJK providing statistical expertise. EDR, AM, JO, SMB, GTC, SWJK, JAG, RF performed the analysis and interpretation of the data. EDR, AM, and JAG drafted the article. EDR, AM, JO, SMB, AKC. MCV, GTC, SWJK, and RF all performed critical revision of the article for important intellectual content.

REFERENCES

1. Tsugawa Y, Jena AB, Figueroa JF, et al. Comparison of hospital mortality and readmission rates for Medicare patients treated by male vs female physicians. *JAMA Intern Med*. 2017;177(2):206-213.
2. Wallis CJ, Ravi B, Coburn N, et al. Comparison of postoperative outcomes among patients treated by male and female surgeons: a population based matched cohort study. *BMJ*. 2017;359:j4366.
3. Amacher SA, Schumacher C, Legeret C, et al. Influence of gender on the performance of cardiopulmonary rescue teams: a randomized, prospective simulator study. *Crit Care Med*. 2017;45(7):1184-1191.
4. Meier A, Yang J, Liu J, et al. Female physician leadership during cardiopulmonary resuscitation is associated with improved patient outcomes. *Crit Care Med*. 2019;47(1):e8-e13.
5. Fernandez R, Rosenman ED, Olenick J, et al. Simulation-based team leadership training improves team leadership during actual trauma resuscitations: a randomized controlled trial. *Crit Care Med*. 2019;48(1):73-82.
6. Yeung JH, Ong G, Davies RP, et al. Factors affecting team leadership skills and their relationship with quality of cardiopulmonary resuscitation. *Crit Care Med*. 2012;40(9):2617-2621.
7. Parker SH, Yule S, Flin R, et al. Surgeons' leadership in the operating room: an observational study. *Am J Surg*. 2012;204(3):347-354.
8. Klein R, Ufere NN, Rao SR, et al. Association of gender with learner assessment in graduate medical education. *JAMA Netw Open*. 2020;3(7):e2010888-e2010888.
9. Klein R, Julian KA, Snyder ED, et al. Gender bias in resident assessment in graduate medical education: review of the literature. *J Gen Intern Med*. 2019;34(5):712-719.
10. Rojek AE, Khanna R, Yim JW, et al. Differences in narrative language in evaluations of medical students by gender and under-represented minority status. *J Gen Intern Med*. 2019;34(5):684-691.
11. Axelson RD, Solow CM, Ferguson KJ, Cohen MB. Assessing implicit gender bias in medical student performance evaluations. *Eval Health Prof*. 2010;33(3):365-385.
12. Ju M, van Schaik SM. Effect of professional background and gender on residents' perceptions of leadership. *Acad Med*. 2019;94(11S):S42-7.
13. Trauma Clinical Guideline: Trauma Team Activation Criteria. 2016. https://www.uwmedicine.org/provider-resource/protocols/trauma-team-activation-criteria. Washington State Department of Health Office of Community Health Systems Emergency Medical Services and Trauma Section. Updated July, 2016. Accessed February, 2020.
14. Baker SP, O'Neill B, Haddon Jr W, et al. The injury severity score: a method for describing patients with multiple injuries and evaluating emergency care. *J Trauma Acute Care Surg*. 1974;14(3):187-196.
15. Cook DA, Beckman TJ. Current concepts in validity and reliability for psychometric instruments: theory and application. *Am J Med*. 2006;119(2):166. e167-166. e116.
16. Fernandez R, Rosenman ED, Brolliar S, et al. An Event-based Approach to Measurement: facilitating observational measurement in highly variable clinical settings. *Acad Emerg Med Educ Train*. 2020;4(2):147-153.
17. Kozlowski SW, Gully SM, McHugh PP, Salas E, Cannon-Bowers JA. A dynamic theory of leadership and team effectiveness: developmental and task contingent leader roles. In: Ferris GR, ed. *Research in Personnel and Human Resources Management*. Greenwich, CT: JAI Press; 1996:253-306.
18. Rosenman ED, Shandro JR, Ilgen JS, et al. Leadership training in health care action teams: a systematic review. *Acad Med*. 2014;89(9):1295-1306.
19. Rosenman ED, Ilgen JS, Shandro JR, et al. A systematic review of tools used to assess team leadership in health care action teams. *Acad Med*. 2015;90(10):1408-1422.
20. American College of Surgeons. *Advanced Trauma Life Support for Doctors: Student Course Manual*. 10th ed. Chicago, IL: American College of Surgeons; 2018.
21. Holcomb JB, Dumire RD, Crommett JW, et al. Evaluation of trauma team performance using an advanced human patient simulator for resuscitation training. *J Trauma Acute Care Surg*. 2002;52(6):1078-1086.

22. Lubbert PH, Kaasschieter EG, Hoorntje LE, et al. Video registration of trauma team performance in the emergency department: the results of a 2-year analysis in a Level 1 trauma center. *J Trauma Acute Care Surg*. 2009;67(6):1412-1420.
23. Ritchie PD, Cameron PA. An evaluation of trauma team leader performance by video recording. *Aust N Z J Surg*. 1999;69(3):183-186.
24. Sugrue M, Seger M, Kerridge R, et al. A prospective study of the performance of the trauma team leader. *J Trauma Acute Care Surg*. 1995;38(1):79-82.
25. Kelleher DC, Bose RJC, Waterhouse LJ, et al. Effect of a checklist on advanced trauma life support workflow deviations during trauma resuscitations without pre- arrival notification. *J Am Coll Surg*. 2014;218(3):459-466.
26. Byrt T, Bishop J, Carlin JB. Bias, prevalence and kappa. *J Clin Epidemiol*. 1993;46(5):423-429.
27. Kalil AC, Sun J. Bayesian methodology for the design and interpretation of clinical trials in critical care medicine: a primer for clinicians. *Crit Care Med*. 2014;42(10):2267-2277.
28. Lewis RJ, Wears RL. An introduction to the Bayesian analysis of clinical trials. *Ann Emerg Med*. 1993;22(8):1328-1336.
29. Kruschke J. *Doing Bayesian Data Analysis: A tutorial with R, JAGS, and Stan*. 2nd ed. Burlington, MA: Academic Press; 2015.
30. Stan Development Team. http://mc-stan.org. RStan: the R interface to Stan. R package version 2.19.2. Accessed September 12, 2019.
31. Sánchez-Meca J, Marín-Martínez F, Chacón-Moscoso S. Effect-size indices for dichotomized outcomes in meta-analysis. *Psychol Methods*. 2003;8(4):448.
32. Cohen J. A power primer. *Psychol Bull*. 1992;112(1):155.
33. Quintana M, Viele K, Lewis RJ. Bayesian analysis: using prior information to interpret the results of clinical trials. *JAMA*. 2017;318(16):1605-1606.
34. Grand JA, Pearce M, Rench TA, et al. Going DEEP: guidelines for building simulation-based team assessments. *BMJ Qual Saf*. 2013;22(5):436-448.
35. Eagly AH, Johnson BT. Gender and leadership style: a meta-analysis. *Psychol Bull*. 1990;108(2):233.
36. Day D. Leadership. *Oxford Handbook of Organizational Psychology. Kozlowski SWJ*. New York, NY: Oxford University Press; 2012:696-729.
37. Badura KL, Grijalva E, Newman DA, et al. Gender and leadership emergence: a meta-analysis and explanatory model. *Pers Psychol*. 2018;71(3):335-367.
38. Kolehmainen C, Brennan M, Filut A, et al. Afraid of Being Witchy with a 'B'": a qualitative study of how gender influences residents' experiences leading cardiopulmonary resuscitation. *Acad Med*. 2014;89(9):1276.
39. Linden JA, Breaud AH, Mathews J, et al. The intersection of gender and resuscitation leadership experience in emergency medicine residents: a qualitative study. *Acad Emerg Med Educ Train*. 2018;2(2):162-168.
40. Rosenman ED, Fernandez R, Wong AH, et al. Changing systems through effective teams: a role for simulation. *Acad Emerg Med*. 2018;25(2):128-143.
41. Klein KJ, Ziegert JC, Knight AP, et al. Dynamic delegation: shared, hierarchical, and deindividualized leadership in extreme action teams. *Adm Sci Q*. 2006;51(4):590-621.
42. Neubert MJ, Taggar S. Pathways to informal leadership: the moderating role of gender on the relationship of individual differences and team member network centrality to informal leadership emergence. *Leadersh Q*. 2004;15(2):175-194.

AUTHOR BIOGRAPHY

Rosemarie Fernandez, MD, is an Associate Professor in emergency medicine at the University of Florida.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Rosenman ED, Misisco A, Olenick J, et al. Does team leader gender matter? A Bayesian reconciliation of leadership and patient care during trauma resuscitations. *JACEP Open*. 2021;2:e12348. https://doi.org/10.1002/emp2.12348