


# Who Benefits Most? A Multisite Study of Coaching and Resident Well-being



Kerri Palamara, MD<sup>1</sup> , Jacqueline T. Chu, MD<sup>1</sup>, Yuchiao Chang, PhD<sup>1</sup>, Liyang Yu, MS<sup>1</sup>, Dominique Cosco, MD<sup>2</sup>, Stacy Higgins, MD<sup>3</sup>, Asher Tulskey, MD<sup>4</sup>, Ronda Mourad, MD<sup>5</sup>, Simran Singh, MD<sup>5</sup>, Karen Steinhauser, PhD<sup>6</sup>, and Karen Donelan, ScD, EdM<sup>7</sup>

<sup>1</sup>Center for Physician Well-Being, Department of Medicine, Massachusetts General Hospital, Fruit Street, Boston, MA, USA; <sup>2</sup>Department of Medicine, Washington University at St. Louis, St. Louis, MO, USA; <sup>3</sup>Department of Medicine, Emory University, Atlanta, GA, USA; <sup>4</sup>Department of Medicine, Boston Medical Center, Boston, MA, USA; <sup>5</sup>Department of Medicine, University Hospitals Cleveland, Cleveland, OH, USA; <sup>6</sup>Departments of Population Health Sciences and Medicine, Duke University, Durham, NC, USA; <sup>7</sup>Mongan Institute Health Policy Center, Massachusetts General Hospital, Boston, MA, USA.

**BACKGROUND:** Coaching has been shown to improve resident well-being; however, not all benefit equally.

**OBJECTIVE:** Assess predictors of changes in resident physician well-being and burnout in a multisite implementation of a Professional Development Coaching Program.

**DESIGN:** Pre- and post-implementation surveys administered to participant cohorts at implementation sites in their intern year. Effect size was calculated comparing pre- and post-intervention paired data.

**PARTICIPANTS:** In total, 272 residents in their intern year at five internal medicine residency programs (Boston Medical Center, University Hospitals Cleveland Medical Center, Duke University, Emory University, Massachusetts General Hospital). Analyses included 129 residents with paired data.

**INTERVENTIONS:** Interns were paired with a faculty coach trained in positive psychology and coaching skills and asked to meet quarterly with coaches.

**MAIN MEASURES:** Primary outcomes included Maslach Burnout Inventory depersonalization (DP) and emotional exhaustion (EE) subscales, and the PERMA well-being scale. Key predictors included site, demographics, intolerance of uncertainty, hardiness-resilience, gratitude, and coping. Program moderators included were reflection, goal setting, and feedback.

**KEY RESULTS:** Well-being (PERMA) changed from baseline to follow-up in all participants; females showed a decline and males an increase (−1.41 vs. .83,  $p = 0.04$ ). Self-reflection was associated with positive change in PERMA (mean positive change 1.93,  $p = 0.009$ ). Burnout (EE) declined in non-Hispanic white residents vs. Black/Asian/Hispanic/other residents (−1.86,  $p = 0.021$ ). Burnout improved with increased goal setting.

**CONCLUSION:** Coaching programs should consider tailored approaches to support residents whose well-being is

impacted by gender and/or race, and who have higher intolerance of uncertainty and lower resilience at baseline. Coaching skills of goal setting and reflection may positively affect interns and teach coping skills.

**KEY WORDS:** coaching; well-being; graduate medical education; faculty development; burnout.

## Abbreviations

|        |   |
|--------|---|
| PDCP   | Professional Development Coaching Program                                 |
| DP     | Depersonalization   |
| EE     | Emotional exhaustion  |
| MBI    | Maslach Burnout Inventory   |
| PERMA  | Positive Emotion, Engagement, Relationships, Meaning, and Accomplishments |
| MGH    | Massachusetts General Hospital  |
| IUS    | Intolerance of uncertainty  |
| HRS    | Hardiness resilience score  |
| DRS-15 | Dispositional-Resilience Scale-15   |
| GR-2   | 2-question Gratitude Questionnaire  |
| GQ-6   | 6-question Gratitude Questionnaire  |
| MOCS3  | Measurement of Current Status   |
| RCT    | Randomized controlled trial   |
| URiM   | Under-represented in medicine   |

J Gen Intern Med 37(3):539–47  
DOI: 10.1007/s11606-021-06903-5  
© Society of General Internal Medicine 2021

## INTRODUCTION

The stresses and challenges of medicine residency threaten learners' well-being and burnout risk is high. Concerns about resident well-being have driven the National Academy of Medicine's Committee on Systems Approaches to Improve Patient Care by Supporting Clinician Well-Being call to action, where they urged healthcare organizations to adopt and apply principles to improve the work environment and balance job demands and job resources.<sup>1, 2</sup> A key recommendation included professional development of faculty to promote positive role model behaviors.<sup>3, 4</sup>

Well-documented challenges affecting residents include excessive workload, inefficient practice, work-life integration,

---

**Prior Presentations** While the tools and implementation process used in the coaching program have been shared in numerous workshops at national conferences (Society of General Internal Medicine, Association of Program Directors in Internal Medicine, American College of Physicians) and invited grand rounds and faculty development workshops, the data included in this paper has not been previously presented.

---

Received November 20, 2020

Accepted April 30, 2021

Published online June 7, 2021

lack of autonomy and control, loss of meaning in work, disconnection with and structural inequities within organizational culture, and high stress environment.<sup>5–8</sup> These challenges affect the learning environment, as well as resident motivation, decision-making, quality of care, and communication at work.<sup>9, 10</sup> In response, residents use a variety of coping strategies, often without prior experience or reflection to help them understand what may or may not work for them. Physicians with lower resilience and who report lacking the skills to help them function well under stress are more likely to experience burnout.<sup>10, 11</sup>

Learning effective coping strategies may improve burnout. Prior research indicates that both individual-focused and systems-based strategies can improve well-being.<sup>4, 12</sup> Few examples exist of programmatic interventions that combine both approaches as levers of change. Coaching is an example of such a resource to support residents in managing stress and adapting to change through development of coping behaviors, increasing positive emotions, and fostering positive relationships with faculty.<sup>5, 13–15</sup> Coaching is a relationship between an individual coachee and a coach with a specific purpose to help the coachee design and achieve professional or personal goals.<sup>14, 15</sup> Coaching allows for both structured and unstructured conversation, emphasizes active listening, and enhances strengths.<sup>13–16</sup> The coaching literature suggests that the process of coaching is highly complex, and multiple moderating and mediating factors in the individual and the approach contribute to the success of a coaching intervention (coachee motivations and expectations, the coaching relationship, using an internal vs. external coach, coachee self-efficacy).<sup>13</sup> To our knowledge, the exploration of these mechanisms is limited in the literature.

By engaging faculty trained in positive psychology and coaching skills, we endeavored to develop and disseminate a systems-based solution focused on individuals to improve physician well-being in the Professional Development Coaching Program (PDCP). Our initial work reported on a single-center positive psychology coaching program and demonstrated that in those residents who participated and reported a positive coaching experience, their opportunities to reflect increased, as did coping and relationship skills.<sup>17</sup> Those residents with poor or fair opportunities to reflect exhibited higher emotional exhaustion scores on the Maslach Burnout Inventory (MBI).<sup>18, 19</sup> While it was clear that some benefitted from the program more than others, program size limited a more comprehensive assessment of other factors that might be used to predict who came into the program with better coping and resiliency practices and might benefit more or less from the program.

In 2016–2017, we expanded the PDCP to five institutions to assess the generalizability of a positive psychology coaching program. In this study, we explored how

coachee attributes and behaviors at baseline contribute to their well-being and impact their experience with a coaching program. In baseline data of participants, there was no correlation between burnout (using MBI) and a well-being scale known as PERMA (Positive Emotion, Engagement, Relationships, Meaning, and Accomplishments).<sup>20, 21</sup> This prompted exploration of the associations between participant characteristics and changes in well-being and burnout in a multivariate model to assess how to maximize the benefit of the coaching program. Using primary *outcome variables* of MBI and PERMA, and controlling for resident coachee demographics and program site, our predictors included skills the PDCP was designed to enhance through the intervention (intolerance of uncertainty, hardiness/resilience, gratitude, and coping) and previously identified outcomes of the coaching relationship (self-reported practice of goal setting, reflection, and feedback).

## METHODS

### Design

This was a prospective, pre-post study of 272 interns in internal medicine training programs at five institutions (Boston Medical Center, University Hospitals Cleveland Medical Center, Duke University, Emory University, and Massachusetts General Hospital) who were enrolled in the previously described PDCP during the 2016–2017 academic year.<sup>18</sup> Participants were surveyed in October of intern year to assess baseline burnout and well-being, as well as modifiers/mediators and key skills necessary to achieve well-being. October was chosen as interns would have 3+ months of work experience to ensure measurement reflected their internship baseline, rather than July which might mistakenly reflect well-being and burnout from medical school. They were then surveyed with the same tools in May of intern year.

Survey data are available for 235 interns at baseline; paired data are available for 129 interns. Effect size was calculated comparing pooled baseline samples to pooled end of year samples and calculated comparing pre- and post-intervention paired data.

### Intervention

The PDCP has been previously described.<sup>18</sup> All interns were paired with a faculty coach at their institution and outside their field of specialty interest. Coaches were provided a 3-h skills-based training in positive psychology and coaching principles and given curricular guides for 1:1 meetings with their coachees. Each coach was assigned 1–2 coachees; pairs were expected to meet quarterly. A typical coaching meeting was 45–60 min in length. Themes and skills to practice reflection, processing

feedback, and goal setting were woven throughout the program. Coaching meetings focused on strengths evaluation, resiliency, and constructing a framework of positive well-being (PERMA).<sup>18</sup> Appendix Figure 1 outlines the curriculum. Survey and program participation were voluntary; there was no consequence for not meeting with their coach. Dr. Palamara led the faculty coach trainings at each site. Each site had an appointed champion(s) present for trainings, creating a “train the trainer” model for sustainability beyond the pilot. Champions were in regular communication with Dr. Palamara to oversee the implementation and administration of the program. Site-specific curricular materials were created and stored in a cloud-based system for coaches to access for meetings.

Surveys were administered by a research team at MGH. Sites did not have access to the data but were asked to encourage survey completion at their site. Sites determined the use of incentives for survey participation; the amount and type varied.

**Measurements**

Survey metrics were grouped into 3 categories: (1) *study outcomes*—MBI score for burnout (7-point frequency [0–6], 9 items for EE and 5 items for DP, range 0–54 for EE and 0–30 for DP); PERMA score as a metric of well-being (5-point frequency [1–5], 15 items, range 0–75); (2) *baseline predictors (modifiers)*—program site and resident demographics, the Intolerance of Uncertainty Scale (IUS, 5-point frequency [1–5], 12 items, range 0–60); HRS or hardiness-resilience scoring (DRS-15, 4-point frequency [0–3], 15 items, range 0–45), gratitude (GR-2, 7-point frequency [1–7], 2 items from the GQ-6 scale, range 0–14) and a 3-item measure of coping (Measurement of Current Status; MOCS3, 5-point frequency [1–5], 3 items, range 0–15), self-reported reflection and goal setting; and (3) *post program predictors (mediators)*—self-reported reflection, goal setting during follow-up.<sup>19, 21–26</sup>

**Statistical Analysis**

Participant characteristics were summarized for all respondents to the baseline survey grouped by completion of follow-up survey. We then restricted the remaining analysis to those who answered both baseline and follow-up surveys. Due to considerable variation in resident survey response and participation level by site, we controlled for program site in all analyses. For the primary outcomes of burnout and well-being, a positive change in PERMA indicates improvement in well-being. For EE and DP, a negative change indicates improvement in burnout. For independent variables, scoring measures that were continuous (HRS, IUS, MOCS, and GR2) were divided the scores into quartiles and dichotomized further as top quartile vs. lower quartiles.

**Table 1 Comparison of Participant Characteristics Between Those Who Did Not Complete Follow-up Survey and Those Who Did Complete the Follow-up Survey**

| Respondents (pre and post)       | Lost to follow-up (pre only) |       | p   |       |         |
|----------------------------------|------------------------------|-------|-----|-------|---------|
|                                  | N                            | %     | N   | %     |         |
| All                              | 106                          | 100.0 | 129 | 100.0 |         |
| Gender                           |                              |       |     |       | 0.19    |
| Male                             | 45                           | 47.4  | 72  | 56.3  |         |
| Female                           | 50                           | 52.6  | 56  | 43.8  |         |
| Race/ethnicity                   |                              |       |     |       | 0.17    |
| Non-Hispanic white               | 53                           | 57.0  | 81  | 63.3  |         |
| Black                            | 4                            | 4.3   | 5   | 3.9   |         |
| Asian                            | 32                           | 34.4  | 28  | 21.9  |         |
| Other                            | 2                            | 2.2   | 7   | 5.5   |         |
| Hispanic                         | 2                            | 2.2   | 7   | 5.5   |         |
| Site                             |                              |       |     |       | <0.0001 |
| A                                | 22                           | 20.8  | 21  | 16.3  |         |
| B                                | 29                           | 27.4  | 16  | 12.4  |         |
| C                                | 12                           | 11.3  | 22  | 17.1  |         |
| D                                | 15                           | 14.2  | 63  | 48.8  |         |
| E                                | 28                           | 26.4  | 7   | 5.4   |         |
| Goal setting                     |                              |       |     |       | 1.00    |
| Never/rarely                     | 38                           | 40.0  | 50  | 40.0  |         |
| Sometimes/frequently             | 57                           | 60.0  | 75  | 60.0  |         |
| Opportunity to reflect           |                              |       |     |       | 0.49    |
| Poor/fair                        | 24                           | 23.3  | 22  | 17.1  |         |
| Good                             | 56                           | 54.4  | 75  | 58.1  |         |
| Excellent                        | 23                           | 22.3  | 32  | 24.8  |         |
| Receiving feedback               |                              |       |     |       | 0.64    |
| Poor/fair                        | 21                           | 20.6  | 32  | 24.8  |         |
| Good                             | 65                           | 63.7  | 81  | 62.8  |         |
| Excellent                        | 16                           | 15.7  | 16  | 12.4  |         |
| Hardiness-resilience score       |                              |       |     |       | 0.81    |
| Low (<Q4)                        | 66                           | 75.9  | 92  | 77.3  |         |
| High (Q4)                        | 21                           | 24.1  | 27  | 22.7  |         |
| Intolerance of uncertainty score |                              |       |     |       | 0.094   |
| Low (<Q4)                        | 62                           | 69.7  | 98  | 79.7  |         |
| High (Q4)                        | 27                           | 30.3  | 25  | 20.3  |         |
| MOCS: Coping score               |                              |       |     |       | 0.37    |
| Low (<Q4)                        | 72                           | 76.6  | 89  | 71.2  |         |
| High (Q4)                        | 22                           | 23.4  | 36  | 28.8  |         |
| Gratitude score                  |                              |       |     |       | 0.72    |
| Low (<Q4)                        | 71                           | 75.5  | 97  | 77.6  |         |
| High (Q4)                        | 23                           | 24.5  | 28  | 22.4  |         |
| Burnout                          |                              |       |     |       | 0.25    |
| No                               | 46                           | 63.9  | 66  | 55.5  |         |
| Yes                              | 26                           | 36.1  | 53  | 44.5  |         |
| PERMA overall score              |                              |       |     |       | 0.10    |
| Low (36–59)                      | 36                           | 39.1  | 60  | 50.4  |         |
| High (60–75)                     | 56                           | 60.9  | 59  | 49.6  |         |
| Meetings                         |                              |       |     |       | <0.0001 |
| 0–1                              | 28                           | 26.4  | 28  | 21.7  |         |
| 2–3                              | 21                           | 19.8  | 73  | 56.6  |         |
| 4–5                              | 3                            | 2.8   | 26  | 20.2  |         |
| Missing                          | 54                           | 50.9  | 2   | 1.6   |         |

Abbreviations: LTFU, lost to follow-up; Q, quartile; MOCS, Measurement of Current Status; PERMA, Positive Emotion, Engagement, Relationships, Meaning, and Accomplishments

We performed bivariate analysis for the primary outcomes to measure three differences in scores outcomes for participants who answered both baseline and follow-up surveys. For each predictor, we performed a one-way ANCOVA test while adjusting for baseline score of each

respective outcome. The sample size limited our ability to use all possible predictors in our models; we used our pre-specified predictors based on prior knowledge and bivariate analyses to perform multiple linear regression modeling for the three outcomes separately to examine predictors including baseline score, demographics (gender, race, site), goal setting, HRS, GR-2, IUS at baseline, and reflection at follow-up. A two-sided  $p$ -value  $<0.05$  was considered statistically significant. All analyses were conducted using SAS version 9.4 (SAS Institute, Cary, NC).

## RESULTS

A total of 235 participants answered the baseline survey. Of those, 129 responded to the post-survey. In Table 1, we compared participant characteristics and survey responses of responders with non-responders. To examine selection bias,

we compared demographics between responders and non-responders and found no significant differences in baseline characteristics. As previously noted, we controlled for site in all multivariate analyses.

In Table 2 and Figure 1, we show the bivariate analysis to assess associations of each predictor with the differences in scores between baseline and follow-up measures of primary outcomes PERMA and MBI subscales of EE and DP. Among all respondents, there were declines in PERMA, MBI, and EE scores, and a small increase in DP. The change is calculated from subtracting baseline score from the follow-up score. A positive change in PERMA and a negative change in EE or DP indicate improvement.

PERMA changed from baseline to follow-up in all participants, with significant differences between male and female coaches—males indicated a small positive change, females a negative change in PERMA (males 0.83 vs. females  $-1.41$ ,  $p=0.04$ ). Of several other predictors, few were significantly

**Table 2 Examination of Change (Post-Pre) in Outcome Scores for PERMA, Emotional Exhaustion, and Depersonalization for Each Predictor. This Analysis Is Limited to the Number of Persons Providing Complete Data**

|                                  | Diff <sup>*</sup> in PERMA |       |       | Diff <sup>*</sup> in emotional exhaustion |       |       | Diff <sup>*</sup> in depersonalization |       |       | Diff <sup>*</sup> in MBI** score |        |       |
|----------------------------------|----------------------------|-------|-------|---|-------|-------|--|-------|-------|----------------------------------|--------|-------|
|                                  | N                          | Mean  | p***  | N   | Mean  | p***  | N                                      | Mean  | p***  | N                                | Mean   | p***  |
| All                              | 116                        | -0.16 |       | 118                                       | -0.92 |       | 117                                    | 0.22  |       | 116                              | -0.73  |       |
| Gender                           |                            |       | 0.040 |   |       | 0.50  |  |       | 0.95  |                                  |        | 0.73  |
| Male                             | 65                         | 0.83  |       | 64  | -0.50 |       | 63                                     | -0.06 |       | 63                               | -0.68  |       |
| Female                           | 51                         | -1.41 |       | 53  | -1.62 |       | 53                                     | 0.49  |       | 52                               | -1.08  |       |
| Race                             |                            |       | 0.28  |   |       | 0.021 |  |       | 0.64  |                                  |        | 0.10  |
| Black/Asian/Hispanic/other       | 43                         | -0.47 |       | 43  | 0.47  |       | 41                                     | 0.22  |       | 41                               | 0.59   |       |
| Non-Hispanic white               | 73                         | 0.03  |       | 74  | -1.86 |       | 75                                     | 0.17  |       | 74                               | -1.66  |       |
| Site                             |                            |       | 0.23  |   |       | 0.81  |  |       | 0.70  |                                  |        | 0.75  |
| A                                | 20                         | -1.70 |       | 21  | -1.38 |       | 21                                     | 0.38  |       | 21                               | -1.00  |       |
| B                                | 12                         | 1.42  |       | 14  | -1.36 |       | 14                                     | 0.43  |       | 14                               | -0.93  |       |
| C                                | 18                         | 1.72  |       | 20  | -0.30 |       | 20                                     | -0.20 |       | 20                               | -0.50  |       |
| D                                | 60                         | -0.57 |       | 61  | -0.66 |       | 60                                     | 0.43  |       | 59                               | -0.29  |       |
| E                                | 6                          | 0.33  |       | 2   | -7.00 |       | 2                                      | -5.00 |       | 2                                | -12.00 |       |
| Opportunity to reflect (pre)     |                            |       | 0.95  |   |       | 0.57  |  |       | 0.42  |                                  |        | 0.57  |
| Poor/fair                        | 20                         | 0.00  |       | 20  | -1.70 |       | 21                                     | 0.52  |       | 20                               | -1.05  |       |
| Good                             | 68                         | 0.07  |       | 68  | -0.66 |       | 66                                     | -0.18 |       | 66                               | -0.94  |       |
| Excellent                        | 28                         | -0.82 |       | 30  | -0.97 |       | 30                                     | 0.90  |       | 30                               | -0.07  |       |
| Opportunity to reflect (post)    |                            |       | 0.009 |   |       | 0.10  |  |       | 0.21  |                                  |        | 0.090 |
| Poor/fair                        | 18                         | -1.72 |       | 18  | 0.78  |       | 19                                     | 0.16  |       | 18                               | 1.06   |       |
| Good                             | 68                         | -0.66 |       | 69  | -0.49 |       | 67                                     | 0.64  |       | 67                               | 0.06   |       |
| Excellent                        | 30                         | 1.93  |       | 31  | -2.84 |       | 31                                     | -0.65 |       | 31                               | -3.48  |       |
| Goal setting (pre)               |                            |       | 0.77  |   |       | 0.056 |  |       | 0.099 |                                  |        | 0.046 |
| Never/rarely                     | 47                         | 0.60  |       | 47  | -3.45 |       | 47                                     | -1.34 |       | 46                               | -4.80  |       |
| Sometimes/frequently             | 69                         | -0.67 |       | 71  | 0.76  |       | 70                                     | 1.27  |       | 70                               | 1.94   |       |
| Goal setting (post)              |                            |       | 0.29  |   |       | 0.96  |  |       | 0.45  |                                  |        | 0.69  |
| Never/rarely                     | 20                         | -1.05 |       | 22  | -1.59 |       | 22                                     | 0.14  |       | 22                               | -1.45  |       |
| Sometimes/frequently             | 96                         | 0.03  |       | 96  | -0.76 |       | 95                                     | 0.24  |       | 94                               | -0.56  |       |
| Hardiness-resilience score       |                            |       | 0.29  |   |       | 0.034 |  |       | 0.38  |                                  |        | 0.074 |
| Low (<Q4)                        | 89                         | -0.27 |       | 87  | -0.55 |       | 86                                     | 0.22  |       | 86                               | -0.42  |       |
| High (Q4)                        | 24                         | 0.17  |       | 25  | -2.28 |       | 26                                     | -0.23 |       | 25                               | -2.44  |       |
| Gratitude score                  |                            |       | 0.31  |   |       | 0.62  |  |       | 0.081 |                                  |        | 0.29  |
| Low (<Q4)                        | 88                         | -0.01 |       | 91  | -1.19 |       | 90                                     | 0.29  |       | 89                               | -0.96  |       |
| High (Q4)                        | 28                         | -0.61 |       | 27  | 0.00  |       | 27                                     | 0.00  |       | 27                               | 0.00   |       |
| Intolerance of uncertainty score |                            |       | 0.79  |   |       | 0.002 |  |       | 0.096 |                                  |        | 0.005 |
| Low (<Q4)                        | 93                         | -0.18 |       | 93  | -1.57 |       | 93                                     | 0.18  |       | 92                               | -1.36  |       |
| High (Q4)                        | 23                         | -0.04 |       | 23  | 1.13  |       | 22                                     | 0.05  |       | 22                               | 0.91   |       |
| Coping score                     |                            |       | 0.90  |   |       | 0.26  |  |       | 0.86  |                                  |        | 0.53  |
| Low (<Q4)                        | 81                         | 0.31  |       | 84  | -1.25 |       | 82                                     | -0.26 |       | 82                               | -1.60  |       |
| High (Q4)                        | 35                         | -1.23 |       | 34  | -0.09 |       | 35                                     | 1.34  |       | 34                               | 1.35   |       |

\*Difference is calculated post-pre

\*\*MBI score is sum of emotional exhaustion and depersonalization

\*\*\*p-value adjusted for baseline score

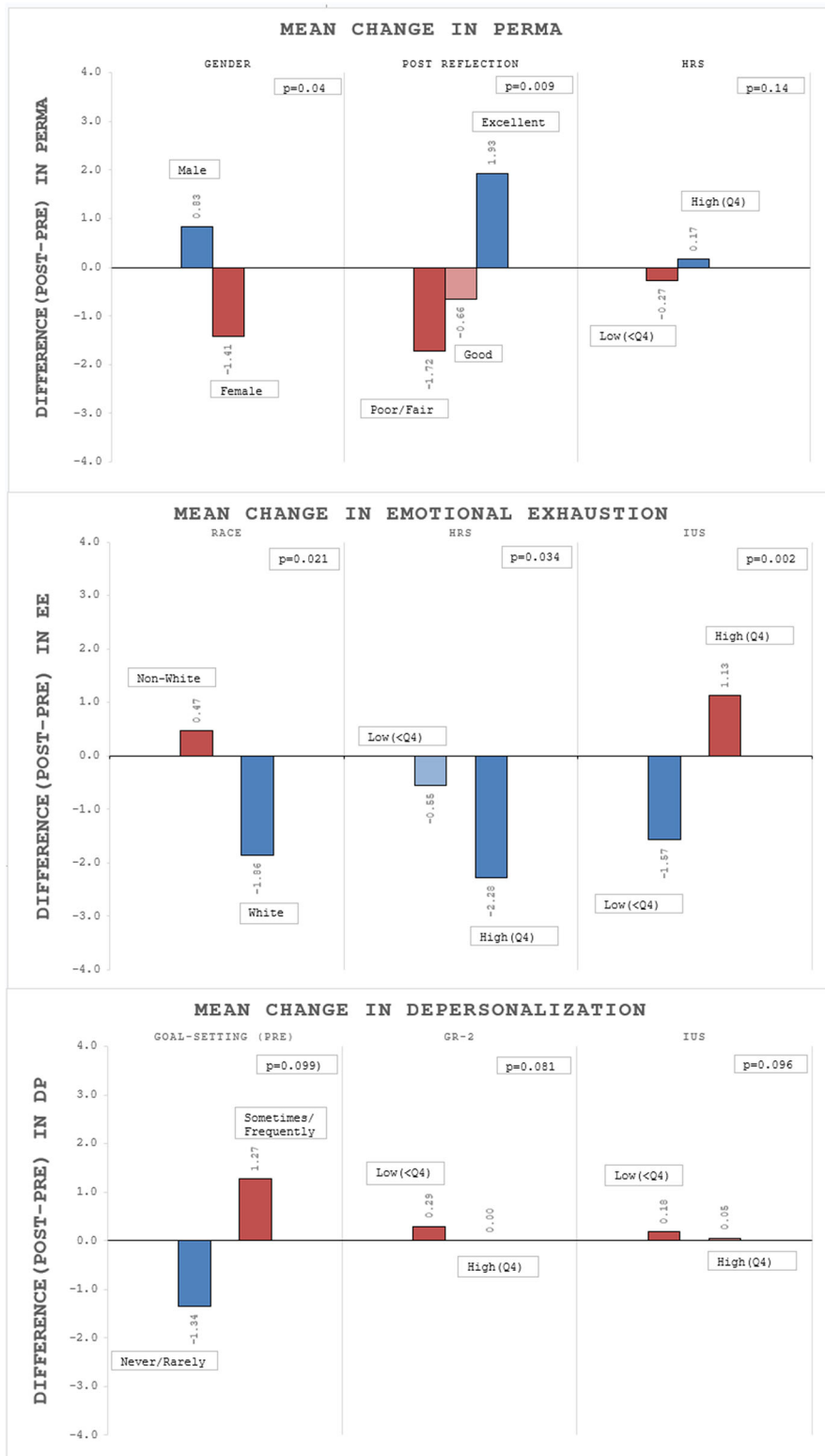


Figure 1 Graphical display of selected predictors that are associated with change in PERMA, emotional exhaustion, and depersonalization scores. PERMA, Positive Emotion, Engagement, Relationships, Meaning, and Accomplishments; HRS, Hardiness-Resilience Scale; IUS, intolerance of uncertainty; GR2, Gratitude Scale. For PERMA, males indicated an average positive change (0.83), and females a negative change (-1.41). Participants who reported their post-program opportunity to reflect was “Excellent” showed mean positive change in PERMA (1.93) while those who answered “Poor/Fair” and “Good” had mean decreases (-1.72, -0.66). Examining burnout in terms of EE, those who were non-Hispanic white had a mean decrease of EE (-1.86). Higher HRS and lower IUS at baseline are associated with decreases in EE on average (-2.28, -1.57). Those who reported that they “Never/Rarely” practiced goal setting at baseline had improved DP with average reduction of -1.34. Lower GR2 and lower IUS at baseline were associated with increases in DP (0.29, 0.18).

associated with a change in PERMA except for participant-reported self-reflection at follow-up, a key program skill. Participants who reported their post-program opportunity to reflect was “Excellent” showed mean positive change in well-being (PERMA) (1.93,  $p = 0.009$ ).

We also examined bivariate associations of changes in the MBI and its subscales. Examining burnout, those who were non-Hispanic white had a mean decrease of EE ( $-1.86$ ,  $p = 0.021$ ). Those who reported that they “Never/Rarely” practiced goal setting at baseline had improved EE and DP during the program year with a respective average reduction of  $-3.45$  ( $p = 0.056$ ) and  $-1.34$  ( $p = 0.099$ ). Also, for EE, higher HRS and lower IUS scores at baseline were associated with lower burnout on average ( $-2.28$ ,  $p = 0.034$ ;  $-1.57$ ,  $p = 0.002$ ).

Given our hypothesis that program site, participant characteristics, and other factors might influence participant experience and receptivity to program mediators, we conducted multiple linear regression modeling to examine these multiple factors in predicting our primary outcomes (see Table 3). Covariates included in these models included baseline scores, gender, race (Black/Asian/Hispanic/other vs. non-Hispanic white), site, post-opportunity to reflect, baseline goal setting, HRS, gratitude score, and IUS score, all of which were significant in bivariate analyses. In general, if a participant had higher well-being (PERMA) at baseline, there was less improvement seen ( $-0.407$ ,  $p < 0.001$ ). Those whose resilience improved as measured by higher HRS score (vs. lower at baseline) also experienced a positive change in well-being ( $3.474$ ,  $p = 0.008$ ). However, controlling for all factors,

females see a significant negative change ( $-2.475$ ,  $p = 0.012$ ) in well-being (PERMA) compared to males. Participants with improved post-reflection (report “Excellent” vs. poor/fair reflection at baseline) experienced a positive change ( $4.025$ ,  $p = 0.011$ ).

Participants experiencing higher burnout at baseline, measured by higher baseline EE and DP scores, were more likely to improve and experience lower burnout in the program year ( $-0.393$ ,  $p < 0.001$ ;  $-0.476$ ,  $p < 0.001$ ). Those who are Black/Asian/Hispanic/other were significantly more likely to have an increase in EE ( $2.608$ ,  $p = 0.046$ ) compared to non-Hispanic white participants. Higher HRS score at baseline is significantly associated with improvement in EE ( $-4.469$ ,  $p = 0.003$ ), whereas higher IUS score at baseline is significantly associated with more EE ( $6.022$ ,  $p = 0.0002$ ) and DP ( $2.555$ ,  $p = 0.025$ ) when compared to lower scores in HRS and IUS, respectively. Respondents with higher gratitude scores vs. lower scores at baseline are more likely to report an improvement in DP ( $-1.865$ ,  $p = 0.065$ ).

Finally, in another aspect of the survey, we included direct measures of the residents’ own perceptions of whether coaching had helped them cope with key stressors in internship year (work-life balance, cultural competence, working relationships, work hour restrictions, administrative burdens, self-confidence, and information processing). Appendix Table 1 reports whether interns “definitely” or “somewhat” agreed or disagreed with the following statement: “In previous research, the following have been noted as major challenges for trainees. For each one, please indicate if you believe the coaching

**Table 3 Modeling of Change (Post-Pre) in Outcome Scores for PERMA, Emotional Exhaustion (EE), and Depersonalization (DP) for Pre-specified Predictors**

| Predictors                       | Outcome Diff in PERMA |         | Diff in EE |         | Diff in DP |         | Diff in MBI |         |
|----------------------------------|-----------------------|---------|------------|---------|------------|---------|-------------|---------|
|                                  | Estimate              | p-value | Estimate   | p-value | Estimate   | p-value | Estimate    | p-value |
| Baseline score                   | -0.407                | <0.0001 | -0.393     | <0.0001 | -0.476     | <0.0001 | -0.430      | <0.0001 |
| Gender                           |                       |         |            |         |            |         |             |         |
| Male (REF)                       | -                     | -       | -          | -       | -          | -       | -           | -       |
| Female                           | -2.475                | 0.012   | -0.706     | 0.56    | 0.201      | 0.81    | -0.395      | 0.83    |
| Race                             |                       |         |            |         |            |         |             |         |
| Black/Asian/Hispanic/other       | -0.870                | 0.40    | 2.608      | 0.046   | 0.056      | 0.95    | 2.643       | 0.17    |
| Non-Hispanic white (REF)         | -                     | -       | -          | -       | -          | -       | -           | -       |
| Site                             |                       |         |            |         |            |         |             |         |
| A                                | -1.051                | 0.44    | -1.194     | 0.49    | -0.924     | 0.43    | -2.149      | 0.40    |
| B                                | 2.139                 | 0.20    | 0.289      | 0.88    | 0.292      | 0.83    | 0.586       | 0.84    |
| C                                | 1.777                 | 0.20    | -0.133     | 0.94    | -0.482     | 0.68    | -0.597      | 0.82    |
| D (REF)                          | -                     | -       | -          | -       | -          | -       | -           | -       |
| E                                | -0.040                | 0.99    | -5.933     | 0.20    | -5.753     | 0.068   | -11.58      | 0.093   |
| Opportunity to reflect (post)    |                       |         |            |         |            |         |             |         |
| Poor/fair (REF)                  | -                     | -       | -          | -       | -          | -       | -           | -       |
| Good                             | 0.898                 | 0.52    | -1.058     | 0.55    | -0.346     | 0.77    | -1.378      | 0.60    |
| Excellent                        | 4.025                 | 0.011   | -3.821     | 0.052   | -1.858     | 0.16    | -5.715      | 0.049   |
| Goal setting (pre)               |                       |         |            |         |            |         |             |         |
| Never/rarely                     | -0.067                | 0.95    | -2.228     | 0.094   | -1.598     | 0.076   | -3.750      | 0.058   |
| Sometimes/frequently (REF)       | -                     | -       | -          | -       | -          | -       | -           | -       |
| Hardiness-resilience score       |                       |         |            |         |            |         |             |         |
| Low (<Q4, REF)                   | -                     | -       | -          | -       | -          | -       | -           | -       |
| High (Q4)                        | 3.474                 | 0.008   | -4.469     | 0.003   | -1.021     | 0.30    | -5.579      | 0.012   |
| Gratitude score                  |                       |         |            |         |            |         |             |         |
| Low (<Q4, REF)                   | -                     | -       | -          | -       | -          | -       | -           | -       |
| High (Q4)                        | 1.573                 | 0.21    | 0.495      | 0.74    | -1.865     | 0.065   | -1.438      | 0.52    |
| Intolerance of uncertainty score |                       |         |            |         |            |         |             |         |
| Low (<Q4, REF)                   | -                     | -       | -          | -       | -          | -       | -           | -       |
| High (Q4)                        | -1.039                | 0.41    | 6.022      | 0.0002  | 2.555      | 0.025   | 8.626       | 0.0005  |

program has improved your ability to cope.” The strongest agreement was seen with self-confidence (89.3%), working relationships (89.1%), information processing (84.7%), cultural competence (75.7%), work hour restrictions (73.9%), and work-life balance (73.8%). The lowest agreement was with administrative burdens (60.5%).

## DISCUSSION

These data are taken from a multisite implementation of the PDCP, a program targeted at residents in medicine to improve well-being and decrease burnout through positive psychology coaching and the introduction of new skills in goal setting, reflection, and receiving feedback. These analyses reveal the complex relationships that are likely to exist in the lives of new physicians and in the global measures of burnout and well-being. If we look at the entire cohort, those outcomes do not change very much during this program year, though the overall minor changes mask significant differences by gender and race, as female residents are more likely to experience decreased well-being and residents who are Black/Asian/Hispanic/other are more likely to experience burnout (EE and DP). Those with higher scores in resilience and gratitude at baseline do better in our primary outcomes, and those with high intolerance of uncertainty experience significantly higher EE and DP.

The positive psychology component of this program allows for the opportunity to explore the modifiable traits that can impact an individual’s well-being and burnout, including effective response to stress and tolerance of uncertainty. Goal setting, reflection, and receiving feedback are core components of a coaching relationship. In our bivariate and multivariate analyses, the increased practice of reflection was significantly associated with improved well-being and goal setting practices with reduced burnout. This finding is consistent with our prior research, which linked high EE with reflecting less often.<sup>17</sup> Reflection is a core coaching skill emphasized in trainings; which helps coachees explore their experiences, motivations, positive emotions, and strengths.<sup>13, 14</sup> Based on this analysis, we propose that a coaching program can offer deliverables to support resident well-being, including time, space, and tools for setting goals, reflecting, and receiving feedback.

Interns in this study showed high resilience at baseline, similar to West et al.’s study.<sup>11</sup> The finding of better EE and PERMA in those with the highest resilience suggests learning skills to help cope with stress effectively may be useful to improve well-being for interns. Previously, the PDCP has been shown to provide coping skills in areas of greatest stress for interns, such as work-life balance, working relationships, information processing, work hours, and self-confidence.<sup>17</sup> Greater than two-thirds of respondents agreed that this program helped them with coping skills to respond to stressors in these areas, suggesting that this program can be useful to

develop skills to improve resilience, which is shown here to be strongly protective against burnout. Managing uncertainty, a key skill throughout physician practice, and developing coping skills through reflection opportunities are target areas for further evolution of the coaching program, and we believe this finding is useful in helping residency programs address well-being.

In previous evaluations of this coaching program, our ability to analyze gender, race, and ethnicity was limited. In this analysis, we learn that Black/Asian/Hispanic/other residents experience higher burnout than their non-Hispanic white peers, and male residents experience higher well-being than female residents. This warrants further exploration in future studies to discover if different coaching approaches are needed to address structural racism and gender bias within the coaching intervention or within organizations. In recent years, multiple studies have explored resident experience of bias.<sup>27–30</sup> These experiences can lead to increased stress, lower well-being, impaired coping, and increased burnout.<sup>31, 32</sup> The coaching relationship may provide an important layer of support for residents to address the role implicit bias, imposter phenomenon, and microaggressions may have in daily intern experiences.

We acknowledge several limitations in this multisite study. This study is not a randomized controlled trial (RCT). This study is also limited by a low survey response rate that varied across sites. Sites are heterogenous in their experiences and local culture, which may impact generalizability in understanding and applying these results. While coaches at all sites were trained by the same subject matter expert, the implementation and support at each site were variable. Based on this experience, we recommend institutions wishing to adopt such a program have an appointed faculty champion with administrative support and incentives to improve participation in measurements. We also do not have a means to evaluate the efficacy of the coach themselves, understanding that this is high-level communication skill that we are teaching over the course of a 3-h workshop. Future studies on the efficacy of the coach as well as the role the coach’s well-being may play in the relationship would be useful. Additionally, in this study, we address differences in outcomes based on race and ethnicity though we are limited by the low number of interns who are under-represented in medicine (URiM). A future study with greater number of URiM interns can help explore these associations and outcomes further.

A potential limitation to these findings might be the perceived impact that small differences in these scales have on physician burnout or well-being. We believe the subtle nuances within these parameters may provide insight into the multifactorial nature of physician well-being. While this cannot be answered in a prospective cohort study such as this, we anticipate that an RCT might further elucidate these subtleties and the impact of coaching. One hypothesis that could be explored in an RCT is whether no change is actually an improvement compared to someone who is not receiving

coaching. Given the many ways that internship can change an individual in this formative time, it may be that stability is in fact success, given what is known about increasing rates of burnout across the trajectory of a physician's career.<sup>3, 4, 11</sup>

The costs of this program were minimized for sustainability. Coaching materials have been published to aid residencies in onboarding a similar program.<sup>18</sup> Still, we recognize that not all programs can implement a program like this. For these programs, we believe that incorporating coaching into existing advising and mentorship programs would benefit resident well-being. The question of the difference between coaching and mentoring is often raised, and why coaching is needed if well-structured mentoring programs exist. It should be noted that this coaching program was initiated in residencies with existing mentorship programs, who were seeking additional ways to support trainee well-being. Coaching provides a non-evaluative, safe space where the personal and professional journeys of these trainees can co-exist and be explored.

Residency programs should consider formal assessments of well-being and predictive factors early in the intern year to help target support for residents who may not arrive at internship with coping, resilience, and other skills. Organizational investments in faculty education in coaching and communication skills may help residents build in time and space for goal setting and reflection. Organizations also need to measure and understand factors in their culture that may contribute to disparities in the resident experience by race and gender. Providing support, coaching, or otherwise, that targets tolerance of uncertainty, resilience, and reflection, may improve resident burnout and well-being.

---

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s11606-021-06903-5>.

**Acknowledgements:** The authors wish to thank the faculty coaches and resident participants who engaged in the program, as well as the Departments of Medicine at each institution for their support in developing and onboarding this program. The authors wish to express gratitude to the Arthur Vining Davis Foundation and the Physicians Foundation, for the financial support for our program evaluation. The authors wish to acknowledge the Institute of Coaching and Carol Kauffman, PhD, for their expertise in designing and implementing this program.

**Corresponding Author:** Kerri Palamara, MD; Center for Physician Well-Being, Department of Medicine, Massachusetts General Hospital, Fruit Street, Boston, MA, USA (e-mail: [kpalamara@partners.org](mailto:kpalamara@partners.org)).

**Funding** The Arthur Vining Davis Foundation and The Physicians Foundation each provided \$100,000 of funding to support the investigation and dissemination of this program.

**Data Availability** The datasets during and/or analyzed during the current study are available from the corresponding author on reasonable request.

**Declarations:**

**Ethical Approval:** This program and its evaluation were exempt by the Institutional Review Board with exemption #45 CFR 46.101(b) (1)

Research conducted in established or commonly accepted educational settings, involving normal educational practices.

**Conflict of Interest:** The authors declare that they do not have a conflict of interest.

## REFERENCES

1. Dyrbye LN, Shanafelt TD, Balch CM, et al. Relationship between work-home conflicts and burnout among American surgeons: A comparison by sex. *Arch Surg*. 2011;146(2):211-217. doi:<https://doi.org/10.1001/archsurg.2010.310>
2. National Academies of Sciences E and M. Taking Action Against Clinician Burnout. In: Taking Action Against Clinician Burnout: A Systems Approach to Professional Well-Being. Washington, DC: The National Academies Press; 2019. <https://doi.org/10.17226/25521>
3. West CP, Dyrbye LN, Shanafelt TD. Physician burnout: contributors, consequences and solutions. *J Intern Med*. 2018;283(6):516-529. doi:<https://doi.org/10.1111/joim.12752>
4. Panagioti M, Panagopoulou E, Bower P, et al. Controlled interventions to reduce burnout in physicians: a systematic review and meta-analysis. *JAMA Intern Med*. 2017;177(2):195-205. doi:<https://doi.org/10.1001/jamainternmed.2016.7674>
5. Dyrbye LN, Shanafelt TD, Gill PR, et al. Effect of a Professional Coaching Intervention on the Well-being and Distress of Physicians: A Pilot Randomized Clinical Trial. *JAMA Intern Med*. 2019;179(10):1406-1414. doi:<https://doi.org/10.1001/jamainternmed.2019.2425>
6. Shanafelt TD, Dyrbye LN, Sinsky C, et al. Relationship Between Clerical Burden and Characteristics of the Electronic Environment With Physician Burnout and Professional Satisfaction. *Mayo Clin Proc*. 2016;91(7):836-848. doi:<https://doi.org/10.1016/j.mayocp.2016.05.007>
7. Shanafelt TD, Gorringer G, Menaker R, et al. Impact of organizational leadership on physician burnout and satisfaction. *Mayo Clin Proc*. 2015;90(4):432-440. doi:<https://doi.org/10.1016/j.mayocp.2015.01.012>
8. Dyrbye LN, Shanafelt TD, Sinsky CA, et al. Burnout Among Health Care Professionals: A Call to Explore and Address This Underrecognized Threat to Safe, High-Quality Care. *NAM Perspect*. 2017;7(7). doi:<https://doi.org/10.31478/201707b>
9. McConnell MM, Eva KW. The role of emotion in the learning and transfer of clinical skills and knowledge. *Acad Med*. 2012;87(10):1316-1322. doi:<https://doi.org/10.1097/ACM.0b013e3182675af2>
10. Ratanawongsa N, Wright SM, Carrese JA. Well-being in residency: A time for temporary imbalance? *Med Educ*. 2007;41(3):273-280. doi:<https://doi.org/10.1111/j.1365-2929.2007.02687.x>
11. West CP, Dyrbye LN, Sinsky C, et al. Resilience and Burnout Among Physicians and the General US Working Population. *JAMA Netw Open*. 2020;3(7):e209385. doi:<https://doi.org/10.1001/jamanetworkopen.2020.9385>
12. West CP, Dyrbye LN, Erwin PJ, et al. Interventions to prevent and reduce physician burnout: a systematic review and meta-analysis. *Lancet*. 2016;388(10057):2272-2281. doi:[https://doi.org/10.1016/S0140-6736\(16\)31279-X](https://doi.org/10.1016/S0140-6736(16)31279-X)
13. Grover S, Furnham A. Coaching as a developmental intervention in organisations: A systematic review of its effectiveness and the mechanisms underlying it. *PLoS One*. 2016;11(7). doi:<https://doi.org/10.1371/journal.pone.0159137>
14. Gazelle G, Liebschutz JM, Riess H. Physician Burnout: Coaching a Way Out. *J Gen Intern Med*. 2015;30(4):508-513. doi:<https://doi.org/10.1007/s11606-014-3144-y>
15. Whitmore SJ. Coaching for Performance: The Principles and Practice of Coaching and Leadership. Fifth. (Whitmore J, ed.). London, UK: John Murray Press; 2020.
16. McGonagle AK, Schwab L, Yahanda N, et al. Coaching for Primary Care Physician Well-Being: A Randomized Trial and Follow-Up Analysis. *J Occup Health Psychol*. 2020. doi:<https://doi.org/10.1037/ocp0000180>
17. Palamara K, Kauffman C, Chang Y, et al. Professional development coaching for residents: Results of a 3-year positive psychology coaching intervention. *Journal of General Internal Medicine*. 2018; 33(11):1842-4.
18. Palamara K, Kauffman C, Stone VE, et al. Promoting Success: A Professional Development Coaching Program for Interns in Medicine. *J*



- Grad Med Educ. 2015;7(4):630-637. doi:<https://doi.org/10.4300/JGME-D-14-00791.1>
19. **Maslach C, Jackson SE, Leiter MP.** The Maslach Burnout Inventory Manual. 3rd ed. Palo Alto, CA: Consulting Psychologists Press; 1996. doi:<https://doi.org/10.1002/job.4030020205>
  20. **Larsen D, Chu JT, Yu L, Chang Y, Donelan K, Palamara K.** Correlating burnout and well-being in a multisite study of internal medicine residents and faculty. *J Gen Intern Med.* 2021. doi: <https://doi.org/10.1007/s11606-021-06653-4>.
  21. **Butler J, & Kern ML** (2016). The PERMA-Profil: A brief multidimensional measure of flourishing. *Int J Wellbeing.* 6(3), 1-48. doi:<https://doi.org/10.5502/ijw.v6i3.526>
  22. **Rafferty JP, Lemkau JP, Purdy RR,** et al. Validity of the Maslach burnout inventory for family practice physicians. *J Clin Psychol.* 1986;42(3):488-492. doi:[https://doi.org/10.1002/1097-4679\(198605\)42:3<488::AID-JCLP2270420315>3.0.CO;2-S](https://doi.org/10.1002/1097-4679(198605)42:3<488::AID-JCLP2270420315>3.0.CO;2-S)
  23. **Carleton RN, Norton MAPJ, Asmundson GJG.** Fearing the unknown: A short version of the Intolerance of Uncertainty Scale. *J Anxiety Disord.* 2007;21(1):105-117. doi:<https://doi.org/10.1016/j.janxdis.2006.03.014>
  24. **Bartone PT.** Test-retest reliability of the dispositional resilience scale-15, a brief hardiness scale. *Psychol Rep.* 2007;101(3 1):943-944. doi:<https://doi.org/10.2466/PRO.101.3.943-944>
  25. **McCullough ME, Emmons RA, Tsang J** (2002). The grateful disposition: A conceptual and empirical topography. *Journal of Personality and Social Psychology.* 82, 112-127.
  26. **Carver CS** (2013) Measure of Current Status (MOCS). Measurement Instrument Database for the Social Science. Retrieved from [www.midss.ie](http://www.midss.ie) Accessed 28 Feb 2015.
  27. **Garcia LC, Shanafelt TD, West CP,** et al. Burnout, Depression, Career Satisfaction, and Work-Life Integration by Physician Race/Ethnicity. *JAMA Netw Open.* 2020;3(8):e2012762. doi:<https://doi.org/10.1001/jamanetworkopen.2020.12762>
  28. **Hu YY, Ellis RJ, Hewitt DB,** et al. Discrimination, abuse, harassment, and burnout in surgical residency training. *N Engl J Med.* 2019;381(18):1741-1752. doi:<https://doi.org/10.1056/NEJMsa1903759>
  29. **Liebschutz JM, Darko GO, Finley EP,** et al. In the minority: Black physicians in residency and their experiences. *J Natl Med Assoc.* 2006;98(9):1441-1448.
  30. **Fnais N, Soobiah C, Chen MH,** et al. Harassment and discrimination in medical training: A systematic review and meta-analysis. *Acad Med.* 2014;89(5):817-827. doi:<https://doi.org/10.1097/ACM.0000000000000200>
  31. **Sue DW, Alsaidi S, Awad MN, Glaeser E, Calle CZ, Mendez N.** Disarming racial microaggressions: Microintervention strategies for targets, White allies, and bystanders. *Am Psychol.* 2019;74(1):128-142. doi: <https://doi.org/10.1037/amp0000296>.
  32. **Zurbrugg L, Miner KN.** Gender, Sexual Orientation, and Workplace Incivility: Who Is Most Targeted and Who Is Most Harmed? *Front Psychol.* 2016;7:565. doi: <https://doi.org/10.3389/fpsyg.2016.00565>.

**Publisher's Note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.