BRIEF CONTRIBUTION

An Investigation of the Relationship Between Emergency Medicine Trainee Burnout and Clinical Performance in a High-fidelity Simulation Environment

Dave W. Lu, MD, MSCI, MBE, Scott M. Dresden, MD, MS, D. Mark Courtney, MD, MSCI, and David H. Salzman, MD, MEd

ABSTRACT

Objective: Burnout is prevalent among emergency medicine (EM) physicians, with physicians experiencing burnout more likely to report committing medical errors or delivering suboptimal care. The relationship between physician burnout and identifiable differences in clinical care, however, remains unclear. We examined if EM trainee burnout was associated with differences in clinical performance using high-fidelity simulation as a proxy for patient care.

Methods: In this cross-sectional study across six institutions, we measured trainee performance over four simulation scenarios based on recognized EM milestones. For each scenario a faculty rater assessed whether the trainee performed predefined critical actions specific to each case. A summation of performed actions across all cases resulted in a cumulative task (CT) score (range = 0–85). Raters also assigned an impression score on a 10-point scale (0 = poor; 10 = outstanding) assessing the trainee's overall performance after each scenario, with the mean of the scores resulting in an overall impression (OI) score. After the simulation assessment, we measured trainees' burnout via the Maslach Burnout Inventory through a confidential, electronic survey. Trainee depression, quality of life (QOL) and daytime sleepiness were also evaluated. Survey results were compared to simulation scores using analysis of variance and covariance.

Results: Fifty-eight of 89 (65.2%) eligible participants completed the survey and simulation assessment. Thirty-one of 58 (53.4%, 95% CI = 40.2% to 66.7%) trainees reported burnout. In trainees with burnout compared to those without, mean CT scores (73.4 vs. 75.2, 95% CI of difference = 0.06 to 3.51) and OI scores (6.4 vs 6.8, 95% CI of difference = 0.03 to 0.79) were negatively associated with burnout after controlling for training program. In contrast, QOL were positively associated with CT [$F(1,48) = 4.796$, $p = 0.033$] and OI [$F(1,48) = 4.561$, $p = 0.038$] scores. There were no significant associations between simulation performance and depression or daytime sleepiness.

Conclusion: Emergency medicine trainees with burnout received lower cumulative performance scores over four high-fidelity simulation scenarios than trainees without burnout.

From the Department of Emergency Medicine (DWL, SMD, MC, DHS) and the Department of Medical Education (DHS), Northwestern University Feinberg School of Medicine, Chicago, IL. Dr. Lu is currently with the Department of Emergency Medicine, Tufts University School of Medicine–Maine Medical Center, Portland, ME.
Received August 1, 2016; revision received September 30, 2016; accepted October 2, 2016.
This study was supported by the Northwestern Memorial Foundation Alvin H. Baum Clinical Simulation Research Grants Initiative.
The authors have no potential conflicts to disclose.
Supervising Editor: Wendy C. Coates, MD.
Address for correspondence and reprints: Dave W. Lu, MD, MSCI, MBE; e-mail: dlu@mmc.org.
AEM EDUCATION AND TRAINING 2017;1:55–59.

doi: 10.1002/aet2.10004

Burnout is a syndrome of depersonalization, emotional exhaustion, and low personal accomplishment.[1] Burnout is common among physicians, with emergency medicine (EM) physicians reporting the highest levels among all specialties.[1] Burnout negatively affects providers and may be associated with poor patient care.[2,3] Much of the research on burnout and patient care has been limited to physicians' self-reports of suboptimal care.[2] It remains unclear if burnout is associated with identifiable differences in clinical care.[3] To investigate this question, our study examined the relationship between burnout and the clinical performance of EM trainees in a high-fidelity simulation environment. We hypothesized that trainees with burnout will have lower performance scores than trainees without burnout.

METHODS

Study Design

Trainee simulation performance scores were compared to their burnout survey results after Northwestern University Human Subjects Review Board approval.

Study Setting and Population

We invited a convenience sample of EM postgraduate year-2 trainees from six Chicago programs to this study in February 2015.

Study Protocol

Simulation cases and assessment tools measuring the completion of technical (e.g., medication administration) and nontechnical skills (e.g., patient communication) were designed by expert faculty based on recognized EM milestones.[4] Cases included placement of a central venous catheter, performance of a lumbar puncture, and two immersive scenarios involving cardiopulmonary chief complaints. Cases and simulation programming (Laerdal, www.laerdal.com/us) were standardized such that participants received identical information.

Trained faculty raters assessed whether trainees performed or failed to perform predefined critical actions for each case. A summation of performed actions across all four cases resulted in a cumulative task (CT) score (range = 0 to 85). Faculty also assigned an impression score on a 10-point scale (0 = poor performance; 10 = outstanding performance) assessing trainees' overall performance at the end of each case. The mean of the four impression scores resulted in an overall impression (OI) score for each trainee. To minimize bias, faculty assessed trainees from different programs. Assessments were completed electronically using Qualtrics (www.qualtrics.com).

Trainees who completed the simulation assessment received e-mail invitations to a confidential electronic survey assessing burnout and wellness. Invitations did not mention burnout and trainees were blinded to study hypotheses. Trainees consented by completing the open survey on a secure and Web-based application (Research Electronic Data Capture). Participants were provided $50 gift cards after survey completion. Nonresponders received up to two reminder e-mails.

Measurements

We assessed burnout using the Maslach Burnout Inventory, with burnout defined by high depersonalization or emotional exhaustion subscale scores.[1] Secondary measures examining other aspects of wellness were quality of life (QOL), depression, and daytime sleepiness. QOL was measured by a 0–10 linear analog scale assessment, with higher scores indicating higher QOL. Depression was screened with the Patient Health Questionnaire-2.[5] Daytime sleepiness was measured using the Epworth Sleepiness Scale (ESS; range = 0 to 24), with higher scores indicating greater daytime sleepiness.[6] Limited demographic information was collected to encourage participation and honest reporting.

Data Analysis

Descriptive statistics were calculated and trainee simulation scores were tested for normality via the Shapiro-Wilk test. Differences in simulation scores among training programs were evaluated using analysis of variance (ANOVA). For our primary analysis two groups were created based on the presence or absence of burnout, after which CT and OI scores adjusted by program were compared between trainees with and without burnout using ANOVA. Secondary analyses compared simulation scores with depression via ANOVA and with QOL and daytime sleepiness via analysis of covariance. A t-test compared trainees' mean ESS score to the

population mean.[6] F-scores (degrees of freedom, number of observations) and p-values are reported. Analyses were performed using SPSS v23.0.

RESULTS

Seventy-seven of 89 (86.5%) eligible trainees completed the simulation assessment, with 58 of 77 (75.3%) responding to the survey, for an overall participation rate of 65.2%. Mean age was 29.0 years and the majority was male (60.3%).

Simulation Performance

Trainees' CT and OI scores tested for normality (0.978, df = 58, p = 0.385; 0.968, df = 58, p = 0.129, respectively). The mean CT score was 74.0 (95% confidence interval [CI] = 72.8 to 75.1) and the mean OI score was 6.6 (95% CI = 6.3 to 6.9). The mean CT and OI scores for trainees who did not participate in the survey portion of the study were 75.5 (95% CI = 73.4 to 77.6) and 6.8 (95% CI = 6.3 to 7.3), respectively. There were no significant differences in the mean CT score (difference in mean = 1.5, 95% CI = –0.8 to 3.8, p = 0.201) and the mean OI score (difference in mean = 0.2, 95% CI = –0.4 to 0.8, p = 0.505) between trainees who did and did not participate in the survey portion of the study. There were, however, significant differences in mean CT scores (range = 70.0 to 79.1) and OI scores (range = 5.5 to 8.4) among the six training programs [F(5,52) = 7.248, p < 0.001; F(5,52) = 15.530, p < 0.001, respectively].

Wellness Measures

Trainees reported a burnout rate of 53.4% (95% CI = 40.2% to 66.7%), with no significant difference between sex (p = 0.985) or programs (p = 0.845). Mean trainee QOL was 6.5 (95% CI = 5.9 to 7.1). Less than half (44.6%, 95% CI = 31.2% to 58.1%) of trainees screened positive for depression. Trainees' mean ESS score was 9.6 (95% CI = 8.6 to 10.7), which is higher than the reference population mean of 4.6 (95% CI of difference = 3.8 to 6.2, p < 0.001).[6]

Comparison of Simulation Performance to Wellness Measures

After training program was controlled for due to significant differences in trainees' simulation scores among programs, there were no significant differences in scores by scenario between trainees with and without burnout (Table 1). When scores from all four scenarios were combined into CT and OI scores, however, burnout was associated with lower CT (73.4 vs. 75.2, 95% CI of difference = 0.06 to 3.51, p = 0.043) and OI scores (6.4 vs. 6.8, 95% CI of difference = 0.03 to 0.79, p = 0.035). There was also a positive association between QOL and CT [F(1,48) = 4.796, p = 0.033] and OI [F(1,48) = 4.561, p = 0.038] scores. We did not find significant associations between simulation scores and a positive depression screen (CT score, 73.3 vs. 74.9, 95% CI of difference = –0.46 to 3.71, p = 0.123; OI score, 6.4 vs. 6.8, 95% CI of difference = –0.08 to 0.76, p = 0.107) or ESS scores [CT score, F(1,50) = 0.120, p = 0.730; OI score, F(1,50) = 0.147, p = 0.703].

DISCUSSION

We believe that this is the first study to examine the relationship between physician burnout and identifiable differences in clinical performance in a high-fidelity simulation environment, which has emerged as a tool to assess clinician performance.[7] Trainees in our study reported high levels of burnout as early as the second year of residency. This rate is consistent with prior studies[1,8] and raises questions of how trainees will sustain their medical careers in light of increasing work pressures and a predicted physician shortage.[9] While there exists motivation to address burnout to improve individual provider wellness,

Table 1
Simulation Scores by Burnout Versus No Burnout

Simulation Score	Burnout	No Burnout	p-value
Clinical scenario 1			
Task score (0–22)	19.2 (±2.32)	20.0 (±1.54)	0.238
Impression score (0–10)	6.5 (±1.18)	7.0 (±1.74)	0.364
Clinical scenario 2			
Task score (0–15)	11.9 (±1.99)	12.6 (±1.60)	0.386
Impression score (0–10)	6.5 (±1.69)	6.8 (±1.74)	0.215
Procedural scenario 1			
Task score (0–25)	20.0 (±1.78)	20.7 (±1.49)	0.306
Impression score (0–10)	6.0 (±1.92)	7.0 (±1.21)	0.065
Procedural scenario 2			
Task score (0–23)	21.4 (±2.41)	22.0 (±2.77)	0.240
Impression score (0–10)	6.3 (±2.01)	6.6 (±2.14)	0.728
Cumulative task score (0–85)	72.9 (±4.45)	75.2 (±3.97)	0.043
Overall impression score (0–10)	6.3 (±1.09)	6.8 (±1.19)	0.035

Data are reported as mean (±SD).

our study aimed to determine if another incentive to ameliorate burnout may be its potential influence on clinical performance.

In our study burned out trainees were more likely to perform fewer critical tasks than non–burned out trainees cumulatively over four simulation scenarios. Burned out trainees were also rated as performing more poorly in general (e.g., professionalism). However, there were no significant differences between burned out and non–burned out trainees for each individual scenario. The clinical significance of the small absolute differences in cumulative scores between the two groups also remains uncertain. Our findings, nonetheless, appear to be consistent with prior work demonstrating an association between burnout and physician self-reported error and suboptimal care.[2] Few studies have examined the relationship between provider burnout and real-world patient care with conflicting results.[10,11] Our study used simulation as a proxy for patient care and was unable to resolve this issue.

Although prior work found a significant relationship between trainee depression and medication errors,[11] we did not find a similar association between depression and simulation performance. This discrepancy may be due to different screening instruments and our evaluation of simulation performance involving multiple tasks rather than specifically medication errors. We also did not find a significant association between daytime sleepiness and simulation performance, despite trainees reporting significantly higher ESS scores than the general population. This suggests that work-hour restrictions focused on time off may not improve burnout or trainee clinical performance.[12] In contrast, trainees reporting higher QOL were more likely to have higher simulation scores. We suspect a higher QOL may ameliorate some of the effects of burnout.

LIMITATIONS

Our results from a small sample of Chicago-based trainees studied during winter may not be generalizable to physicians in other environments. The impact of our survey response rate on outcomes is unknown, since we were unable to compare burnout between respondents and nonrespondents. The simulation cases we used have not been rigorously validated, and there were no second raters to determine inter-rater reliability. It also remains unclear if burnout led to poor simulation performances or if poor simulation performances led to burnout. Finally, we were unable to determine trainees' baseline knowledge and skills prior to simulation assessment.

CONCLUSIONS

Emergency medicine trainees with burnout received lower cumulative performance scores over four simulation scenarios than trainees without burnout. Future work to elucidate the relationship between physician burnout and clinical performance is needed.

The authors thank Alyssa Czerniak for her assistance with data collection.

References

1. Shanafelt TD, Boone S, Tan L, et al. Burnout and satisfaction with work-life balance among US physicians relative to the general US population. Arch Intern Med 2012;172:1377–85.
2. Williams ES, Manwell LB, Konrad TR, Linzer M. The relationship of organizational culture, stress, satisfaction, and burnout with physician-reported error and suboptimal patient care: results from the MEMO study. Health Care Manage Rev 2007;32:203–12.
3. Halbesleben JR, Rathert C. Linking physician burnout and patient outcomes: exploring the dyadic relationship between physicians and patients. Health Care Manage Rev 2008;33:29–39.
4. The Emergency Medicine Milestones Project. Available at: http://www.acgme.org/Portals/0/PDFs/Milestones/EmergencyMedicineMilestones.pdf. Accessed Sep 30, 2016.
5. Dyrbye LN, West CP, Satele D, et al. Burnout among U.S. medical students, residents, and early career physicians relative to the general U.S. population. Acad Med 2014;89:443–51.
6. Johns M, Hocking B. Daytime sleepiness and sleep habits of Australian workers. Sleep 1997;20:844–9.
7. Gordon JA, Alexander EK, Lockley SW, et al. Does simulator-based clinical performance correlate with actual hospital behavior? The effect of extended work hours on patient care provided by medical interns. Acad Med 2010;85:1583–8.
8. Kimo Takayesu J, Ramoska EA, Clark TR, et al. Factors associated with burnout during emergency medicine residency. Acad Emerg Med 2014; 21:1031–5.

9. Dewa CS, Jacobs P, Thanh NX, Loong D. An estimate of the cost of burnout on early retirement and reduction in clinical hours of practicing physicians in Canada. BMC Health Serv Res 2014;14:254.
10. Linzer M, Manwell LB, Williams ES, et al. Working conditions in primary care: physician reactions and care quality. Ann Intern Med 2009;151(28–36): W6–9.
11. Fahrenkopf AM, Sectish TC, Barger LK, et al. Rates of medication errors among depressed and burnt out residents: prospective cohort study. BMJ 2008;336:488–91.
12. Ahmed N, Devitt KS, Keshet I, et al. A systematic review of the effects of resident duty hour restrictions in surgery: impact on resident wellness, training, and patient outcomes. Ann Surg 2014;259:1041–53.