

Impact of autonomic regulation on burnout and performance in thoracic surgery residents



Lauren K. Barron, MD,^a and Marc R. Moon, MD^b

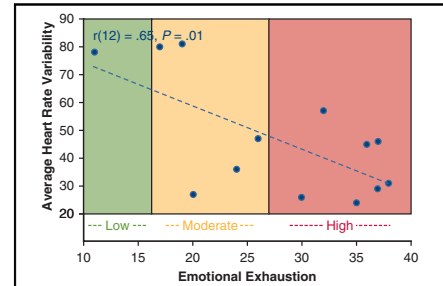
ABSTRACT

Objective: This study sought to determine the feasibility of collecting physiologic data in thoracic surgery residents and whether it would correlate with burnout and burnout with performance.

Methods: This was a prospective study of thoracic surgery residents over a 5-month period. Participants were evaluated with a wearable biometric device (heart rate variability and sleep) and the Maslach Burnout Inventory. Resident performance was quantified using Accreditation Council for Graduate Medical Education Milestones (scale, 1-5) normalized to program-designated targets (3 for postgraduate year 6 or lower residents and 4 for postgraduate year 7 residents).

Results: The cohort consisted of 71% female participants (5/7) with 86% of residents having 1 or more children. High levels of emotional exhaustion (median, 30 [interquartile range, 20-36], where >26 is high) and high levels of depersonalization (median, 16 [interquartile range, 14-22], where >12 is high) were common, but personal accomplishment was also uniformly high (median, 43 [interquartile range, 41-46], where >38 is high). There was a significant correlation between heart rate variability and emotional exhaustion ($r(12) = 0.65, P = .01$) but not depersonalization ($P = .28$) or personal accomplishment ($P = .24$). Depersonalization and personal accomplishment did not correlate with resident performance ($P = .12$ and $P = .75$, respectively); however, increased emotional exhaustion showed a significant correlation with higher resident performance during periods when burnout was reported ($r(6) = 0.76, P = .047$).

Conclusions: Dynamic measurement of resting heart rate variability may offer an objective measure of burnout in thoracic surgery residents. Thoracic surgery residents who report high levels of burnout in this cohort maintained the ability to meet program-designated milestones at or above the level expected of their postgraduate year. (JTCVS Open 2022;11:229-40)



HRV may offer an objective measure of burnout in thoracic residents.

CENTRAL MESSAGE

Dynamic measurement of resting HRV may offer an objective measure of burnout in thoracic residents. Residents reporting burnout are able to meet program-specific performance goals.

PERSPECTIVE

Despite progress in recognizing burnout and efforts to mitigate its effects, it has become endemic in cardiothoracic surgeons. The complex etiology and limited methods of measurement have made definitive interventions for burnout elusive. This proof of concept study shows that dynamic measurement of resting HRV may offer an objective measure of burnout in thoracic surgery residents.

From the ^aDivision of Cardiothoracic Surgery, Washington University School of Medicine, St Louis, Mo; and ^bDivision of Cardiothoracic Surgery, Michael E. DeBakey Department of Surgery, Baylor College of Medicine, Houston, Tex.

Institutional Review Board Approval: 201907040, January 31, 2022.

The editorial review of this article was handled by Associate Editor James I Fann, MD.

Read at the 102nd Annual Meeting of The American Association for Thoracic Surgery, Boston, Massachusetts, May 14-17, 2022.

Received for publication Feb 10, 2022; revisions received April 3, 2022; accepted for publication April 28, 2022; available ahead of print Aug 1, 2022.

Address for reprints: Lauren K. Barron, MD, Division of Thoracic Surgery, Department of Surgery, Washington University, 660 South Euclid Ave, St Louis, MO 63110 (E-mail: muckleroy@wustl.edu).


2666-2736

Copyright © 2022 The Author(s). Published by Elsevier Inc. on behalf of The American Association for Thoracic Surgery. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

<https://doi.org/10.1016/j.xjon.2022.04.044>

Abbreviations and Acronyms

ACGME	= Accreditation Council for Graduate Medical Education
FOR	= functional overreach
HR	= heart rate
HRV	= heart rate variability
MBI	= Maslach Burnout Inventory
NFOR	= nonfunctional overreach
OTS	= overtraining syndrome
PGY	= postgraduate year
REM	= rapid eye movement

 Video clip is available online.

To view the AATS Annual Meeting Webcast, see the URL next to the webcast thumbnail.

Physician burnout, defined as emotional exhaustion, depersonalization, and decreased sense of personal accomplishment, affects the majority of thoracic surgical residents and approximately half of the thoracic surgeons in practice.¹⁻⁴ Burnout has been shown to be an independent risk factor for cardiovascular disease, type 2 diabetes, and all-cause mortality.⁵⁻⁷ Among surgical residents and attending surgeons, burnout was associated with increased incidence of severe depressive symptoms and suicidal ideation.^{8,9} Suicide is 1.4 to 2.3 times more common in physicians than in the general population and the second most common cause of death in resident physicians in the United States.¹⁰ Although there is clearly crossover between symptoms of depression and burnout, multiple objective measures have demonstrated that the 2 entities are similar but separate.^{3,11}

Burnout not only affects the health and performance of physicians but also affects our patients and healthcare systems. Physician burnout is associated with twice the likelihood of patient safety incidents and decreased quality of care.¹² Stanford estimated that their financial burden caused by physician burnout is between \$15.5 and \$55.5 million dollars over 2 years.¹³ When these data are applied to the 950,000 US physicians, the cost increases to \$17 billion with \$28 billion in lost revenue.¹⁴

The data amassed over the last decade documenting physician burnout are extensive. Despite this, there are no preventive measures or interventions that consistently address burnout, and it continues to rise among physicians.

The cause of burnout is multifaceted, and the significance of each contributing variable may change as the level of training increases and differs between medical specialties.² The complex etiology and limited methods of measurement have made definitive interventions for burnout elusive.

The current gold standard for identifying and measuring both burnout and any response to interventions is the Maslach Burnout Inventory (MBI).¹ The MBI is a scale designed to assess various aspects of the burnout syndrome. Although this is a well-validated resource to assess the prevalence of burnout among physicians, the nature of the instrument limits its sensitivity and reliability in measuring the impact of interventions. There is a critical need to identify a minimally invasive approach to objectively measure the physiologic as well as the psychologic impact of interventions aimed at reducing physician burnout.

Prolonged exposure to stressful environments is one of the key drivers of burnout.¹⁵ Stressful stimuli induce a shift toward overriding sympathetic input (tone) that alters the cardiac rhythm, increasing the heart rate (HR) and decreasing the variability in the time between heartbeats. Heart rate variability (HRV) quantifies the variation in the time interval between consecutive heartbeats in milliseconds. The heart does not beat rhythmically like a metronome; there is constant variability, and that variability mirrors autonomic nervous system activity. In healthy adults exposed to an acute stressor, sympathetic tone is quickly downregulated by parasympathetic input, slowing the HR and increasing the HRV as the stress is reduced. In situations of prolonged stress, decreased parasympathetic tone and lower HRV occur as the individual can no longer balance sympathetic stimulation appropriately.¹⁶ As a measurable parameter, HRV offers a noninvasive biomarker of stress and recovery. With sympathetic activation during stress, HRV decreases. With increased parasympathetic tone during recovery, HRV increases. HRV has been correlated to burnout in cardiac patients, college students, and athletes who use HRV to design training regimens to improve performance.^{17,18} We hypothesized that HRV data could be applied to a cohort of thoracic surgical residents and might offer an objective physiologic measure of burnout. This study intends to report the feasibility of using the whoop strap to collect physiologic data from thoracic surgical trainees and whether these data correlate with burnout. The second objective was to determine whether burnout would correlate with performance as described by the Accreditation Council for Graduate Medical Education (ACGME) milestones.

MATERIALS AND METHODS

This was a prospective study of full-time clinical thoracic surgery residents in an early specialization (4 + 3) or standard integrated (5 + 2) training program over a 5-month period ending January 2022. Participants were evaluated with a wearable biometric device (Whoop strap), the MBI,

and the ACGME Milestones scores. Demographic data (sex, ethnicity, marital status, and number of children) and academic data (clinical training year, training focus, ACGME milestones report) were collected via survey and residency files. Physiologic data (resting HR, HRV, and sleep parameters) were measured using the Whoop strap. Selection of the physiologic data reported was based on the timing of survey responses. Of the enrolled trainees, 7 of 9 (78%) completed survey data. Participants who were pregnant, breastfeeding, or within 12 weeks of delivery were excluded from physiologic data analysis.¹⁹ The entire cohort of ACGME fellows ($n = 6$, postgraduate year [PGY] 6-7) was included in performance and burnout assessments. Performance and burnout, but not physiologic data, were collected on 1 PGY 5 (4 + 3) resident. The remaining 2 residents enrolled but ultimately chose not to participate; 1 completed training and left the institution before data collection, and 1 was not compliant with wearing the device or answering survey. The Institutional Review Board of the Washington University reviewed our protocol before data collection and approved the study (201907040, January 31, 2022), and all study participants signed informed consent.

Physiologic Assessment

The Whoop strap is a commercially available device worn on the wrist or bicep that measures HR and HRV via photoplethysmography and movement via a 3-axis accelerometer 100 times per second on a continual basis. These measures are used to detect sleep and sleep stage. For each resident, HRV was quantified as the root mean square of successive beat to beat interval differences and was calculated during slow-wave, deep sleep.²⁰ Residents were asked to wear the devices for the duration of rest periods to capture resting HRV. The devices could be worn continuously providing feedback through the app to the resident, but nonrest periods were not tracked by the study.

Sleep

Normal sleep in a healthy adult consists of 2 primary sleep states: rapid eye movement (REM) sleep and non-REM sleep. Non-REM sleep comprises approximately 75% to 80% of total sleep time broken into 4 distinct stages with stages 1 and 2 labeled "light sleep" and stages 3 and 4 labeled slow-wave or deep sleep. Cyclic REM sleep occurs 4 to 6 times nightly accounting for approximately 20% to 25% of total sleep.²¹ For this study, REM sleep was reported as a percentage of total sleep as detected by the Whoop strap.

Assessment of Burnout Using the Maslach Burnout Inventory

The MBI is a 22-item survey scored on a 7-point Likert scale designed to measure burnout on 3 dimensions: emotional exhaustion (9 questions, score range 0-54), depersonalization (5 questions, range 0-30), and personal accomplishment (8 questions, range 0-48). The MBI was offered to participants multiple times over the course of 5 months. Participants were given the MBI at enrollment and could record responses to the survey weekly thereafter. There was no minimum response rate required after the initial survey, and participants were sent notifications of opportunity via e-mail with a unique link to the survey. Physiologic data were averaged over a 7-day period that included the day the MBI was submitted for each occurrence. MBI scores for each of the 3 dimensions were categorized as low, moderate, or high based on the scoring and interpretation key provided with the survey using common cutoff points. Emotional exhaustion was categorized as low if less than 17 and high if greater than 26. Depersonalization was categorized as low if less than 7 and high if greater than 12. Personal accomplishment was categorized as low if less than 32 and high if greater than 38. Burnout is defined as reporting a high score for emotional exhaustion, high score for depersonalization, or low score for personal accomplishment.

Performance Assessment

Resident performance was quantified using ACGME Thoracic Surgery-Independent Milestones 2.0, a criterion-based competency

assessment of residents that occurs every 6 months. Specialty-specific milestones were scored on a scale of 1-5 normalized to a program-specific target (3 for PGY 6 or less residents and 4 for PGY 7 residents). Our 2-year program has evolved into one in which most residents track into general thoracic or adult cardiac, so chief residents spend very little, if any time in the nonspecialty rotations. As such, we analyzed only specialty-specific milestones to assess progress (Table E1).

Statistical Analysis

Burnout and physiologic data were analyzed using univariate linear regression analysis (for continuous variables) or chi-square test (for categorical variables). The 3 MBI dimension scores were treated as continuous variables for HRV analysis. For analysis of burnout based on demographic characteristics, MBI dimensions were treated as categorical variables categorized into low, moderate, and high based on the highest value reported by each subject during the study period. Statistical analysis was performed using Sigmaplot version 12.5 for Windows (Systat Software, Inc).

RESULTS

The cohort consisted of 71% female participants (5 of 7), with 86% of residents having 1 or more children. No demographic characteristic was associated with any dimension of burnout (Table 1). Table 2 summarizes MBI scores for each dimension. Residents most often reported high levels of emotional exhaustion and depersonalization; however, a positive sense of personal accomplishment persisted. Burnout was experienced by 86% of residents on at least 1 MBI measurement and at least 1 of the criteria was met on 93% of all MBI measurements during the study period.

The average resting HRV in this cohort was 47 ± 21 ms and fell within normal range for age and gender.²² The average resting HR was 61 ± 6 beats/min. There was no correlation between emotional exhaustion or depersonalization and resting HR ($P > .5$ for all). In contrast, personal accomplishment correlated with a decrease in resting HR ($r(12) = 0.57$, $P = .04$); as the resident's sense of personal accomplishment increased, resting HR decreased. There was a significant correlation between HRV and emotional exhaustion ($r(12) = 0.65$, $P = .01$, Figure 1) but did not correlate with depersonalization ($P = .28$) or personal accomplishment ($P = .24$). Likewise, depersonalization and personal accomplishment did not correlate with resident performance on normalized ACGME milestones ($P = .12$ and $P = .75$, respectively); however, increased emotional exhaustion showed a significant correlation with higher resident performance on program-specific normalized ACGME milestones ($r(6) = 0.76$, $P = .047$, Figure 2).

On average, thoracic surgery residents got 5 hours and 52 minutes (± 76 minutes) of sleep nightly. Total sleep did not correlate with any dimension of burnout ($P > .4$ for all). When analyzing sleep stages, residents averaged $24\% \pm 4\%$ of total sleep time in REM, $22\% \pm 5\%$ of total sleep time in slow-wave sleep, and $54\% \pm 5\%$ in light sleep. When analyzing sleep stages, HRV positively correlated with REM sleep ($r(11) = 0.71$, $P = .01$) (Figure 3, A),

TABLE 1. Association of thoracic surgery resident demographics with the 3 dimensions of the Maslach Burnout Inventory

	n (%)	Emotional exhaustion	Depersonalization	Personal accomplishment
Age (y)	34.3 ± 1.7	<i>P</i> = .88	<i>P</i> = .47	<i>P</i> = .91
Gender		<i>P</i> = .76	<i>P</i> = .7	<i>P</i> = .76
Male	2 (29%)			
Female	5 (81%)			
Ethnicity		<i>P</i> = .91	<i>P</i> = .61	<i>P</i> = .91
White	6 (86%)			
Asian	1 (14%)			
Year in training		<i>P</i> = .8	<i>P</i> = .57	<i>P</i> = .8
Senior Fellow	3 (43%)			
Junior Fellow	3 (43%)			
4 + 3 Resident	1 (14%)			
Training focus		<i>P</i> = .54	<i>P</i> = .17	<i>P</i> = .54
Cardiac	3 (43%)			
Thoracic	4 (57%)			
Married or in a domestic partnership	7 (100%)	—	—	—
No. of children		<i>P</i> = .8	<i>P</i> = .52	<i>P</i> = .09
0	1 (14%)			
1	2 (29%)			
2	3 (43%)			
3+	1 (14%)			

but had no correlation with any other sleep stage (Video Abstract). Additionally, there was a significant negative correlation between emotional exhaustion and percent time spent in REM sleep ($r(11) = 0.67, P = .02$) (Figure 3, B).

DISCUSSION

Cardiovascular Measurements

HRV has been used as an objective measure of alterations in stress-related physiology across a multitude of professions, including physicians.^{3,16,23} In surgeons, The and co-authors²⁴ found that HRV reliably detected differences in mental stress in an operative setting. The review included 17 studies of surgeons (n = 1-20 surgeons) and showed that HRV could be used to pinpoint stressors within the operative environment with regard to time, role, and technique. Within this collection of studies, there were no designs aimed at using HRV to measure long-term stress or burnout specifically. The Dresden burnout study, a large-scale longitudinal study of burnout in adults, measured burnout (MBI), and HRV (root mean square of

successive beat to beat interval differences) in 410 subjects.³ Similar to our results, they found a correlation between emotional exhaustion and HRV. Further, high emotional exhaustion at the initial time point predicted reductions in HRV at 12 months.³ Within the thoracic surgery residents included in the current report, HRV decreased as emotional exhaustion increased.

When looking at resting HR, previous work found that subjects with significant burnout (n = 22), defined as requiring 2 weeks to 6 months away from work, had increased resting HR compared with healthy controls (n = 23).²⁵ In line with these data, we found that a decrease in personal accomplishment was correlated to an increase in resting HR, further supporting the hypothesis that burnout has physiologic changes that can be measured.

Sleep

In addition to HR and HRV, the Whoop device detects sleep stages and has been validated against polysomnography. The Swedish National Institute for Psychosocial

TABLE 2. Distribution of the 3 dimensions of the Maslach Burnout Inventory among thoracic surgery residents

	Median [IQR]	Low n (%)	Moderate n (%)	High n (%)
Emotional exhaustion	30 [20-36]	2 (15%)	5 (38%)	6 (46%)
Depersonalization	16 [14-22]	1 (8%)	1 (8%)	11 (85%)
Personal accomplishment	43 [41-46]	1 (8%)	2 (15%)	10 (77%)

IQR, Interquartile range.

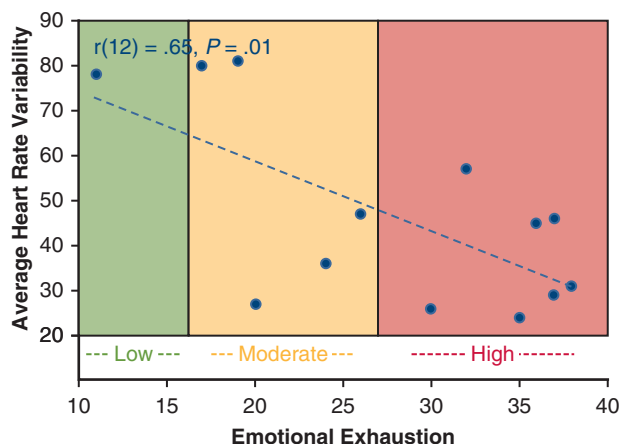


FIGURE 1. Emotional exhaustion correlates with HRV in thoracic surgery residents ($r(12) = 0.65, P = .01$).

Factors and Health studied sleep in subjects with severe burnout compared with healthy controls. They classified severe burnout as having Shirom Melamed Burnout Questionnaire greater than 4.5 on scale of 1 to 7 and being on full-time sick leave for 3 months. This approach offers a unique perspective given that subjects in most studies on burnout are recruited from the “working sick” excluding the most severely affected who are no longer able work. When compared with controls, subjects with high burnout had 6% less total REM sleep (19% vs 24% of total sleep).²⁶ These subjects also exhibited the classic pattern of sleep fragmentation, previously shown to be associated with burnout, further corroborating the notion that REM sleep is decreased in burnout.²⁷ The present study is limited in that sleep fragmentation cannot be accurately documented because frequent sleep disturbances are the hallmark of call shifts. However, REM sleep was found to correlate negatively with increased emotional exhaustion and positively with HRV, strengthening hypothesis that HRV may

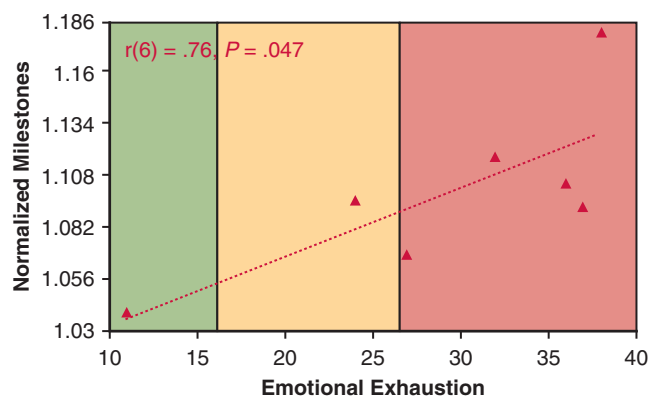


FIGURE 2. Emotional exhaustion correlated with thoracic surgery resident performance assessed using program-specific normalized ACGME milestones ($r(6) = 0.76, P = .047$).

offer an objective measure of burnout in thoracic surgical residents.

Performance

The ACGME Milestones project was designed to provide a framework for evaluating performance throughout training. Residents are expected to move from novice (a score of 1) toward expert (a score of 5) aiming to achieve a score of 4 at the completion of training.²⁸ Unexpectedly, the current report demonstrates high levels of emotional exhaustion had a significant positive correlation with higher scores on ACGME Milestones performance. One explanation for this finding is that both performance and burnout are a reflection of workload in thoracic surgery residents. The Areas of Worklife Scale is a measure aimed at analyzing the work environment as components that play a role in altering the continuum between engagement and burnout.¹⁵ Of the 6 areas defined by Leiter and Maslach,¹⁵ workload has been shown to have a consistent relationship with emotional exhaustion. Within surgical training, workload, as quantified by operative volume, correlates with increased performance, as measured by ACGME milestones and simulated procedural tasks.²⁹⁻³² With workload leading to both increased performance and emotional exhaustion, it is something of a double-edged sword for surgical training programs. HRV may offer a mechanism by which we can allow the resident to push to the limits of their capacity and safely monitor for adequate recovery allowing for maximum efficiency and improved performance.

Applications and Future Directions

In athletes, the phenomena of decreased performance despite increased or maintained levels of effort and training is referred to as “overtraining syndrome” (OTS). Also originally termed “burnout,” OTS is a continuum. The burnout continuum begins with undertraining, where demand is less than allostatic load and the athlete meets demands without a required recovery period. The consequence is maintenance of performance but no further improvement. The burnout continuum advances to OTS, where the allostatic load exceeds athletic capacity and the athlete is unable to recover, even with prolonged rest. The ultimate consequence is that performance continually declines. In successful athletic training programs, planned workloads beyond the athlete’s current capacity called “overtraining” are often used as a means to stimulate adaptation resulting in improved performance. Planned overtraining is termed “over-reaching” and is followed by a period of relative rest or tapered effort to allow for compensation and increased capacity.³³ The European College of Sport Science and the American College of Sports Medicine consensus statement further supports this construct endorsing the use of the terms functional overreach (FOR) and nonfunctional overreach (NFOR). FOR is determined by the athlete’s ability to recover in a period of days

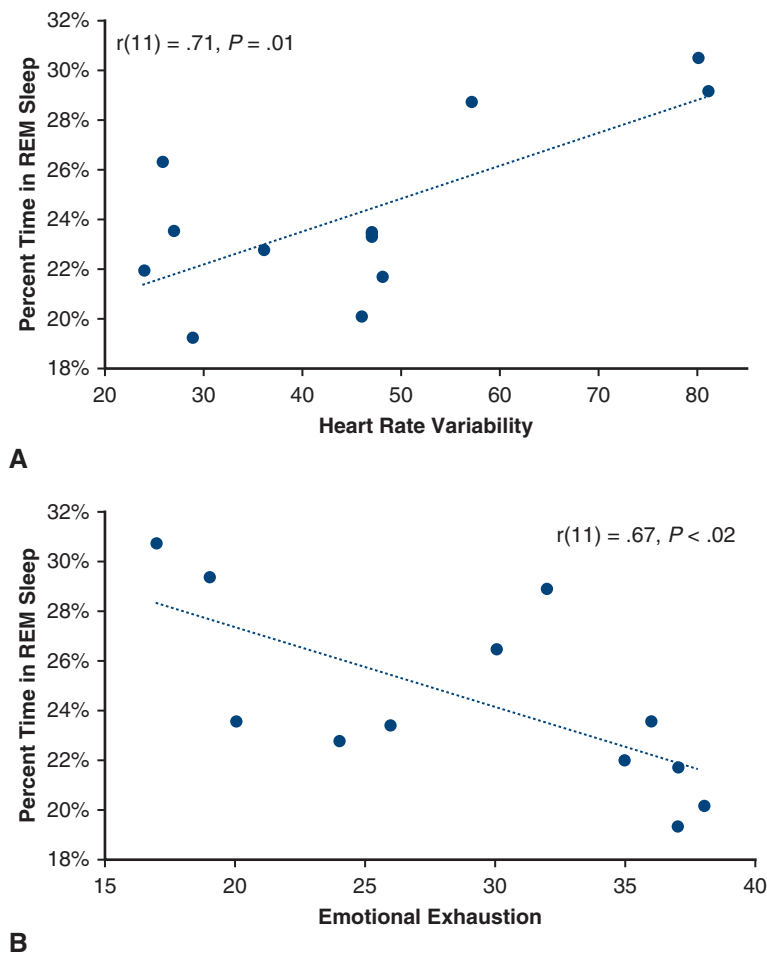


FIGURE 3. A, Percent time in REM sleep correlated positively with HRV ($r(11) = 0.71, P = .01$). B, Percent time in REM sleep correlated negatively with emotional exhaustion ($r(11) = 0.67, P = .02$). *REM*, Rapid eye movement.

to weeks. This is in contrast to NFOR or continued intense training into the period of declined performance requiring weeks to months to recover to baseline performance.³⁴ This model also recognizes that increased workload is tied to both enhanced performance and OTS and burnout.

Kajaia and associates³⁵ investigated whether HRV could be used to distinguish athletes with NFOR and OTS ($n = 43$) from those in the same discipline with the same training load without it ($n = 40$) and compared all athletes with 35 sedentary subjects of the same age. Similar to what is seen in our cohort with respect to emotional exhaustion, they found that athletes with NFOR/OTS had a lower HRV compared with athletes without NFOR/OTS. Despite obvious superior physical conditioning compared with sedentary controls, the group of NFOR/OTS athletes had HRV similar to sedentary controls. The similarities among the technical demands, performance measures, and required endurance of training in thoracic surgery and elite athletes are remarkable when viewed through the lens of the physiologic response to

stress. This parallel offers an opportunity to further explore whether surgical training might benefit from being modeled on the construct of FOR and NFOR.

We hypothesize that adjusting training intensity based on a physiologic indicator of physical and psychological allosteric load (HRV in the case of our thoracic surgery residents) would allow us to maximize the efficiency of training (Figure 4, A). Having real-time data on resident capacity would allow faculty to optimize the timing of increases in responsibility and autonomy. Adjusting expectations on this level would maximize efficiency for both the resident and faculty given that faculty demonstrate more stress when supervising than when performing as the primary operator.²⁴ Having continuous, longitudinal HRV data could alert programs when a resident is unlikely to improve from continued maximal effort and might benefit from a pivot in focus similar to cross-training (Figure 4, B). At these points, residents might be encouraged to make progress in less intense requirements such as simulation, patient safety and quality improvement

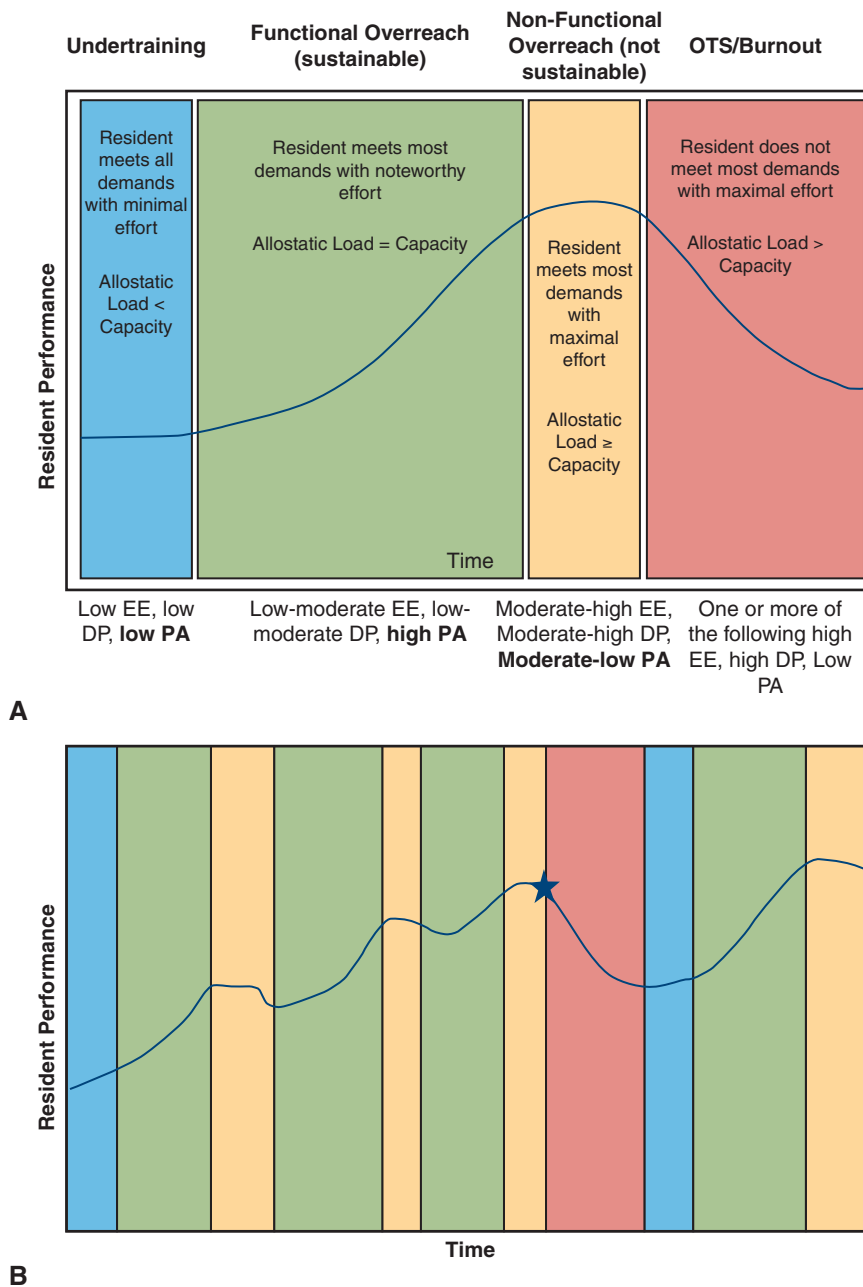


FIGURE 4. Thoracic surgery residents training on the OTS continuum. A, Heart rate variability can be used to allow residents to safely overreach with periods of recovery as they move into NFOR. B, Opportunities to increase responsibility and autonomy (blue and green blocks), to evaluate progress and adjust focus to target areas (yellow), or to pivot/intervene (starred timepoint) to prevent prolonged decrease in performance in response to overtraining and burnout (red). OTS, Overtraining syndrome; EE, emotional exhaustion; DP, depersonalization; PA, personal accomplishment.

projects, or less physically demanding procedures such as endoscopy or endovascular interventions. These pivots would allow for a period of recovery while continuing to meet the educational goals of the program. We could optimize the targets to better match capacity, that is, work smarter, not harder. Novel contingency plans for flexing coverage will need to be

developed to balance program needs to individual workload capacity.

HRV could be used not only to guide training but also to measure the impact of proposed interventions aimed at both the individual and systems levels. Having a physiologic measure of effectiveness would allow programs and

hospital systems to focus investments (time and funding) on data-driven interventions. By eliminating perfunctory measures aimed at burnout that are found to be ineffective, we could free up additional resources for the more costly system-level changes. Further investigation into the utility and reliability of HRV as an objective measure of burnout is warranted and has the potential for stimulating positive changes in the efficacy of our training.

Study Limitations

To accomplish data-driven interventions, a multi-institutional study would be required for several reasons. The cohort size and single institution of this pilot study limit the ability to report many factors known to influence HRV (gender, health conditions, and medications) due to the substantial risk of unintentionally reporting identifiable health information on female and underrepresented minority subjects. Given the disproportionately small number of these individuals currently in training, only a large sample would allow for meaningful analysis and keeping any results truly anonymous.

Additionally, the majority of residents had high levels of emotional exhaustion or depersonalization at each MBI assessment, not allowing for comparisons between HRV in subjects with and without burnout. For this feasibility study, a control group was not identified. To offer a meaningful assessment, control subjects would need to be trainees with similar roles, responsibilities, and attending

physicians. Using trainees from other fields or at junior levels would not assess device performance or study design feasibility in the target environment. Additionally, milestones are specialty-specific, limiting comparisons between subjects and nonthoracic surgery controls. For these reasons, we propose that a large multi-institutional study would provide enough data to select a control cohort of individuals or time points where burnout is not reported. Finally, we were unable to compare changes from baseline because of the inability to calculate a true baseline HRV because all subjects have been in training for 7 to 10 years at the time of this study. To address this limitation, a longitudinal study following residents over the course of the training period would be required and is recommended by the authors.

CONCLUSIONS

These data suggest that HRV may offer an objective measure of burnout in a homogenous population (Figure 5). A large multi-institutional study design would be necessary for a representative control group to be identified and to report health information known to affect HRV while maintaining the anonymity of the subjects. Additional longitudinal studies are needed to determine whether HRV changes may precede or are the result of changes in emotional exhaustion. Further characterization of HRV as a measure of recovery may allow for a more targeted and efficacious approach to maximizing the training period in thoracic surgery.

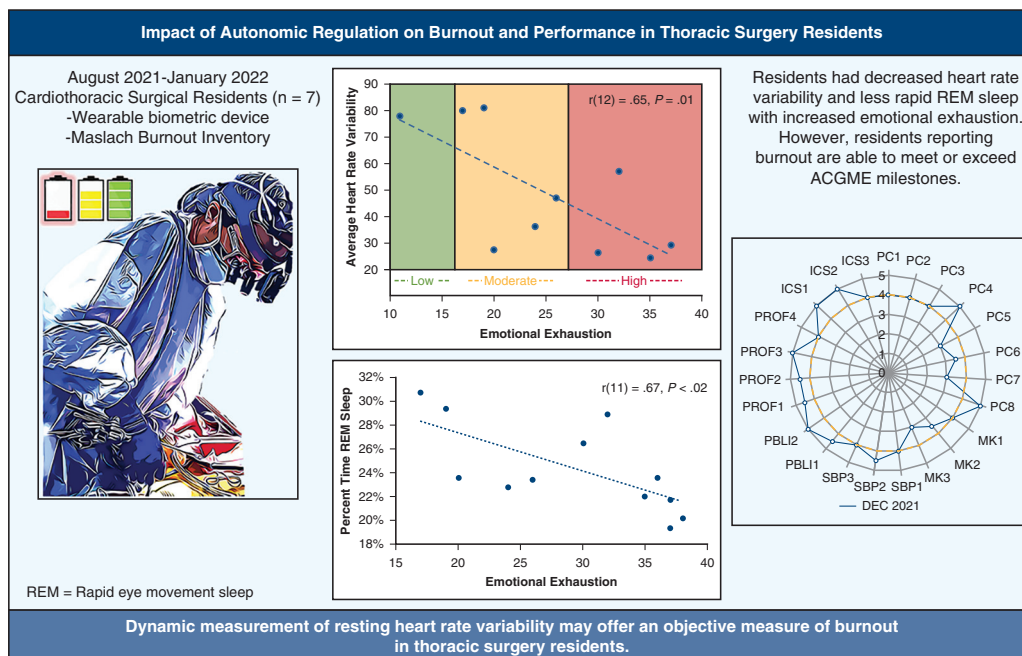


FIGURE 5. Dynamic measurement of resting HRV may offer an objective measure of burnout in thoracic surgery residents. Residents had decreased HRV and less REM sleep with increased emotional exhaustion but maintained the ability to meet milestones at or above the level expected of their postgraduate year. REM, Rapid eye movement; ACGME, Accreditation Council for Graduate Medical Education.

Webcast

You can watch a Webcast of this AATS meeting presentation by going to: <https://www.aats.org/resources/2886>.



Conflict of Interest Statement

M.R.M. is a consultant/advisory board member for Medtronic. The other author reported no conflicts of interest.

The *Journal* policy requires editors and reviewers to disclose conflicts of interest and to decline handling or reviewing manuscripts for which they may have a conflict of interest. The editors and reviewers of this article have no conflicts of interest.

References

- Maslach C, Jackson SE. The measurement of experienced burnout. *J Organiz Behav*. 1981;2:99-113.
- Medscape National Physician Burnout, Depression & Suicide Report 2019*. Medscape; 2019.
- Wekenborg MK, Hill LK, Thayer JF, Penz M, Wittling RA, Kirschbaum C. The longitudinal association of reduced vagal tone with burnout. *Psychosom Med*. 2019;81:791-8.
- Chow OS, Sudarshan M, Maxfield MW, Seese LM, Watkins AA, Fleishman A, et al. National survey of burnout and distress among cardiothoracic surgery trainees. *Ann Thorac Surg*. 2021;111:2066-71.
- Kitaoka-Higashiguchi K, Morikawa Y, Miura K, Sakurai M, MasaIshizaki M, Kido T, et al. Burnout and risk factors for arteriosclerotic disease: follow-up study. *J Occup Health*. 2009;51:123-31.
- Ahola K, Vaananen A, Koskinen A, Kouvonen A, Shirom A. Burnout as a predictor of all-cause mortality among industrial employees: a 10-year prospective register-linkage study. *J Psychosom Res*. 2010;69:51-7.
- Melamed S, Shirom A, Tokar S, Shapira I. Burnout and risk of type 2 diabetes: a prospective study of apparently healthy employed persons. *Psychosom Med*. 2006;68:863-9.
- Lebares CC, Guvva EV, Ascher NL, O'Sullivan PS, Harris HW, Epel ES. Burnout and stress among US surgery residents: psychological distress and resilience. *J Am Coll Surg*. 2018;226:80-90.
- Shanafelt TD, Balch CM, Dyrbye L, Beachamps G, Russell T, Satele D, et al. Special report: suicidal ideation among American surgeons. *Arch Surg*. 2011;146:54-62.
- Yaghmour NA, Brigham TP, Richter T, Miller RS, Philibert I, Baldwin DC Jr, et al. Causes of death of residents in ACGME-accredited programs 2000 through 2014: implications for the learning environment. *Acad Med*. 2017;92:976-83.
- Koutsimani P, Montgomery A, Georganta K. The relationship between burnout, depression, and anxiety: a systematic review and meta-analysis. *Front Psychol*. 2019;10:284.
- Panagioti M, Geraghty K, Johnson J, Zhou A, Panagopoulou E, Chew-Graham C, et al. Association between physician burnout and patient safety, professionalism, and patient satisfaction: a systematic review and meta-analysis. *JAMA Intern Med*. 2018;178:1317-31.
- Hamidi MS, Bohman B, Sandborg C, Smith-Coggins R, de Vries P, Albert MS, et al. Estimating institutional physician turnover attributable to self-reported burnout and associated financial burden: a case study. *BMC Health Serv Res*. 2018;18:851.
- Krippner J. National Taskforce for Humanity in Healthcare, position paper: the business case for humanity in healthcare; 2022. Accessed November 6, 2021. https://www.healthcareexcellence.org/wp-content/uploads/2018/04/NTH-Business-Case_2018.pdf
- Leiter MP, Maslach C. Areas of worklife: a structured approach to organizational predictors of job burnout. In: Perrewé P, Ganster DC, eds. *Emotional and Physiological Processes and Positive Intervention Strategies*. Elsevier Science/JAI Press; 2004:91-134.
- Kanthak MK, Stalder T, Hill LK, Thayer JF, Penz M, Kirschbaum C. Autonomic dysregulation in burnout and depression: evidence for the central role of exhaustion. *Scand J Work Environ Health*. 2017;43:475-84.
- Zhang M, Liu L, Shi Y, Yang Y, Yu X, Angerer P, et al. Longitudinal associations of burnout with heart rate variability in patients following acute coronary syndrome: a one-year follow-up study. *Gen Hosp Psychiatry*. 2018;53:59-64.
- May RW, Seibert GS, Sanchez-Gonzalez MA, Fincham FD. School burnout and heart rate variability: risk of cardiovascular disease and hypertension in young adult females. *Stress*. 2018;21:211-6.
- Groer MW, Jevitt CM, Sahebzamani F, Beckstead JW, Keefe DL. Breastfeeding status and maternal cardiovascular variables across the postpartum. *J Womens Health (Larchmt)*. 2013;22:453-9.
- Miller DJ, Lastella M, Scanlan AT, Bellenger C, Halson SL, Roach GD, et al. A validation study of the WHOOP strap against polysomnography to assess sleep. *J Sports Sci*. 2020;38:2631-6.
- Carskadon MA, Dement WC. Monitoring and staging human sleep. In: Kryger MH, Roth T, Demet WC, eds. *Principles and Practice of Sleep Medicine*. 5th ed. Elsevier Saunders; 2011:16-26.
- van den Berg ME, Rijnbeek PR, Niemeijer MN, Hofman A, van Herpen G, Bots ML, et al. Normal values of corrected heart-rate variability in 10-second electrocardiograms for all ages. *Front Physiol*. 2018;9:424.
- de Vente W, van Amsterdam JG, Olf M, Kamphuis JH, Emmelkamp PM. Burnout is associated with reduced parasympathetic activity and reduced HPA axis responsiveness, predominantly in males. *Biomed Res Int*. 2015; 2015:431725.
- The AF, Reijmerink I, van der Laan M, Cnossen F. Heart rate variability as a measure of mental stress in surgery: a systematic review. *Int Arch Occup Environ Health*. 2020;93:805-21.
- De Vente W, Olf M, Van Amsterdam JG, Kamphuis JH, Emmelkamp PM. Physiological differences between burnout patients and healthy controls: blood pressure, heart rate, and cortisol responses. *Occup Environ Med*. 2003;60(Suppl 1): i54-61.
- Ekstedt M, Soderstrom M, Akerstedt T, Nilsson J, Sondergaard HP, Aleksander P. Disturbed sleep and fatigue in occupational burnout. *Scand J Work Environ Health*. 2006;32:121-31.
- Soderstrom M, Ekstedt M, Akerstedt T, Nilsson J, Axelsson J. Sleep and sleepiness in young individuals with high burnout scores. *Sleep*. 2004;27:1369-77.
- The thoracic surgery milestone project. *J Grad Med Educ*. 2014;6:332-54.
- Sayegh F, Margulies I, Zoghbi Y, Taub PJ. Plastic surgery resident operative volume and milestones performance: what is the correlation? *Plast Reconstr Surg Glob Open*. 2020;8:131-2.
- Sheahan M, Lalani AM, Lee JT, Shames M. The fundamentals of vascular surgery: when do vascular trainees achieve basic open surgical competency? *J Vasc Surg*. 2020;72:e24-5.
- Doneza JA, Palvia V, Lerner VT, Overbey J, Levie M, Brodman M, et al. Correlation of surgical case volume and fellowship training with performance on simulated procedural tasks. *Am J Obstet Gynecol*. 2021;225:548.e1-10.
- Olson JJ, Zhang B, Zhu D, Zheng ET, Dyer GSM, Rozental TD, et al. Do resident surgical volumes and level of training correlate with improved performance on psychomotor skills tasks: construct validity testing of an ASSH training platform (STEP)? *JB JS Open Access*. 2021;6:e20.00123.
- Budgett R. Overtraining syndrome. *Br J Sports Med*. 1990;24:231-6.
- Meeusen R, Duclos M, Foster C, Fry A, Gleeson M, Nieman D, et al. Prevention, diagnosis, and treatment of the overtraining syndrome: joint consensus statement of the European College of Sport Science and the American College of Sports Medicine. *Med Sci Sports Exerc*. 2013;45:186-205.
- Kajaia T, Maskhulia L, Chelidze K, Akhalkatsi V, Kakhbrishvili Z. The effects of non-functional overreaching and overtraining on autonomic nervous system function in highly trained athletes. *Georgian Med News*. 2017;97-103.

Key Words: burnout, heart rate variability, milestones, performance, residency

Discussion

Presenter: Dr Lauren Barron



Dr Mara Antonoff (*Houston, Tex*). As experts in cardiothoracic physiology and pathophysiology, this topic is highly relevant to all of our practices. As educators, these issues highlight important considerations in our trainee learning environment as well as our culture of wellness. As you described,

you evaluated 7 trainees using a broad range of metrics, correlating physiologic data, burnout survey results, and milestone-based performance data. I have to say that this topic, as I mentioned to you, struck a personal note for me because it gave me the opportunity to recall that when I was a trainee in the institution from which these data were derived, during morning rounds one day, I actually went into really rapid supraventricular tachycardia and had to be treated for it. So, of course, I was curious to see your results and how that might have applied. In the study of your current trainees, you found that high levels of emotional exhaustion and depersonalization were common, although I did note in your presentation and in your article that 5 of the 7 study subjects were women, and most of them had kids, while the study was conducted in a peri-pandemic period. I suspect that there are a lot of contributing factors for the findings, potentially not all related to the fact that they were going through different aspects of their training. I was surprised to see that there's no significant relationship between gender and number of children, amount of sleep, or training track with the measurements of burnout. I think, as we've discussed, the small sample size can certainly be a limitation. Regardless, you suggested that dynamic measurements of HRV can serve as an objective measure and important surrogate for burnout in cardiothoracic surgical fellows. Although we might not have reached statistical significance on some of those other aspects that you've discussed, this proof-of-concept project is valuable, touching on a prevalent problem in our work environment.

As you highlight in your presentation, emotional exhaustion, depersonalization, and other symptoms of burnout have been shown to impact the majority of trainees in almost half of the practicing surgeons in our specialty. It's associated with negative health consequences and potential detriments to physical and emotional well-being. Burnout impacts our colleagues, our friends, ourselves, and the patients whom we all aim to serve. We are living in a modern and technology-rooted world that allows us to keep track of our sleep-wake cycles, our HRV, our steps, our body composition, and so much more. Our watches now are essentially Holter monitors. This study does a great job setting a foundation for suggestions that we may use such novel sources of data to get clarity on what's going on with our trainees and the

members of our workforce. The current study may be limited, again, by sample size, but it really does set the stage where we might be able to conduct future studies looking at different objective variables and understanding our responses to the day-to-day stresses that we encounter at work. As a specialty, we're understanding more and more about how important it is to consider wellness. As we contemplate strategies to improve trainee wellness, one thought is that perhaps we should be more focused on HRV than 80 hours per week. Maybe we should care more about emotional exhaustion than 1 day off in 28. There are certainly different ways to think about it. I do have a few questions for you. You talked about how we can use these results to optimize performance for our trainees. Do you have any thoughts about how we can also use data generated from this type of novel technology to guide resident wellness initiatives?



Dr Lauren Barron (*St Louis, Mo*). I think the first step will be establishing what a normative value looks like in our population. From there, we will have to consider the many different metrics of performance in a trainee. Training requires developing both our technical and professional skill sets,

as well as teaching us how to recover from high stress demands so that we can take on the next set of challenges. Otherwise, we will get to the attending level, experience stress at an intensity different from our previous experience, and fail to progress. We have to continue to make progress. I think having the ability to monitor real time gives us the capability to take a trainee who may be flagging in one aspect, but has significant potential for progress in another and redirect their training focus. In doing this, we essentially cross-train them. When you've worn someone out technically, maybe you flip them over and challenge them with a leadership position. In that way, I think you're right. Maybe 80 hours a week or 1 in 7 is not the right way to train somebody who's supposed to be a dynamic performer.

Dr Antonoff. That's a great idea, and I'm looking forward to see what you're going to be doing with all of these metrics going forward. Surprisingly, I think to many of us, you showed increased emotional exhaustion that significantly correlated with higher resident performance scores. I'm wondering to what you attribute those findings—and which do you think is the chicken, and which is the egg?

Dr Barron. I think we are similar to professional athletes with our performance response to stress. In a professional athlete, training programs are actually designed to over-train, reach a max, then recover and start again because the next time they reach their maximum, it's going to be higher. I think that is how we should be looking at training. Although the word burnout implies that a person is no longer capable of doing their job well, our data show that

this is not true in thoracic surgery residents. It may be true in other types of physicians, but in this cohort, and likely in the attendings, it's not the case.

Dr Antonoff. I suspect there are a lot of potential confounders. What do you think the impact might be about from people who potentially work out a lot or people who are on different medications, or those who have thyroid disorders, are pregnant, or are breastfeeding? How do you think those contribute? Do you think any of those characteristics could have masked or overstated any of your findings?

Dr Barron. I absolutely think that not knowing gives us the potential for masking or overstating the findings. I think a large cohort is necessary for us to be able to start drilling down on the active issues allowing us to individualize our wellness plans. Once we have the data to show what happens in trainees with uncommon circumstances, for example, trainees who are breastfeeding, it's going to be easier for us to help support them.

Dr Antonoff. Terrific. Thank you so much.

Dr Marc Moon (*Houston, Tex*). Thank you, Dr Antonoff, for a very insightful discussion. Can we ask Dr Goldstone to also discuss this article?



Dr Andrew Goldstone (*New York, NY*). A great presentation, really provocative and goes along with the theme of the meeting. Trying to do new things is really interesting. Your parallel of surgical training to athletic training, I think, is not unique. I think it's very tough training. It's tough outside of training, I can tell you that. As a first-year attending, it's stressful. Anyway, from my understanding, HRV varies substantially between individuals. So, from athletic training, they've learned that HRV can't really be compared between individuals but really is used more as a change from baseline or within-person metric. This study, although prospective, is effectively cross-sectional in design. So, what is your next step, and what's your next study for this?

Dr Barron. That is an excellent question, and yes to that observation. We designed the study prospectively, but when we got our data from the subject-driven methods, it looked more like cross-sectional data with many data points. We decided to proceed based on the findings from the Dresden Burnout Study. This group also applied HRV in a similar manner in a much larger population and were able to show that there's something there. In the next study, like you mentioned, we plan to have repeated measures on

each subject. After all, that's the next step in being able to target an intervention and measure its effectiveness. Because with the data we have right now, targeting an intervention is still just an idea.

Dr Goldstone. Excellent. It was surprising to me that only emotional exhaustion correlated with resident performance but not the other components for assessing burnout. Why do you think that's the case? Is the definition of burnout not significant enough, or?

Dr Barron. I think this happens because the definition of burnout is nebulous. We're really looking at the response of many contributing factors, and we're measuring one, then calling it burnout. It's a bit like saying someone has hypertension and not knowing why they have hypertension. We have a plethora of drugs that allow us to target different causes of hypertension. If you give someone a drug that does not target the etiology of their hypertension, it's ineffective. I think in that way, burnout has 3 dimensions, and sometimes the potential target that we're aiming for is the wrong one for an individual.

Dr Goldstone. Approximately 90% of the trainees were burned out by definition, 86%, but all of them performed well. I think their normalized scores were over 1. From my understanding, they're overperforming, right? So, WashU just beating down the residents well enough, putting them into supraventricular tachycardia or whatever? And is that helpful? Or is burnout just not a good metric here? Maybe you should just see if HRV or other metrics are just correlated straight to performance and get rid of the burnout assessment.

Dr Barron. Those 2 observations may be correct, but I suspect what we're seeing is only the first half of a performance curve because the data were not collected. There are not enough data to show the other half of the performance curve. The next step for us is to look at our entire general surgery population's ACGME milestones and MBI to try and define that performance curve. Either we're going to get to the top of it and see that we actually do peak and then fall off or it's just going to keep going up, and we have to figure out at what point do we stop and that the milestones are not measuring critical aspects of what it means to be both well trained and human.

Dr Goldstone. Congratulations on the article for everyone. It's available for publication today, right?

Dr Barron. *JTCVS Open*.

Dr Moon. Right. So please join me in thanking Drs Barron, Antonoff, and Goldstone for this innovative work and the creative variability in your discussion.

TABLE E1. Program-specific milestones

Cardiac track milestones analyzed	Thoracic track milestones analyzed
<ul style="list-style-type: none"> • Patient Care 1: Ischemic Heart Disease • Patient Care 2: Mechanical Circulatory Support • Patient Care 3: Valvular Disease • Patient Care 4: Great Vessel Disease • Patient Care 8: Critical Care • Medical Knowledge 1: Cardiovascular • Systems-Based Practice 1: Patient Safety and Quality Improvement • Systems-Based Practice 2: System Navigation for Patient-Centered Care • Systems-Based Practice 3: Physician Role in Health Care Systems • Practice-Based Learning and Improvement 1: Evidence-Based and Informed Practice • Practice-Based Learning and Improvement 2: Reflective Practice and Commitment to Personal Growth • Professionalism 1: Ethical Principles • Professionalism 2: Professional Behavior and Accountability • Professionalism 3: Administrative Tasks • Professionalism 4: Well-Being • Interpersonal and Communication Skills 1: Patient- and Family-Centered Communication • Interpersonal and Communication Skills 2: Interprofessional and Team Communication • Interpersonal and Communication Skills 3: Communication within Health Care Systems 	<ul style="list-style-type: none"> • Patient Care 5: Esophagus • Patient Care 6: Lung and Airway • Patient Care 7: Chest Wall/Pleura/Mediastinum/Diaphragm • Patient Care 8: Critical Care • Medical Knowledge 2: General Thoracic • Systems-Based Practice 1: Patient Safety and Quality Improvement • Systems-Based Practice 2: System Navigation for Patient-Centered Care • Systems-Based Practice 3: Physician Role in Health Care Systems • Practice-Based Learning and Improvement 1: Evidence-Based and Informed Practice • Practice-Based Learning and Improvement 2: Reflective Practice and Commitment to Personal Growth • Professionalism 1: Ethical Principles • Professionalism 2: Professional Behavior and Accountability • Professionalism 3: Administrative Tasks • Professionalism 4: Well-Being • Interpersonal and Communication Skills 1: Patient- and Family-Centered Communication • Interpersonal and Communication Skills 2: Interprofessional and Team Communication • Interpersonal and Communication Skills 3: Communication within Health Care Systems
Cardiac track exclusions	Thoracic track exclusions
<ul style="list-style-type: none"> • Patient Care 5: Esophagus • Patient Care 6: Lung and Airway • Patient Care 7: Chest Wall/Pleura/Mediastinum/Diaphragm • Medical Knowledge 2: General Thoracic • Medical Knowledge 3: Congenital. 	<ul style="list-style-type: none"> • Patient Care 1: Ischemic Heart Disease • Patient Care 2: Mechanical Circulatory Support • Patient Care 3: Valvular Disease • Patient Care 4: Great Vessel Disease • Medical Knowledge 1: Cardiovascular • Medical Knowledge 3: Congenital.