Burnout in Pediatric Residents: Comparing Brief Screening Questions to the Maslach Burnout Inventory

Kathi J. Kemper, MD, MPH; Paria M. Wilson, MD, MEd; Alan Schwartz, PhD;
John D. Mahan, MD; Maneesh Batra, MD, MPH; Betty B. Staples, MD;
Hilary McClafferty, MD; Charles J. Schubert, MD; Janet R. Serwint, MD;
on behalf of the Pediatric Resident Burnout-Resilience Study Consortium

From the College of Medicine, The Ohio State University (KJ Kemper); Nationwide Children's Hospital and the Department of Pediatrics, College of Medicine, The Ohio State University College of Medicine (JD Mahan), Columbus, Ohio; Department of Pediatrics, University of Cincinnati and Cincinnati Children's Hospital Medical Center (CJ Schubert), Cincinnati, Ohio; Department of Pediatrics, Division of Emergency Medicine (P Wilson), UPMC Children's Hospital of Pittsburgh, Pittsburgh, Pa; Department of Medical Education and Pediatrics (A Schwartz), University of Illinois College of Medicine, Chicago, Ill; Department of Pediatrics (M Batra), Seattle Children's Hospital, Seattle, Wash; Department of Pediatrics (B Staples), Duke University Medical Center, Durham, NC; Department of Medicine (H McClafferty), University of Arizona College of Medicine, Tucson, Ariz; and Department of Pediatrics (JR Serwint), The Johns Hopkins University, Baltimore, Md
The authors have no conflicts of interest to disclose.
Address correspondence to Kathi J. Kemper, MD, MPH, College of Medicine,The Ohio State University, 370 W 9th Ave, Columbus, OH 43210 (e-mail: kathi.kemper@osumc.edu).
Received for publication June 29, 2018; accepted November 2, 2018.

ABSTRACT

BACKGROUND: Measuring burnout symptoms is important, but the Maslach Burnout Inventory (MBI) has 22 items. This project compared 3 single-item measures with the MBI and other factors related to burnout.

METHODS: Data were analyzed from the 2016 and 2017 Pediatric Resident Burnout-Resilience Study Consortium surveys, which included standard measures of perceived stress, mindfulness, resilience, and self-compassion; the MBI; and the 1- and 2-item screening questions.

RESULTS: In 2016 and 2017, data were collected from 1785/2723 (65%) and 2148/3273 (66%) eligible pediatric residents, respectively. Burnout rates on the MBI were 56% in 2016 and 54% in 2017. The Physician Work Life Study item generated estimates of burnout prevalence of 43% to 49% and, compared with the MBI for 2016 and 2017, had sensitivities of 69% to 72%, specificities of 79% to 82%, positive likelihood ratios of 3.4 to 3.8, and negative likelihood ratios of 0.35 to 0.38. The combination of an emotional exhaustion

item and a depersonalization item generated burnout estimates of 53% in both years and, compared with the full MBI, had sensitivities of 85% to 87%, specificities of 84% to 85%, positive likelihood ratios of 5.7 to 6.4, and negative likelihood ratios of 0.18 for both years. Both items were significantly correlated with their parent subscales. The single items were significantly correlated with stress, mindfulness, resilience, and self-compassion.

CONCLUSIONS: The 1- and 2-item screens generated prevalence estimates similar to the MBI and were correlated with variables associated with burnout. The 1- and 2-item screens may be useful for pediatric residency training programs tracking burnout symptoms and response to interventions.

KEYWORDS: burnout; education; pediatric; resident; resilience; screening

WHAT'S NEW

Little is known about the performance of brief screening tools for burnout in pediatric residents. We found that 1- and 2-item screening questions have reasonable sensitivities, specificities, likelihood ratios, and construct validity, and they may be useful for tracking burnout symptoms in pediatric residency training programs.

BURNOUT IS AN increasing focus of pediatric physician training, highlighting the importance of accurately measuring burnout.[1] The most widely used instrument for assessing burnout is the 22-item Maslach Burnout Inventory (MBI), which has 3 subscales: emotional exhaustion (EE), depersonalization (DP), and personal accomplishment (PA).[2] Because burnout is often measured in lengthy surveys that assess risk and protective factors associated

with it, a screening instrument with fewer items and lower participant burden is desirable.

In 2009, West and colleagues[3] reported that 2 items from the MBI—"I feel burned out from my work" (an item from the EE subscale) and "I have become more callous to other people since I took this job" (an item from the DP subscale) —were consistent with the full MBI in predicting major medical errors and suicidality in medical students and internal medicine residents; however, they did not include a large sample of pediatric residents. In 2015, Dolan and colleagues[4] reported that a single non-proprietary item from the Physician Work Life Study (PWLS)—"Overall, based on your definition of burnout, how would you rate your level of burnout?" with a 5-point scale ranging from 1 ("I enjoy my work. I have no symptoms of burnout.") to 5 ("I feel completely burned out and often wonder if I can go on."), where scores of 3 or more are considered positive for burnout—had a sensitivity of 83% and specificity of 87% compared with the full MBI in a sample of practicing primary care physicians and nurses, which did not include pediatric residents.[4] Waddimba and colleagues[5] tested all 3 single-item measures against the full MBI in a sample of 308 physicians and advanced practice clinicians in rural upstate New York and reported that the 3 single-item measures varied in their effectiveness as screeners for EE or DP dimensions of burnout. Thus, although using just one or two items is appealing, there is a gap in understanding the test parameters of these 2 questions in pediatric residents who may differ from other groups of practicing physicians; for example, pediatric residents typically include a higher percentage of women and are likely to be younger than practicing clinicians. Being able to use single-item screens may be useful at the program level and in larger epidemiologic studies looking at trends and responses to interventions. Furthermore, there are limited data on the correlation between these screening questions and commonly measured risk and protective factors for burnout (ie, stress, mindfulness, resilience, and self-compassion), and it would be useful to understand these relationships before relying on single-item measures alone to track burnout symptoms.

The Pediatric Resident Burnout-Resilience Study Consortium (PRB-RSC) has been conducting annual national surveys of burnout in pediatric residents since 2016, and we were considering using 1 or 2 screening questions rather than the full 22-item MBI in future studies. Before doing so, however, we aimed to answer 2 research questions: 1) What are the sensitivity, specificity, positive and negative predictive values, and likelihood ratios of these screening items compared to the MBI in pediatric residents? 2) How do the correlations compare between the screening questions and full MBI to the variables of interest (stress, mindfulness, resilience, and self-compassion) in pediatric residents?

METHODS

The samples and survey instrument have been described fully elsewhere.[6−8] Briefly, pediatric residents were surveyed in the spring of 2016 and 2017 from 34 and 46 programs, respectively. Resident responses were eligible if they answered all 22 items on the MBI. For the MBI, burnout as a dichotomous variable was defined as having high subscale scores for personal emotional exhaustion ($\geq$27) and/or depersonalization ($\geq$10).[9,10] Burnout as a continuous variable was defined as the total of the EE score plus the DP score. The continuous variable was used not to identify burnout but to assess correlations with other variables. For this study, consistent with the PWLS study, we considered the single screening question positive if the score was $\geq$3 on the 5-point scale (1 or more symptoms of burnout).[4] Consistent with the West et al[3] study, the 2 screening items were considered positive for burnout if the responses on either the EE or DP item were endorsed at least weekly ($>$3 on the 7-point Likert scale from never to daily). We determined the sensitivity, specificity, positive and negative predictive values, and likelihood ratios of the PWLS screen compared with the full MBI in 2016 and 2017, as well as the EE and DP items alone and together compared with the EE and DP subscales and the full MBI for respondents in 2016 and 2017.

The survey also included widely used measures of stress (Perceived Stress Scale),[11] mindfulness (Cognitive and Affective Mindfulness Scale-Revised),[12] resilience (Brief Resilience Scale),[13] and self-compassion (Neff's Self-Compassion Scale, short form).[14] To answer the second study question, consistent with the study by West et al,[3] we multiplied the numeric response to the first screening question by 9 (because the emotional exhaustion subscale of the MBI has 9 items), we multiplied the numeric response to the second screening question by 5 (because the depersonalization subscale of the MBI has 5 items), and then we added those results together.[3] We then compared correlations of the full MBI versus the 2-item screen to standard measures of stress, mindfulness, resilience, and self-compassion. Because there was substantial overlap in respondents from 2016 and 2017, we analyzed data from each year separately.

Survey data were exported from LimeSurvey to R 3.3.1 for analysis. Descriptive statistics were used to characterize participants. Spearman's rank correlations were used to calculate correlations between variables. The Nationwide Children's Hospital Institutional Review Board and other participating programs approved this study.

RESULTS

Among the 1758/2723 (65%) eligible pediatric resident respondents in 2016 and 2148/3273 (66%) eligible respondents in 2017, 72% were women, 60% were married, and 62% reported being more than $100,000 in debt. In 2016, 56% met the criteria for burnout on the full MBI; in 2017, 54% met the criteria. Table 1 shows the value of 1- and 2-item screens for predicting subscales and full-scale scores of the MBI. Results from the 2016 survey are presented on top, and a replication based on the 2017 survey is presented on the bottom. The PWLS single-item screen generated an estimated prevalence of burnout of 46% in 2016 and 49% in 2017; it had a sensitivity of 69% and specificity of 82% (positive likelihood ratio [LR

Table 1. Sensitivity, Specificity, Likelihood Ratios, and Positive and Negative Predictive Values of Screening Questions in 2016 and 2017

	Burnout Rate (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)	LR+ (95% CI)	LR− (95% CI)	PPV	NPV
2016	56% on MBI						
PWLS screening question compared with MBI	46% (44−49)	69% (66−72)	82% (79−84)	3.8 (3.2−4.4)	0.38 (0.35−0.42)	83%	67%
EE question compared with EE subscale of MBI	44% (43−46)	72% (69−74)	82% (79−84)	4.0 (3.5−4.5)	0.34 (0.31−0.38)	79%	75%
DP question compared with DP subscale of MBI	38% (36−40)	69% (66−72)	82% (80−84)	6.4 (5.3−7.8)	0.18 (0.16−0.21)	89%	71%
Combined EE and DP questions compared with full MBI	53% (50−55)	84% (82−86)	87% (84−89)	6.4 (5.3−7.8)	0.18 (0.16−0.21)	89%	81%
2017	54% on MBI						
PWLS screening question compared with MBI	49% (47−51)	72% (69−75)	79% (76−81)	3.4 (3.0−3.9)	0.35 (0.32−0.39)	80%	71%
EE question compared with EE subscale of MBI	41% (39−43%)	78% (75−80)	90% (88−92)	7.7 (6.5−9.2)	0.25 (0.22−0.28)	87%	82%
DP question compared with DP subscale of MBI	35% (33−37)	77% (74−80)	88% (86−90)	6.6 (5.6−7.6)	0.26 (0.23−0.30)	78%	87%
Combined EE and DP questions compared with full MBI	53% (50−55)	85% (83−87)	85% (83−87)	5.7 (4.9−6.6)	0.18 (0.16−0.21)	87%	82%

CI indicates confidence interval; LR+, positive likelihood ratio; LR−, negative likelihood ratio; PPV, positive predictive value; NPV, negative predictive value; PWLS, Physician Work Life Study; MBI, Maslach Burnout Inventory; EE, emotional exhaustion subscale of the MBI; DP, depersonalization subscale of the MBI.

Table 2. Spearman's Rank Correlations of 2 Screening Items and Full-Scale MBI with Stress, Mindfulness, Resilience, and Self-Compassion among Pediatric Residents in 2016 and 2017

	Full MBI	Screen MBI (2 Questions)	MBI EE Subscale	MBI DP Subscale	Stress	Mindfulness	Resilience	Self-Compassion
				2016 Residents				
Full MBI	1	0.90	0.95	0.85	0.61	−0.44	−0.37	−0.46
PWLS item	0.64	0.68	0.68	0.43	0.59	−0.39	−0.36	−0.42
EE single item	0.81	0.92	0.86	0.52	0.61	−0.42	−0.37	−0.42
DP single item	0.73	0.79	0.84	0.84	0.36	−0.24	−0.19	−0.29
Combined EE+DP items	0.90	1	0.86	0.75	0.59	−0.40	−0.34	−0.42
				2017 Residents				
Full MBI	1	0.90	0.95	0.84	0.61	−0.46	−0.38	−0.47
PWLS item	0.65	0.68	0.69	0.44	0.61	−0.43	−0.40	−0.44
EE single item	0.82	0.92	0.87	0.54	0.61	−0.42	−0.38	−0.42
DP single item	0.73	0.79	0.85	0.85	0.36	−0.31	−0.24	−0.35
Combined EE+DP items	0.90	1	0.86	0.76	0.59	−0.43	−0.36	−0.43

MBI indicates Maslach Burnout Inventory; EE, emotional exhaustion subscale of the MBI; DP, depersonalization subscale of the MBI; PWLS, Physician Work Life Study.
All correlations are statistically significant at $P < .01$.

+] = 3.8, 95% confidence interval [CI] 3.2−4.4; LR − = 0.38, 95% CI 0.32−0.42) compared with the full MBI in 2016 and similar values in 2017. The combined EE +DP screening questions generated a burnout prevalence of 53% in both 2016 and 2017. In 2016, for the combined EE+DP screen, the specificity was 87% and sensitivity was 84% (LR+ = 6.4, 95% CI 5.3−7.8; LR− = 0.18, 95% CI 0.16−0.21), with similar values in 2017. Predictive values in both years ranged between 81% and 89%.

Table 2 presents correlations among different measures of burnout, the MBI, and measures of related constructs. The top row presents results for the full MBI, which as the industry standard presents a basis of comparison for the 1- and 2-item screens presented in subsequent rows. The single PWLS item correlations with widely used measures of stress, mindfulness, resilience, and self-compassion were similar to the full MBI, statistically significant, and in expected directions. Correlations between the combined EE and DP items and other variables were also similar to the full MBI.

DISCUSSION

This study of large national samples of pediatric residents over 2 years showed that 3 single-item screens, particularly the combined EE+DP screening questions, generated similar prevalence estimates for burnout as the full MBI. The single EE and DP items showed moderate to strong correlations with the EE subscale and the DP subscale, respectively. Furthermore, the single items also had similar, significant correlations with other variables associated with burnout: stress, mindfulness, resilience, and self-compassion. These findings extend previous research by focusing specifically on pediatric residents and testing both PWLS items and the single items for EE and DP against the MBI and its subscales, as well as by assessing their correlations with other factors thought to be associated with burnout.

Although the MBI has been widely used among physicians,[15] its length limits its use for repeated assessment of burnout and other risk and protective factors in large groups of individuals. West et al[3] demonstrated that the 2-item measure of burnout could identify the likelihood of high burnout with a Spearman's rank correlation coefficient of 0.76 to 0.83 across the groups of individuals tested, which is similar to the correlations we found in both years between single items for each domain (EE and DP). The single item from the PWLS has been reportedly correlated to emotional exhaustion in a large sample of residents of different specialties with $r = 0.64$,[16] which echoes our findings in pediatric residents.

Dolan et al[4] reported on the test parameters of the PWLS screening item relative to the full MBI, but the reported prevalence of burnout in their study was almost half of what was found in our study, which can alter the negative and positive predictive values of this screen. Similar to our findings of high correlations with measures of stress, mindfulness, resilience, and self-compassion, West et al[17] found that the EE+DP items exhibited a strong association with outcomes of suicidality in medical

students, residents, and practicing surgeons. Our data are also consistent with Waddimba et al,[5] who reported that the 3 single items varied in effectiveness as screeners for the EE and DP dimensions of burnout; for example, the DP item alone generated an estimated prevalence of burnout of 34% in 2016, substantially below the prevalence of 56% generated by the full MBI, and it showed lower correlations with other variables, such as stress, mindfulness, resilience, and self-compassion, than the PWLS item, the EE item alone, or the combination of EE and DP. In our study, the combined EE+DP item performed better than either single item alone compared with the full MBI.

The use of a single- or 2-item screen is appealing as a substitute for the full MBI, and, based on our findings replicated across 2 years, these measures can be considered to measure epidemiologic trends over time in pediatric residency programs. Pediatric program directors may opt to use the PWLS item or the 2-item (EE+DP) screening questions to gain insight into the magnitude of the problem in their programs and response to programmatic interventions. However, no screening test or even the industry standard, MBI, can perfectly measure risk in individual residents, and if program directors are concerned about the potential for depression, substance abuse, or suicidality, then additional assessments should be considered.[18]

This study is limited by having been conducted among pediatric residents who were part of training programs in the Pediatric Resident Burnout-Resilience Study Consortium, and it is unclear whether there was a non-response bias. This study assessed burnout, not consequences of burnout such as suicidality, substance abuse, or medical errors. The study included only pediatric residents, and it is not clear that results can be generalized to students, other specialties, or practicing pediatricians. It is also not clear that these results can be generalized to other groups of health professionals with markedly different rates of burnout. Finally, these questions were included as part of a larger survey on burnout in pediatric residents, and the items may function differently when used alone or in other contexts.

CONCLUSIONS

The single PWLS item and the combined EE+DP screening items have good correlations with variables thought to be associated with burnout, and they generate prevalence estimates similar to those of the full MBI. They may be considered for tracking epidemiologic trends in burnout and the response to programmatic interventions in pediatric residencies.

ACKNOWLEDGMENTS

Pediatric Resident Burnout-Resilience Study Consortium in 2018: Megan McCabe, Albert Einstein College of Medicine (Montefiore); Molly Shane; Sharon Unti, Ann & Robert H. Lurie Children's Hosp of Chicago; Sanghamitra Misra, Baylor College of Medicine (Houston); Colin Sox, Boston Children's Hospital/Boston Medical Center; Violet Borowitcz, Carillion Clinic - Virginia Tech Carillion School of Medicine; Sydney Primis, Carolinas Medical Center; Ross Myers, Kathy Mason, Case Western Reserve University/University Hospital Case Medical Center/Rainbow Babies and Children's Hospital; Miriam Stewart, Lisa Zaoutis, Oana Tomescu, Children's Hospital of Philadelphia; Chuck Schubert, Cincinnati Children's Hospital Medical Center/University of Cincinnati College of Medicine; Frances Zappalla, Crozer-Chester Medical Center; Kim Gifford, Dartmouth Hitchcock Medical Center; Betty Staples, Duke University Hospital Medical Center; Jerry Rushton, Adam Hill, Kim Schneider, Indiana Univeristy School of Medicine; Kathleen Donnelly, Inova Fairfax Medical Campus/Inova Children's Hospital; Jordan Watson, Jefferson Medical College/duPont Hospital for Children; Janet R. Serwint, Johns Hopkins University School of Medicine; Jerussa Levy, Louisiana State University (New Orleans); Pam Dietz, Maine Medical Center; Jay Homme, Mayo Clinic College of Medicine (Rochester); Leah Phillipi, Medical College of Wisconsin; John Mahan, Nationwide Children's Hospital/Ohio State University; Jennifer DiPace, New York Presbyterian Hospital (Cornell Campus); Bridget VoigtRush University Medical Center; Nicole Paradise Black, Denise McIntyre, Shands Hospital at the University of Florida; Katherine Baker, St. Christopher's Hospital for Children; Priya Garg, Tufts Medical Center; Lauren Nassetta, University of Alabama Medical Center; Hilary McClafferty, University of Arizona Health Sciences Center; Albina Gogo, University of California (Davis) Health System; Kate Perkins, Jessica Lloyd, Alan Chin, University of California (Los Angeles) Medical Center; Deborah Rana, University of California (San Diego); Joseph Hageman, University of Chicago Corner Children's Hospital; Jennifer Reese, University of Colorado Denver Children's Hospital; Amanda Osta, University of Illinois College of Medicine at Chicago; Lisa Gilmer, University of Kansas School of Medicine; Tara McKinley, Michelle Stevenson, University of Louisville School of Medicine; Hilary Haftel, Thomas Saba, University of Michigan Medical Center; Maren Olson, Michael Pitt, University of Minnesota Medical School; Carmen Herrera; Lanier Lopez, University of New Mexico Children's Hospital; Keith Mather, University of Oklahoma College of Medicine (Tulsa); Casey Hester, University of Oklahoma Health Sciences Center (Oklahoma City); Stephanie Dewar, University of Pittsburgh Children's Hospital; Sophia Goslings, University of South Alabama Children's and Women's Hospital; Maneesh Batra, University of Washington-Seattle Children's Hospital; Grant Syverson, University of Wisconsin Hospitals and Clinics; Ann Burke, Wright State University.

REFERENCES

1. Shanafelt T, Hasan O, Dyrbye L, et al. Changes in burnout and satisfaction with work-life balance in physicians and the general US working population between 2011 and 2014. *Mayo Clin Proc*. 2015;90:1600–1613.

2. Maslach C, Jackson S, Leiter M. *Maslach Burnout Inventory Manual*. 3rd ed. Palo Alto, Calif: Consulting Psychologists Press; 1996.

3. West CP, Dyrbye LN, Sloan JA, et al. Single item measures of emotional exhaustion and depersonalization are useful for assessing burnout in medical professionals. *J Gen Intern Med*. 2009;24:1318–1321.

4. Dolan ED, Mohr D, Lempa M, et al. Using a single item to measure burnout in primary care staff: a psychometric evaluation. *J Gen Intern Med*. 2015;30:582–587.

5. Waddimba AC, Scribani M, Nieves MA, et al. Validation of single-item screening measures for provider burnout in a rural health care network. *Eval Health Prof*. 2016;39:215–225.

6. Batra M, Kemper KJ, Serwint JR, et al. Burnout in pediatric residents: a national survey to inform future interventions (platform presentation). *Acad Pediatr*. 2017;17:e39–e40.

7. Batra M, Kemper KJ, Mahan JD, et al. Promoting resilience among pediatric trainees and practitioners Paper presented at: In: Pediatric Academic Societies Meeting; May 6, 2018; Toronto.

8. Kemper KJ, McClafferty H, Wilson PM, et al. Do mindfulness and self-compassion predict burnout in pediatric residents? *Acad Med*. 2018; in press.

9. Dyrbye LN, Massie Jr FS, Eacker A, et al. Relationship between burnout and professional conduct and attitudes among US medical students. *JAMA*. 2010;304:1173–1180.

10. 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;172:1377–1385.

11. Cohen S. Perceived stress in a probability sample of the United States. In: Spacapan S, Oskamp S, eds. *The Claremont Symposium on Applied Social Psychology: The Social Psychology of Health*, Thousand Oaks, Calif: Sage Publications; 1988:31–67.

12. Feldman G, Hayes A, Kumar S, et al. Mindfulness and emotion regulation: the development and initial validation of the Cognitive and Affective Mindfulness Scale-Revised (CAMS-R). *J Psychopathol Behav Assess*. 2007;29:177–190.

13. Smith BW, Dalen J, Wiggins K, et al. The brief resilience scale: assessing the ability to bounce back. *Int J Behav Med*. 2008;15:194–200.

14. Raes F, Pommier E, Neff KD, et al. Construction and factorial validation of a short form of the self-compassion scale. *Clin Psychol Psychother*. 2011;18:250–255.

15. Rafferty JP, Lemkau JP, Purdy RR, et al. Validity of the Maslach Burnout Inventory for family practice physicians. *J Clin Psychol*. 1986;42:488–492.

16. Rohland BM, Kruse GR, Rohrer JE. Validation of a single-item measure of burnout against the Maslach Burnout Inventory among physicians. *Stress Health*. 2004;20:75–79.

17. West CP, Dyrbye LN, Satele DV, et al. Concurrent validity of single-item measures of emotional exhaustion and depersonalization in burnout assessment. *J Gen Intern Med*. 2012;27:1445–1452.

18. Omurtag A, Fenton AA. Assessing diagnostic tests: how to correct for the combined effects of interpretation and reference standard. *PLoS ONE*. 2012;7:e52221.