BMC Health Services Research

SYSTEMATIC REVIEW

Open Access

Effectiveness of wellness program interventions to improve physician wellness: a systematic review

Nikola Kolobaric[1,6], Benjamin Milone[1,3], Maria Salman[1,6], Amos Buh[1], Nandini Biyani[1], Gamal Wafy[1], Stephen G Fung[2], Mekaylah Scott[1,4], Rohan Kiska[1,5], Rachel Kang[1,4], Kathryn Lee[1,3], Sarah Syed[1,6], Maya Gibb[1], Shan Dhaliwal[1,6], Daniel Myran[1,2,6], Courtney Maskerine[2,6], Pierre Antoine Brown[1,6], Ayub Akbari[1,6], Gregory L. Hundemer[1,6], Peter Tanuseputro[7], Deena Fremont[1], Edward Spilg[1,2,6†] and Manish M. Sood[1,6*†]

Abstract

Background Physician wellness programs are being implemented to offset rises in physician burnout. Insight into the effectiveness of these programs and to whom they are being offered, remains unclear.

Objectives To identify and characterize wellness program interventions to improve physician wellness.

Methods A PRISMA-P 2020-compliant systematic review as conducted, searching PubMed, Scopus, and Medline from May 2006 to July 2024. Search terms included Medical Subject Headings terms and keywords related to physicians and wellness program interventions. Peer reviewed published studies that qualitatively and/or quantitively measured outcomes of wellness interventions for practicing physicians were included.

Results Thirty-six studies involving 6,708 total participants were included. Interventions were heterogenous and included group therapy, stress reduction strategies, time off/workload reductions, education, and peer support. The efficacy of interventions varied, with sixteen studies (44.4%) demonstrating some measurable degree of effectiveness, with statistically significant changes ($p < 0.05$) post-intervention. Few studies reported improvements by physician sex, age groups, or comparisons across specialities.

Conclusion Studies examining physician wellness program interventions are highly heterogenous in terms of intervention, study design and methods of outcome assessment, limiting definitive conclusions about their general effectiveness.

Trial registration The review protocol has been registered on Open Science Framework (https://doi.org/10.17605/OSF.IO/8SDM9).

Keywords Physician health, Physician wellness, Interventions

†Edward Spilg and Manish M. Sood have contributed equally as co-senior authors.

*Correspondence:
Manish M. Sood
msood@toh.ca; Msood@toh.on.ca
[1]Ottawa Hospital Research Institute, The Ottawa Hospital, Civic Campus, 2-014 Administrative Services Building, 1053 Carling Avenue, Box 693, Ottawa, ON K1Y 4E9, Canada
[2]Bruyère Research Institute, Ottawa, ON, Canada
[3]Faculty of Science, McMaster University, Hamilton, ON, Canada
[4]Faculty of Health Sciences, University of Ottawa, Ottawa, ON, Canada
[5]Faculty of Science, Carleton University, Ottawa, ON, Canada
[6]Faculty of Medicine, University of Ottawa, Ottawa, ON, Canada
[7]Deparment of Family Medicine and Primary Care, Li Ka Shing Faculty of Medicine, University of Hong Kong, Hong Kong, China

Introduction

It is well established that in recent years, there has been a noticeable decline in physician wellness across all phases of education and training [1–4]. This decline is evident through measurable factors such as high rates of emotional exhaustion, mental health concerns, and burnout. Moreover, the prevalence of burnout appears to be on the rise, with longitudinal studies showing an upward trajectory [5–7]. Various factors appear to be contributing to this decline in physician wellness, including demanding hours, administrative burdens, and difficulties with work-life balance [1, 3, 8–12]. Additionally, systemic issues within healthcare systems, such as limited resources, inadequate support structures, and a culture of perfectionism create additional stressors for physicians [3, 6, 9, 12]. The identification and characterization of such factors could aid in informing the development of support initiatives for the growing number of physicians who are challenged by fatigue and burnout [13].

To address issues related to physician burnout, multiple health facilities and physician organizations have allocated considerable resources to developing and implementing physician wellness programs. These programs often address one key contributor of burnout to a more significant degree or implement a wider scope of service that intends to generally improve the overall health and wellness of physicians [3, 14]. Wellness program interventions may range from individual-focused strategies, such as mindfulness training, stress management workshops, and resilience-building programs, to organizational-level initiatives, such as redesigning work processes, improving work environments, and fostering a culture of support and well-being [3, 15–18].

Promoting physician wellness benefits both their health and the quality of patient care, contributing to a sustainable healthcare workforce [3, 18–23]. However, the current literature regarding physician wellness lacks clear and current characterization of which types of interventions are consistently effective. Assessing the efficacy of interventions aimed at improving physician wellness is crucial in addressing the present decline in physician wellness [13, 24, 25]. As a result, this systematic review aims to synthesize contemporary literature on physician wellness program interventions from May 2006 to July 2024. By systematically analyzing studies conducted over this timeframe, we seek to provide insights into the effectiveness of various interventions aimed at improving physician well-being. Through rigorous methods of study selection, data extraction, and synthesis, we aim to identify trends, patterns, and gaps in the literature, ultimately informing strategies for promoting physician wellness and mitigating the negative impacts of burnout and stress.

Methods

Study design

This systematic review assessed studies that investigated the efficacy of implemented wellness program interventions for physicians. The review protocol has been registered on Open Science Framework (https://doi.org/10.17605/OSF.IO/8SDM9). We incorporated an a-priori protocol (OSF protocol) developed following Preferred Reporting Items for Systematic Review and Meta-Analysis Protocols (PRISMA-P) 2020 guidelines (Supplemental Material).

Population of interest

This review included studies concerning physicians, which includes resident/intern, fellowship, and fully trained staff physicians.

Intervention

This review included studies implementing physician wellness programs.

Comparison

The efficacy of the program interventions was evaluated by comparison of the population's wellness before and after their implementation, including qualitative feedback from the study participants regarding their perceived efficacy.

Outcomes

This review considered studies which assessed one or more components of the wellness among the study participants. Some common outcomes examples include burnout, depression, perceived stress, resiliency, job satisfaction, and perceived happiness. Study outcomes concerning feedback regarding the perceived efficacy of the interventions reported by participants were also included.

Types of studies included

This review included cross-sectional studies that were either before-after studies or qualitative studies, as well as randomized and non-randomized control trials. Peer-reviewed published studies that reported on (i) efficacy of wellness program interventions, (ii) in medical residents, fellows and/or staff physicians and (iii) published in English were included.

Search strategy

The search phase of the review was conducted using the databases PubMed, Scopus and Medline, with keyword terms being "Physician/Residents", "Wellness Programs/Benefits", "Patient contact healthcare workers", "Physician Wellness", "Physician Health", "Health Initiatives", "Healthy Lifestyle" and "Work-Life Balance", while MeSH

terms used were "Workplace Health Issues", "Mental Health", "Malpractice Stress", and "Depression". Scopus and Medline search results were obtained July 4 2024, and PubMed results were obtained on July 9 2024.

This systematic search of PubMed, Medline, and Scopus was conducted from May 2006 to July 2024, using a combination of MeSH terms and keywords related to physician wellness, burnout, and intervention outcomes. The search was limited to peer-reviewed studies published in English. The full search strategy is presented in the Supplementary Material.

Study screening and selection process

All articles collected from the initial search were organized in Covidence for screening. Screening was conducted using multiple screeners with conflicts being resolved using a third opinion. The first stage of screening focused on title and abstract, before full-text screening in the second stage. Studies that did not report any outcomes, included health care professionals other than physicians, did not implement an actual intervention in a study population or did not provide details on the intervention implemented were excluded.

Assessment of methodological quality

A risk of bias assessment was conducted for each study type by considering any missing data, loss to follow-up and validity of measurements. This was completed by two independent reviewers, with discrepancies being resolved by a third reviewer. Due to the broad inclusion criteria, we anticipated a variety of study designs. Therefore, we used a variety of quality assessment tools provided by the National Heart, Lung, and Blood Institute (NHLBI) and JBI's critical appraisal tools. From the NHLBI tools we used the Quality Assessment of Controlled Intervention Studies, as well as the Quality Assessment of Before-After Studies with no Control Group where applicable [26]. For qualitative research, we used the JBI Quality Assessment of Qualitative Research [27].

Data extraction

Data extraction was completed by two independent reviewers on the following information: study details (title, last name of first author, year of publication, journal, country), study characteristics (design, participant type, missing data) study population data and intervention information (main analysis method, sample size, mean age at baseline in years, demographics reported, intervention description, area of focus of wellness program) intervention group statistics (sample size, mean age in years, pre-intervention score, post-intervention score, within group difference), control group statistics (description, sample size, mean age in years, pre-intervention score, post-intervention score, within group difference) and statistical results (method of outcome measurement, between-group difference, relevant findings). Discrepancies were resolved by a third reviewer.

Data synthesis

The diversity of interventions and assessment methods prevented data synthesis through meta-analysis or meta-regression. However, effect size estimates were calculated for wellness metrics that reported standard deviations, using Hedge's g due to correction of bias for small sample sizes in the studies evaluated. For studies reporting a single total value for wellness, effect size was estimated for solely that metric. For positive wellness metrics – those with an intended increase in value over time – effect sizes with a Hedge's g value greater than 0.8 denoted a large effect, those between 0.8 and 0.5 had a medium effect and anything less than 0.5 had a small effect. For negative wellness metrics – those with an intended decrease in value over time – effect sizes with a Hedge's g value less than – 0.8 denoted a large effect, those between – 0.8 and – 0.5 had a medium effect and anything greater than – 0.5 had a small effect. The full set of effect size estimates can be found in ETable 5 and ETable 6 in the Supplement.

Confidence in cumulative evidence

The evidence in this review has been evaluated using the Grading of Recommendations, Assessment, Development and Evaluation (GRADE) approach [28]. For quantitative studies, GRADE criteria were applied to assess risk of bias, inconsistency, indirectness, imprecision, and publication bias. Randomized controlled trials (RCTs) began as high-quality evidence and observational studies as low-quality, though this was subject to upgrading or downgrading based on the study characteristics. For qualitative studies, GRADE was adapted to evaluate methodological rigor, coherence, data adequacy and relevance to the research question [29].

Ethics and dissemination

Due to the lack of primary data being collected, no ethical clearance was required for the conduction of this review. The study has adhered to the procedures outlined by these methods. The results of this study will be disseminated through presentations at conferences to inform further research on physician wellness interventions.

Results

Study characteristics

Our literature search yielded 1191 results, after which 80 references were removed due to being duplicate studies (EFigure 1). During the title and abstract screening stage, studies were predominantly excluded for factors such as a lack of intervention implementation or evaluation. Overall, 36 studies were determined to be adequate for

inclusion in this systematic review. The characteristics of each study are summarized in Table 1.

Study quality appraisal

A risk assessment adapted from the NHLBI critical appraisal tools determined that the majority (85.7%) [31–36] of RCTs were identified as fair quality and only one (14.3%) [30] was of poor quality. For before-after studies with no randomization or control, 60% [44, 50, 52, 55–57, 60, 63, 65] were identified as fair quality, while the remaining 40% [37, 39, 45, 46, 59, 62] were deemed poor quality. Finally, using the JBI critical appraisal tool, 21.4% [43, 54, 61] of the included qualitative studies were deemed to be of fair quality, while the remaining 78.6% [38, 40–42, 47–49, 51, 53, 58, 64] were deemed poor quality. Most of the quality rating penalties were related to low response rates or small, unrepresentative sample sizes (Supplementary Tables 1–3). Quality and risk of bias assessment for the included studies determined the relevance and quality of the methods utilized were satisfactory for the purposes of this review.

Study interventions

Each wellness program intervention was unique in its implementation (Tables 2 and 3). Sixteen (44.4%) [32, 34, 36–38, 42, 48, 49, 51–53, 58–61, 64] of the wellness programs were administered one-on-one, including included interventions such as peer mentoring, wellness resources and individual counselling services. Seventeen (47.2%) [30, 31, 33, 35, 39–41, 43, 44, 47, 50, 55–57, 62,

Table 1 General characteristics of the included studies assessing interventions for physician wellness

Author, Year	Location	Distribution	Type of Intervention	Study Design
Kavanaugh et al., [30] 2022	USA	Single centre	Physical Activity	RCT
Huang et al., [31] 2020	China	Single centre	Social Connection	RCT
Taylor et al., [32] 2020	Australia	Single centre	Physical Activity	RCT
Loewenthal et al., [33] 2021	USA	Multi centre (3)	Physical Activity	RCT
Sood et al., [34] 2011	USA	Single centre	Educational	RCT
Spilg et al., [35] 2022	Canada	Single centre	Educational	RCT
Saadat et al., [36] 2012	USA	Single centre	Educational	RCT
Williamson et al., [37] 2021	USA	Multi centre (10)	Educational	NRCT
Jacob & Lambert [38], 2021	USA	Single centre	Educational	CSS
Garcia et al., [39] 2021	USA	Single centre	Administrative	CSS
Jung et al., [40] 2021	USA	Single centre	Educational	CSS
Mohamed et al., [41] 2022	USA	Single centre	Educational	CSS
Dabrow et al., [42] 2006	USA	Single centre	Social Connection	CSS
Hategan & Riddell [43], 2020	Canada	Single centre	Multifactorial	CSS
Babbar et al., [44] 2019	USA	Single centre	Physical Activity	CSS
Petrie et al., [45] 2022	Australia	Multi centre (2)	Multifactorial	CSS
Wothe et al., [46] 2022	USA	Single centre	Multifactorial	CSS
Baumann et al., [47] 2020	USA	Single centre	Educational	CSS
Ward et al., [48] 2022	Australia	Multi centre (4)	Social Connection	CSS
Keyser et al., [49] 2021	USA	Multi centre (8)	Social Connection	CSS
Fabreau et al., [50] 2013	Canada	Multi centre (2)	Multifactorial	CSS
Awadallah et al., [51] 2023	USA	Single centre	Social Connection	CSS
Ricker et al., [52] 2021	USA	Single centre	Educational	CSS
Fournier & Tourian [53], 2020	Canada	Single centre	Social Connection	CSS
Ey et al., [54] 2016	USA	Single centre	Multifactorial	CSS
Chang et al., [55] 2023	USA	Single centre	Social Connection	CSS
Letica-Kriegel et al., [56] 2023	USA	Single centre	Social Connection	CSS
Runyan et al., [57] 2016	USA	Multi centre (3)	Educational	CSS
Johnson et al., [58] 2022	USA	Single centre	Physical Activity	CSS
Cawyer et al., [59] 2022	USA	Single centre	Educational	CSS
Shahid et al., [60] 2018	USA	Single centre	Educational	CSS
Aggarwal et al., [61] 2017	USA	Single centre	Educational	CSS
Bisgaard et al., [62] 2021	USA	Single centre	Administrative	CSS
Hart et al., [63] 2019	USA	Single centre	Social Connection	CSS
Bui et al., [64] 2020	USA	Multi centre	Multifactorial	CSS
Mari et al., [65] 2019	USA	Single centre	Multifactorial	CSS

RCT Randomized Controlled Trial, *NRCT* Non-Randomized Controlled Trial, *CSS* Cross-Sectional Study

Table 2 Intervention design, method of measurement and outcomes for randomized controlled trials assessing interventions for physician wellness

Author, Year	Sample (*n*)	Mean Response Rate (%)	Description of Intervention	Use of Validated Tool	Outcome Assessment	Intervention Group Score(s) Pre-test	Intervention Group Score(s) During or Post-test	Control Group Score(s) Pre-test	Control Group Score(s) During or Post-test	Within Group *p*-value	Between Group *p*-value
Kavanaugh et al., [30] 2022	56	60.7	3 h walk in nature with various activities. Reflect and interact with surroundings. Participants in the control group were placed on a waitlist for the intervention.	yes	OBI & Mini-Z Questionnaire	OBI (mean) 39 Mini-Z (mean) 35	OBI (mean) 38 Mini-Z (mean) 34	OBI (mean) 42 Mini-Z (mean) 36	OBI (mean) 41 Mini-Z (mean) 37	OBI Intervention: 0.30 Control: 0.44 Mini-Z Intervention: 0.43 Control: 0.27	OBI Pre-test: 0.07 Post-test: 0.17 Mini-Z Pre-test: 0.21 Post-test: 0.03
Huang et al., [31] 2020	36	100.0	Balint groups; discussions with patients about perceptions of interactions with patients. Participants in the control group were placed on a waitlist for future Balint sessions.	yes	Burnout measured by MBI and Satisfaction measured by MSQ	MBI (mean ± SD) EE: 16.89 ± 4.825 DP: 7.00 ± 3.361 PA: 28.50 ± 7.139 MSQ (mean ± SD) JS: 71.56 ± 5.913	MBI (mean ± SD) EE: 15.00 ± 5.1111 DP: 6.17 ± 3.130 PA: 30.33 ± 6.843 MSQ (mean ± SD) JS: 72.17 ± 5.586 Change (mean ± SD) EE: −1.89 ± 6.057 DP: −0.83 ± 4.866 PA: 1.83 ± 10.870 JS: 0.61 ± 9.185	MBI (mean ± SD) EE: 15.83 ± 5.533 DP: 6.89 ± 2.847 PA: 28.06 ± 7.174 MSQ (mean ± SD) JS: 70.72 ± 6.524	MBI (mean ± SD) EE: 22.17 ± 9.482 DP: 9.72 ± 3.801 PA: 26.22 ± 6.839 MSQ (mean ± SD) JS: 70.11 ± 6.668 Change (mean ± SD) EE: 6.33 ± 11.621 DP: 2.83 ± 4.148 PA: −1.83 ± 10.853 JS: −0.61 ± 10.738	MBI Intervention EE: 0.203 DP: 0.477 PA: 0.484 Control EE: 0.034 DP: 0.010 PA: 0.483 MSQ Intervention JS: 0.781 Control JS: 0.812	MBI EE: 0.013 DP: 0.020 PA: 0.318 MSQ JS: 0.716
Taylor et al., [32] 2020	21	85.7	Personalized, trauma-informed yoga sessions with workshops and eHealth program. Compared to generic group fitness program as control.	yes	MBI, ProQOL, PTSD-C, LEC, DSM-5, SIAS, MAAS & SDS	MBI (mean ± SD) EE: 27 ± 11 DP: 14 ± 6 PA: 33 ± 6 ProQOL (mean ± SD) CE: 34 ± 3 BO: 27 ± 4 STS: 24 ± 6 (mean ± SD) PTSD-C: 10 SIAS: 9 ± 3 MAAS: 3 ± 1 SDS: 4	MBI (mean ± SD) EE: 25 ± 10 DP: 10 ± 6 PA: 36 ± 6 ProQOL (mean ± SD) CE: 36 ± 4 BO: 25 ± 5 STS: 24 ± 7 (mean ± SD) PTSD-C: 10 SIAS: 11 ± 1 MAAS: 4 ± 1 SDS:2	MBI (mean ± SD) EE: 21 ± 9 DP: 10 ± 7 PA: 32 ± 8 ProQOL (mean ± SD) CE: 36 ± 3 BO: 27 ± 4 STS: 22 ± 6 (mean ± SD) PTSD-C: 8 SIAS: 11 ± 6 MAAS: 4 ± 1 SDS: 2	MBI (mean ± SD) EE: 21 ± 12 DP: 11 ± 7 PA: 36 ± 7 ProQOL (mean ± SD) CE: 38 ± 4 BO: 25 ± 6 STS: 22 ± 9 (mean ± SD) PTSD-C: 8 SIAS: 10 ± 4 MAAS: 3 ± 1 SDS: 0	MBI Intervention EE: 0.42 DP: 0.15 PA: <0.001 Control EE: 0.95 DP: 0.25 PA: 0.004 ProQOL Intervention CE: 0.003 BO: <0.001 STS: 0.62 Control CE: 0.09 BO: 0.003 STS: 0.71 Intervention PTSD-C: <0.001 SIAS: 0.92 MAAS: 0.29 SDS: 0.09 Control PTSD-C: <0.001 SIAS: 0.37 MAAS: 0.37 SDS: 0.10	MBI EE: 0.49 DP: 0.05 PA: 0.52 ProQOL CE: 0.22 BO: 0.47 STS: 0.89 PTSD-C: <0.001 SIAS: 0.47 MAAS: 0.13 SDS: 0.85

Table 2 (continued)

Author, Year	Sample (*n*)	Mean Response Rate (%)	Description of Intervention	Use of Validated Tool	Outcome Assessment	Intervention Group Score(s) Pre-test	Intervention Group Score(s) During or Post-test	Control Group Score(s) Pre-test	Control Group Score(s) During or Post-test	Within Group *p*-value	Between Group *p*-value
Loewenthal et al., [33] 2021	56	75.9	Resilience, integration, self-awareness, engagement (RISE) yoga intervention, by Kripalu Center for Yoga and Health. 1 h weekly for 6 weeks. The control group comprised participants who were put on a waitlist for RISE and given an overview of RISE methods at the end of the study period.	yes	MBI, FFMQ, RS-14, PSS, PFI, PROMIS (D/A & Sleep) & RWBI	NR	Changes (mean) 0–6 weeks Resilience: 2.66 Mindfulness: 6.15 Stress: −2.68 Anxiety: NR Depression: NR Sleep disturbance: −2.64 Exhaustion: NR Cynicism: NR Total burnout (MBI): −3.08 Fulfillment: NR Work exhaustion: −4.66	NR	Changes (mean) 0–6 weeks NR 0–2 months Anxiety: 1.36 Sleep Disturbance: −2.04 All others NR	Intervention 0–6 weeks Resilience: 0.013 Mindfulness: 0.0001 Stress: 0.013 Anxiety: >0.08 Depression: >0.08 Sleep disturbance: 0.014 Exhaustion: >0.08 Cynicism: >0.08 Total burnout (MBI): 0.021 Fulfillment: >0.08 Work exhaustion: 0.001	0–6 weeks Mindfulness: 0.004 0–2 months Work exhaustion: 0.052 All others > 0.08
							Interpersonal disengagement:−2.01 Total burnout (PFI): NR Well-being: −2.560–2 months Resilience: 1.86 Mindfulness: 3.03 Stress: −3.24 Anxiety: NR Depression: NR Sleep disturbance:−2.04 Exhaustion: NR Cynicism: NR Total burnout (MBI): 0.005 Fulfillment: NR Work exhaustion −3.78			Interpersonal disengagement: 0.05 Total burnout (PFI): >0.08 Well-being: 0.0170–2 months Resilience: 0.075 Mindfulness: 0.006 Stress: 0.003 Anxiety: >0.08 Depression: >0.08 Sleep disturbance: 0.053 Exhaustion: >0.08 Cynicism: >0.08 Total burnout (MBI): 0.005 Fulfillment: >0.08 Work exhaustion: 0.001	

Table 2 (continued)

Author, Year	Sample (*n*)	Mean Response Rate (%)	Description of Intervention	Use of Validated Tool	Outcome Assessment	Intervention Group Score(s) Pre-test	Intervention Group Score(s) During or Post-test	Control Group Score(s) Pre-test	Control Group Score(s) During or Post-test	Within Group *p*-value	Between Group *p*-value
							Interpersonal disengagement:−2.07 Total burnout (PFI): NR Well-being: −2.38				Interpersonal disengagement: 0.049 Total burnout (PFI): >0.08 Well-being: 0.026 Control 0–6-weeks All >0.08 0–2 months Anxiety: 0.036 Sleep Disturbance: 0.031 All others > 0.08
Sood et al., [34] 2011	40	80.0	One 90 min SMART session, adapted from AIT by Mayo Clinic. The control subjects were placed on a waitlist for the SMART intervention.	yes	CDRS, PSS, SAS & LASA	(mean ± SD) CDRS: 69.6 ± 13.3 PSS: 28.2 ± 5.9 SAS: 55.2 ± 13.6 LASA QOL: 7.6 ± 1.2 Fatigue: 4.5 ± 2.6	(mean ± SD) CDRS: 79.4 ± 11.3 PSS: 22.8 ± 5.5 SAS: 43.4 ± 14.1 LASA QOL: 8.0 ± 1.3 Fatigue: 4.3 ± 2.5	(mean ± SD) CDRS: 68.0 ± 11.2 PSS: 26.2 ± 6.9 SAS: 50.5 ± 23.0 LASA QOL: 7.8 ± 1.1 Fatigue: 4.9 ± 2.7	(mean ± SD) CDRS: 67.2 ± 11.6 PSS: 28.3 ± 6.3 SAS: 53.4 ± 23.1 LASA QOL: 7.2 ± 1.2 Fatigue: 5.4 ± 2.4	Intervention CDRS: <0.05 PSS: <0.05 SAS: <0.05 LASA QOL: >0.05 Fatigue: >0.05 Control CDRS: >0.05 PSS: >0.05 SAS: >0.05 LASA QOL: >0.05 Fatigue: >0.05	CDRS: 0.003 PSS: 0.010 SAS: 0.001 LASA QOL: 0.029 Fatigue: 0.462
Spilg et al., [35] 2022	40	89.2	2-hour SMART workshop followed by 6-month optional online program. Participants in the control group received no SMART training.	yes	CDRS, SHS, PSS & GAD-7	(mean ± SD) CDRS: 69.35 ± 7.71 SHS: 5.24 ± 0.93 PSS: 17.75 ± 5.48 GAD-7 (Median, IQR): 5.00 (4.00, 6.00)	3 month (mean ± SD) CDRS: 73.21 ± 9.05 SHS: 5.34 ± 1.04 PSS: 15.21 ± 6.61 GAD-7 (Median, IQR): 6.00 (2.40, 7.45) 6 month (mean ± SD) CDRS: 73.69 ± 10.90 SHS: 5.25 ± 1.22 PSS: 14.88 ± 8.16 GAD-7 (Median, IQR): 5.48 (2.00, 9.53)	(mean ± SD) CDRS: 68.42 ± 11.82 SHS: 5.21 ± 0.98 PSS: 15.74 ± 4.02 GAD-7 (Median, IQR): 4.00 (2.00, 6.00)	3 month (mean ± SD) CDRS: 66.59 ± 10.10 SHS: 5.13 ± 0.92 PSS: 16.82 ± 5.39 GAD-7 (Median, IQR): 5.00 (3.00, 8.75) 6 month (mean ± SD) CDRS: 66.88 ± 10.01 SHS: 5.09 ± 1.01 PSS: 17.75 ± 7.81 GAD-7 (Median, IQR): 6.32 (2.21, 10.72)	NR	Scores (3 & 6 month) CDRS: >0.05 SHS: >0.05 PSS: >0.05 GAD-7: >0.05 LSMDs 3 month CDRS: 0.059 SHS: 0.508 PSS: 0.214 GAD-7 (log-scale): 0.167 6 month CDRS: 0.090 SHS: 0.316 PSS: 0.111 GAD-7 (log-scale): 0.068

Table 2 (continued)

Author, Year	Sample (*n*)	Mean Response Rate (%)	Description of Intervention	Use of Validated Tool	Outcome Assessment	Intervention Group Score(s) Pre-test	Intervention Group Score(s) During or Post-test	Control Group Score(s) Pre-test	Control Group Score(s) During or Post-test	Within Group *p*-value	Between Group *p*-value
Saadat et al., [36] 2012	60	96.7	Based on Pearlin and Schooler's hierarchy of coping mechanisms. 1 year program to identify stressful situations, modify appraisal processes and eliminate sources of stress. 2 control groups; NTC-RT participants were relieved of clinical duties for the same time as the intervention group during the study period, while the NTC-RD group continued to go about their routine duties.	yes	RQS, CSI, HMPSS, SSTAI, CES-DS. CHI-PS & NSDA	(mean ± SD) RQS Job: 36.1 ± 7.6 Partner: 22.9 ± 6.5 Parent: 17.0 ± 3.8 CSI Problem solving: 26.3 ± 4.7 Seek support: 24.6 ± 5.6 Avoidance: 18.5 ± 4.3 HMPSS Work: 25.1 ± 6.5 Home: 49.1 ± 10.3	(mean ± SD) RQS Job: 36.1 ± 7.4 Partner: 21.7 ± 4.6 Parent: 18.3 ± 3.1 CSI Problem solving: 27.7 ± 4.1 Seek support: 25.6 ± 4.6 Avoidance: 16.5 ± 3.5 HMPSS Work: 27.3 ± 6.9 Home: 52.1 ± 9.8	NTC-RT (mean ± SD) RQS Job: 35.4 ± 7.0 Partner: 23.0 ± 3.8 Parent: 16.7 ± 4.4 CSI Problem solving: 26.1 ± 5.2 Seek support: 25.4 ± 5.1 Avoidance: 19.4 ± 3.9	NTC-RT (mean ± SD) RQS Job: 36.6 ± 8.7 Partner: 23.6 ± 4.9 Parent: 20.1 ± 5.8 CSI Problem solving: 27.8 ± 4.4 Seek support: 24.5 ± 4.9 Avoidance: 19.3 ± 4.1	NR	NTC-RD vs. WIG RQS Job: 0.33 Partner: 0.09 Parent: 0.03 CSI Problem solving: 0.03 Seek support: 0.41 Avoidance: 0.07
						SSTAI Anxiety: 44.4 ± 8.8 CES-DS Depression: 20.4 ± 9.2 CHI-PS Somatic symptoms: 56.3 ± 11.7 NSDA Alcohol: 13.9 ± 16.5	SSTAI Anxiety: 38.4 ± 8.6 CES-DS Depression: 19.8 ± 7.3 CHI-PS Somatic symptoms: 55.1 ± 14.9 NSDA Alcohol: 10.4 ± 13.9	HMPSS Work: 26.4 ± 6.7 Home: 50.1 ± 8.9 SSTAI Anxiety: 48.6 ± 10.1 CES-DS Depression: 20.5 ± 7.4 CHI-PS Somatic symptoms: 54.3 ± 11.3	HMPSS Work: 25.5 ± 7.2 Home: 49.5 ± 10.6 SSTAI Anxiety: 46.7 ± 12.1 CES-DS Depression: 22.3 ± 6.9 CHI-PS Somatic symptoms: 53.4 ± 10.4 NSDA Alcohol: 13.6 ± 17.6		HMPSS Work: 0.02 Home: 0.42 SSTAI Anxiety: 0.02 CES-DS Depression: 0.09 CHI-PS Somatic symptoms: 0.09 NSDA Alcohol: 0.46
								NSDA Alcohol: 11.4 ± 12.6 NTC-RD (mean ± SD) RQS Job: 36.5 ± 8.3 Partner: 21.7 ± 5.1 Parent: 17.6 ± 5.6 CSI Problem solving: 28.7 ± 3.4 Seek support: 26.8 ± 5.0 Avoidance: 18.8 ± 4.7	NTC-RD (mean ± SD) RQS Job: 38.7 ± 10.0 Partner: 24.1 ± 10.2 Parent: 26.7 ± 11.2 CSI Problem solving: 27.1 ± 4.7 Seek support: 25.4 ± 6.1 Avoidance: 19.5 ± 5.8 HMPSS Work: 26.7 ± 7.5 Home: 53.0 ± 10.2		NTC-RT vs. WIGRQS Job: 0.61 Partner: 0.26 Parent: 0.35 CSI Problem solving: 0.81 Seek support: 0.26 Avoidance: 0.14 HMPSS Work: 0.02 Home: 0.23

Table 2 (continued)

Author, Year	Sample (*n*)	Mean Response Rate (%)	Description of Intervention	Use of Validated Tool	Outcome Assessment	Intervention Group Score(s) Pre-test	Intervention Group Score(s) During or Post-test	Control Group Score(s) Pre-test	Control Group Score(s) During or Post-test	Within Group *p*-value	Between Group *p*-value
								HMPSS Work: 27.6 ± 6.9 Home: 51.5 ± 9.7 SSTAI Anxiety: 44.1 ± 6.9 CES-DS Depression: 21.1 ± 7.6 CHI-PS Somatic symptoms: 56.0 ± 12.0 NSDA Alcohol: 11.3 ± 16.3	SSTAI Anxiety: 45.6 ± 11.6 CES-DS Depression: 23.9 ± 7.6 CHI-PS Somatic symptoms: 63.2 ± 23.2 NSDA Alcohol: 9.3 ± 11.1	SSTAI Anxiety: 0.20 CES-DS Depression: 0.44 CHI-PS Somatic symptoms: 0.75 NSDA Alcohol: 0.06 NTC-RD vs. NTC-RT RQS Job: 0.62 Partner: 0.45 Parent: 0.14 CSI Problem solving: 0.02 Seek support: 0.79 Avoidance: 0.68 HMPSS Work: 0.96 Home: 0.72 SSTAI Anxiety: 0.23 CES-DS Depression: 0.39 CHI-PS Somatic symptoms: 0.14 NSDA Alcohol: 0.23

OBI Oldenburg Burnout Inventory, *MBI* Maslach Burnout Inventory, *MSQ* Minnesota Satisfaction Questionnaire, *EE* Emotional Exhaustion, *DP* Depersonalization, *PA* Personal Achievement, *JS* Job Satisfaction, *ProQOL* Professional Quality of Life, *PTSD-C* Post-Traumatic Stress Disorder Checklist, *LEC* Life Events Checklist, *DSM-5* Diagnostic and Statistical Manual of Mental Disorders 5 (extended criterion A), *SIAS* Suicidal Ideation Attributes Scale, *MAAS* Mindful Attention Awareness Scale, *SDS* Shutdown Dissociation Scale, *CE* Compassion Satisfaction, *BO* Burnout, *STS* Secondary Traumatic Stress, *FFMQ* Five Facet Mindfulness Questionnaire, *RS-14* Resilience Scale, *PSS* Perceived Stress Scale, *PFI* Professional Fulfilment Index, *PROMIS* Patient Reported Outcomes Measurement Information System, *D/A* Depression & Anxiety, *RWBI* Resident Well-Being Index, *CDRS* Connor Davidson Resiliency Scale, *SMART* Stress Management and Resiliency Training, *SAS* Smith Anxiety Scale, *LASA* Linear Analog Self-Assessment Scale, *QOL* Quality of Life, *SHS* Subjective Happiness Scale, *GAD-7* Generalized Anxiety Disorder 7-item scale, *IQR* Inter-Quartile Range, *LSMD* Least Square Mean Differences, *NTC-RT* No-Treatment Control Group with Release Time, *NTC-RD* No-Treatment Control Group with Routine Duties, *RQS* Role Quality Scale, *CSI* Coping Strategy Indicator, *HMPSS* House's Measure of Perceived Stress Scale, *SSTAI* Spielberger State-Trait Anxiety Inventory, *CES-DS* Center for Epidemiological Studies Depression Scale, *CHI-PS* Cohen Hoberman Inventory of Physical Symptoms, *NSDA* National Survey on Drug Abuse

Table 3 Intervention design, method of measurement and outcomes for non-randomized and cross-sectional studies assessing interventions for physician wellness

Author, Year	Sample (*n*)	Mean Response Rate (%)	Description of Intervention	Use of Validated Tool	Outcome Assessment	Intervention Group Score(s) Pre-test	Intervention Group Score(s) During or Post-test	Control Group Score(s) Pre-test	Control Group Score(s) During or Post-test	Within Group *p*-value	Between Group *p*-value
Williamson et al., [37] 2021	386	65.9	1 year wellness curriculum with structured lectures, individualized assignments and wellness resources at 5 different sites. Control group contained 5 non-randomized sites that did not implement the curriculum.	yes	PRIME-MD (Depression) & MBI to correlate burnout to depression.	Depression 36.9%	Depression 6 month: 21.8% 1 year: 41.4%	Depression Pre-test: 43%	Depression 6 month: 32.2% 1 year: 32.6%	NR	Depression Pre-test: 0.35 6 month: 0.09 1 year: 0.22
Jacob & Lambert [38], 2021	292	20.2	1 h discussion on trauma exposure response, self-reflection and coping strategies	no	Basic Satisfaction Survey	N/A	Improved understanding Agree: 25% Strongly agree: 64% Understand signs of trauma Agree: 64% Strongly agree: 34% Understand coping mechanisms Agree: 57% Strongly agree: 31% Valuable Somewhat: 15% Moderate: 54% Extremely: 27%	N/A	N/A	N/A	N/A
Garcia et al., [39] 2021	36	56.9	2 protected personal days per year, modernized workspace and additional meal funds.	no	5PL Wellness Survey	5PL (mean) 3.43	5PL (mean) 6-month: 3.62 15-month: 3.55	N/A	N/A	0–6 months: 0.25 0–15 months: 0.62 6–15 months: 0.87	N/A
Jung et al., [40] 2021	30	66.7	4 h-long interactive educational sessions on the residents' specialty with their children present.	yes	5PL-Modified PFI	N/A	5PL (mean) Happy: 4.6 Worthwhile: 4.7 Satisfied: 4.75 Meaningful: 4.75 Connection: 4.2	N/A	N/A	N/A	N/A
Mohamed et al., [41] 2022	88	65.9	Peer support presentation to highlight ways to support each other during COVID-19 case surge.	no	Basic Satisfaction Survey	N/A	Check-ins per week 0 times: 16% 1–2 times: 43% 3–5 times: 24% 5–7 times: 17% How beneficial? Not at all: 10% Somewhat: 28% Moderately: 31% Very: 31% How did it help? Anxiety: 23% Load sharing: 19% Personal life: 37% Other: 21%	N/A	N/A	N/A	N/A

Table 3 (continued)

Author, Year	Sample (*n*)	Mean Response Rate (%)	Description of Intervention	Use of Validated Tool	Outcome Assessment	Intervention Group Score(s) Pre-test	Intervention Group Score(s) During or Post-test	Control Group Score(s) Pre-test	Control Group Score(s) During or Post-test	Within Group *p*-value	Between Group *p*-value
Dabrow et al., [42] 2006	535, 522*	47.1	24/7 service providing access to confidential evaluation, brief counselling, and referral services.	no	Basic Satisfaction Survey	N/A	Awareness of program 2004: 96.6% 2005: 92%	N/A	N/A	N/A	N/A
Hategan & Riddell [43], 2020	22	86.4	Wellness newsletters, virtual curriculum on resilience, and peer support groups.	no	Basic Satisfaction Survey	Perceived stress (mean) 5.5	Perceived stress (mean) 2.75	N/A	N/A	NR	N/A
Babbar et al., [44] 2019	29	100.0	Weekly 1 h yoga classes and nutrition/physical challenges (ex. increasing restful sleep).	yes	MBI, DASS-21 & FFMQ	MBI (mean ± SD) EE: 11.0 ± 3.2 DP: 6.0 ± 2.3 PA: 15.0 ± 3.0 DASS-21 (mean ± SD) Depression: 2.5 ± 1.4 Anxiety: 3.0 ± 1.3 Stress: 7.0 ± 2.4 FFMQ (mean ± SD) 128.0 ± 16.5	MBI (mean ± SD) EE: 9.5 ± 2.3 DP: 4.0 ± 3.8 PA: 15.0 ± 2.4 DASS-21 (mean ± SD) Depression: 2.0 ± 1.5 Anxiety: 2.0 ± 1.5 Stress: 5.0 ± 3.7 FFMQ (mean ± SD) 127.0 ± 15.1	N/A	N/A	MBI EE: 0.13 DP: 0.04 PA: 0.32 DASS-21 Depression: 0.47 Anxiety: 0.02 Stress: 0.08 FFMQ 0.32	N/A
Petrie et al., [45] 2022	1455, 1110**	25.1	Organizational: Reduce overtime, Doctors Wellness Committee and improved claiming overtime process. Individual: Mental health presentations, training courses, mentoring program, and resilience workshops.	yes	K10 & MABEL-based survey	K10 Psychological distress(mean ± SD) 18.6 ± 6.3 MABEL (mean ± SD) Suicidal ideation: 26 ± 11.8 Help-seeking confidence: 131 ± 58.2 Help-seeking behaviours: 38 ± 17.3	K10 Psychological distress(mean ± SD) 17.7 ± 6.5 MABEL (mean ± SD) Suicidal ideation: 20 ± 7.8 Help-seeking confidence: 157 ± 61.3 Help-seeking behaviours: 33 ± 12.9	N/A	N/A	K10 Psychological distress: 0.203 MABEL Suicidal ideation: 0.182 Help-seeking confidence: 0.376 Help-seeking behaviours: 0.511	N/A
Wothe et al., [46] 2022	NR	NR	First Tuesday off in months with 5 Tuesdays, easy-access counselling service and Vertical Mentoring Program.	no	Basic Satisfaction Survey	2017/18 Access to counselling: 85% Ability to schedule: 62% Mentorship opportunity: 77%	2018/19 Access to counselling: 95% Ability to schedule: 73% Mentorship opportunity: 81% 2019/20 Access to counselling: 98% Ability to schedule: 94% Mentorship opportunity: 88%	N/A	N/A	Access to counselling: 0.39 Ability to schedule: 0.005 Mentorship opportunity: 0.52	N/A

Table 3 (continued)

Author, Year	Sample (*n*)	Mean Response Rate (%)	Description of Intervention	Use of Validated Tool	Outcome Assessment	Intervention Group Score(s) Pre-test	Intervention Group Score(s) During or Post-test	Control Group Score(s) Pre-test	Control Group Score(s) During or Post-test	Within Group *p*-value	Between Group *p*-value
Baumann et al., [47] 2020	21	NR	Small group facilitated discussions on imposter syndrome. Part of larger series of resident wellness sessions.	no	5PL Wellness Survey	N/A	5PL (mean) Efficacy: 3.71 Helpful facilitator: 4.19 Recognize in self: 4.23 Recognize in colleagues: 3.24 Comfort discussing imposter syndrome: 3.76 Next steps knowledge: 3.48 Wellness during clinic block preference: 3.37	N/A	N/A	N/A	N/A
Ward et al., [48] 2022	55	50.9	Peer support program for wellness goals, interactive workshops, and participant-led wellness initiatives.	yes	CBI & WHO-5	CBI (mean) 48.1 WHO-5 (mean) 46.7	CBI (mean) 2 months: 46.1 7 months: 41.3 WHO-5 (mean) 2 months: 57.2 7 months: 61.6	N/A	N/A	CBI 0–2 months: 0.454 0–7 months: 0.077 WHO-5 0–2 months: 0.006 0–7 months: 0.006	N/A
Keyser et al., [49] 2021	254	NR	Peer Support Program like that of Brigham and Women's Hospital. Conversations about 15–30 min, in person or over telephone.	no	Basic Satisfaction Survey	N/A	Reported a representative sampling of qualitative feedback; all entirely positive	N/A	N/A	N/A	N/A
Fabreau et al., [50] 2013	67	74.6	Night float system, educational sessions on sleep hygiene, electronic handover tool, and simulation-based medical education curriculum.	no	5PL Wellness Survey regarding 24 h shifts	Resident wellness (mean ± SD) General wellness: 2.2 ± 0.1 Exposure to harm: 4.0 ± 0.1 Conflicting role demands: 2.9 ± 0.1 Healthy relationships: 2.6 ± 0.1 Feelings of isolation: 2.3 ± 0.1	Resident wellness (mean ± SD) General wellness: 2.6 ± 0.1 Exposure to harm: 2.8 ± 0.2 Conflicting role demands: 3.7 ± 0.1 Healthy relationships: 2.7 ± 0.2 Feelings of isolation: 1.9 ± 0.1	N/A	N/A	Resident wellness General wellness: 0.07 Exposure to harm: <0.001 Conflicting role demands: 0.002 Healthy relationships: 0.81 Feelings of isolation: 0.06	N/A

Table 3 (continued)

Author, Year	Sample (*n*)	Mean Response Rate (%)	Description of Intervention	Use of Validated Tool	Outcome Assessment	Intervention Group Score(s) Pre-test	Intervention Group Score(s) During or Post-test	Control Group Score(s) Pre-test	Control Group Score(s) During or Post-test	Within Group *p*-value	Between Group *p*-value
						Ability to deliver quality health care (mean ± SD) Potential for error: 3.2 ± 0.1 Skills: 3.5 ± 0.1 Continuity of care: 3.6 ± 0.1 Emotional labour: 2.2 ± 0.1 Efficiency: 3.4 ± 0.1	Ability to deliver quality health care (mean ± SD) Potential for error: 2.5 ± 0.1 Skills: 3.8 ± 0.1 Continuity of care: 3.6 ± 0.1 Emotional labour: 2.2 ± 0.1 Efficiency: 3.6 ± 0.1			Ability to deliver quality health care Potential for error: 0.003 Skills: 0.07 Continuity of care: 0.80 Emotional labour: 0.57 Efficiency: 0.15	
						Medical education experience (mean ± SD) Teaching: 3.0 ± 0.1 Skills: 3.3 ± 0.1 Learning: 3.1 ± 0.1 Supervision: 3.1 ± 0.1 Rotation disruptions: 4.1 ± 0.2	Medical education experience (mean ± SD) Teaching: 3.5 ± 0.1 Skills: 3.5 ± 0.1 Learning: 3.5 ± 0.1 Supervision: 2.7 ± 0.1 Rotation disruptions: 3.3 ± 0.1			Medical education experience Teaching: 0.0097 Skills: 0.14 Learning: 0.06 Supervision: 0.0099 Rotation disruptions: 0.003	
Awadallah et al., [51] 2023	32	59.4	1 year mentorship program; interns matched with coaches providing quarterly sessions for wellness and academic advice.	no	5PL Satisfaction Survey	N/A	5PL (mean ± SD) Coach Expectations: 4.47 ± 0.51 Encouragement (goals): 4.89 ± 0.32 Encouragement (reflect): 4.79 ± 0.42 Helpful: 4.77 ± 0.43 Experience Enjoyed: 4.44 ± 0.86 Reached goals: 4.42 ± 0.84 Effective (growth): 4.37 ± 0.89 Effective (well-being): 4.32 ± 0.95 Recommend: 4.44 ± 0.86	N/A	N/A	N/A	N/A
Ricker et al., [52] 2021	87	61.0	AWCIM 4.5-hour online Physician Well-being course on resilience, managing stress, preventing burnout. Had daily, self-selected, 10-minute, resilience activity for 2 weeks.	yes	MBI, DRS, CDRS, RSE, GS, Wellness & Satisfaction Surveys	MBI (mean ± SD) EE: 14.63 ± 1.6 DP: 6.06 ± 5.0 PA: 40.42 ± 8.1 (mean ± SD) DRS: 32.57 ± 4.0 CDRS: 31.77 ± 4.7 RSE: 14.3 ± 1.6 GS: 38.83 ± 4.3	MBI (mean ± SD) EE: 11.85 ± 8.3 DP: 4.21 ± 5.1 PA: 36.56 ± 15.3 (mean ± SD) DRS: 34.3 ± 4.7 CDRS: 33.23 ± 4.6 RSE: 14.98 ± 1.7 GS: 38.67 ± 5.0	N/A	N/A	MBI EE: 0.021 DP: 0.002 PA: 0.028 DRS: <0.001 CDRS: 0.012 RSE: 0.024 GS: 0.754	N/A

Table 3 (continued)

Author, Year	Sample (*n*)	Mean Response Rate (%)	Description of Intervention	Use of Validated Tool	Outcome Assessment	Intervention Group Score(s) Pre-test	Intervention Group Score(s) During or Post-test	Control Group Score(s) Pre-test	Control Group Score(s) During or Post-test	Within Group *p*-value	Between Group *p*-value
Fournier & Tourian [53], 2020	17	62.0	Peer mentoring program, matched by university based on preferences for mentor interests.	no	Basic Satisfaction Survey	N/A	Helpful Extremely: 0% Helpful: 36% Moderately: 18% Slightly: 36% Not: 9%	N/A	N/A	N/A	N/A
Ey et al., [54] 2016	139	92.5	Combination of educational outreach (ex. wellness promotion workshops) and direct care and consultation (ex. individual counselling).	no	Basic Satisfaction Survey	N/A	Quality Excellent/good: 95% Met expectations: 95% Met needs: 92% Recommend: 98% Satisfied (help): 92% Satisfied (service): 95% Helpful: 97% Come back: 95%	N/A	N/A	N/A	N/A
Chang et al., [55] 2023	84	85.7	Peer recognition program; acknowledged for achievements by colleagues on a public forum with points redeemable for actual prizes.	yes	SPFI	SPFI (mean) Total PF: 2.26 Overall BO: 1.55	SPFI (mean) Total PF: 2.39 Overall BO: 1.48	N/A	N/A	SPFI Total PF: 0.69 Overall BO: 0.82	N/A
Letica-Kriegel et al., [56] 2023	37	94.6	Psychologist- facilitated process group (6-week rotations) with open meditation and discussion of mental health topics	yes	2-item MBI	MBI (mean ± SD) BO: 2.56 ± 1.17 Callousness: 2.45 ± 1.82	MBI (mean ± SD) BO: 2.73 ± 1.50 Callousness: 2.41 ± 1.85	N/A	N/A	MBI BO: 0.64 Callousness: 0.91	N/A
Runyan et al., [57] 2016	12	75.0	4-week Physician as Leader rotation, with a wellness curriculum and regular wellness sessions on topics like burnout.	yes	MBI, SCS, PSS & JES	MBI (mean ± SD) PE: 24.78 ± 7.68 EE: 20.44 ± 9.36 Cynicism: 15.67 ± 8.94 (mean ± SD) SCS: 35.33 ± 6.23 PSS: 18.11 ± 6.70 JES: 110.56 ± 18.30	MBI (mean ± SD) PE: 26.89 ± 4.29 EE: 18.00 ± 9.88 Cynicism: 15.33 ± 8.07 (mean ± SD) SCS: 38.67 ± 4.82 PSS: 14.78 ± 7.05 JES: 122.11 ± 5.49	N/A	N/A	NR (due to small sample size)	N/A

Table 3 (continued)

Author, Year	Sample (*n*)	Mean Response Rate (%)	Description of Intervention	Use of Validated Tool	Outcome Assessment	Intervention Group Score(s) Pre-test	Intervention Group Score(s) During or Post-test	Control Group Score(s) Pre-test	Control Group Score(s) During or Post-test	Within Group *p*-value	Between Group *p*-value
Johnson et al., [58] 2022	254	16.8	"Take 10"; meditate or exercise for minimum of 10 min at least 3 times a week. Control group contains those who did not complete an average of 3 activities on 3 separate days per week, over a 4-week period.	yes	RWBI, self-reported burnout & PHQ-9	RWBI (mean): 2.93 Burnout score: 2.69 PHQ-9: 4.3 (minimal severity)	RWBI Burnout score 1 month: 2.79 2 month: 3.17 3 month: 3.67 4 month: 2.94	RWBI (mean): 3.77 Burnout score: 3.15 PHQ-9: 4.3 (minimal severity)	RWBI Burnout score 1 month: 3.5 2 month: 4.46 3 month: 4.1 4 month: 3.92	NR	RWBI: <0.01 Self-reported burnout Pre-test: 0.39 1 month: 0.14 2 month: 0.01 3 month: 0.48 4 month: 0.18
Cawyer et al., [59] 2022	35	60.0	1-year personal financial literacy curriculum (debt planning, investing and financial effects on wellbeing).	no	EWBI & FSS-CV	EWBI (median, IQR) 2 (1–5) FSS-CV (median, IQR) 22 (17–26)	EWBI (median, IQR) 1 (0–3) FSS-CV (median, IQR) 20 (18–23)	N/A	N/A	EWBI: 0.049 FSS-CV: 0.06	N/A
Shahid et al., [60] 2018	31	100.0	4-hour emotional intelligence training program to develop self-awareness, self-management, social awareness and social skills.	no	BOEQI	EI (median, IQR) Total: 110 (102–115) Self-perception: 106 (102–114) Self-expression: 106 (94–113) Interpersonal: 114 (108–119) Decision making: 112 (101–117) Stress management: 105 (98–111) Wellness: 104 (94–113)	EI (median, IQR) Total: 114 (108–123) Self-perception: 112 (106–119) Self-expression: 109 (100–115) Interpersonal: 117 (108–119) Decision making: 113 (104–124) Stress management: 111 (103–123) Wellness: 111 (99–118)	N/A	N/A	EI Total: 0.004 Self-perception: 0.016 Self-expression: 0.026 Interpersonal: 0.568 Decision making: 0.033 Stress management: <0.001 Wellness: 0.003	N/A
Aggarwal et al., [61] 2017	272	69.1	Lectures on burnout and wellness followed by 12-week curriculum of weekly wellness and happiness exercises.	no	Basic Satisfaction Survey	N/A	Subjective feedback from 4 of 5 medical specialties involved Internal: Overall well, but enthusiasm varied thru day. Anesthesia: Lacked organization in who's leading sessions Psychiatry: Overall well and good effort to engage, want more diversity in exercises Phys Med: Very well, included more time for stretching and enjoyed quiet exercises	N/A	N/A	N/A	N/A

Table 3 (continued)

Author, Year	Sample (*n*)	Mean Response Rate (%)	Description of Intervention	Use of Validated Tool	Outcome Assessment	Intervention Group Score(s) Pre-test	Intervention Group Score(s) During or Post-test	Control Group Score(s) Pre-test	Control Group Score(s) During or Post-test	Within Group *p*-value	Between Group *p*-value
Bisgaard et al., [62] 2021	59	51.0	Weekly healthy snack delivery, time off for doctors' visits, a holiday schedule, and fun events like book club, quarterly movie outings and running club.	yes	MBI & PGWBI	MBI (mean ± SD) PA: 32.81 ± 6.36 EE: 23.48 ± 11.20 DP: 10.94 ± 5.14 PGWBI (mean ± SD) Global: 68.10 ± 14.73 Anxiety: 14.06 ± 4.68 Depressed: 9.90 ± 1.11 Well-being: 11.48 ± 4.01 Self-control: 11.35 ± 3.57 General health: 11.29 ± 2.82 Vitality: 10.00 ± 2.16	MBI 6 months (mean ± SD) PA: 32.39 ± 5.77 EE: 23.00 ± 9.90 DP: 9.61 ± 4.39 2 year (mean ± SD) PA: 35.58 ± 8.10 EE: 24.42 ± 12.74 DP: 11.12 ± 7.70	N/A	N/A	MBI PA: 0.163 EE: 0.888 DP: 0.534 PGWBI Global: 0.097 Anxiety: 0.142 Depressed: 0.000 Well-being: 0.364 Self-control: 0.622 General health: 0.481 Vitality: 0.187	N/A
							PGWBI 6 months (mean ± SD) Global: 71.48 ± 11.18 Anxiety: 15.27 ± 4.16 Depressed: 11.85 ± 2.03 Well-being: 11.03 ± 3.10 Self-control: 11.76 ± 2.97 General health: 10.85 ± 2.48 Vitality: 10.73 ± 2.71 2 year (mean ± SD) Global: 76.62 ± 18.14 Anxiety: 16.35 ± 4.01 Depressed: 12.42 ± 3.05 Well-being: 12.42 ± 4.16 Self-control: 12.15 ± 2.52 General health: 11.73 ± 3.07 Vitality: 11.54 ± 4.37				

Table 3 (continued)

Author, Year	Sample (*n*)	Mean Response Rate (%)	Description of Intervention	Use of Validated Tool	Outcome Assessment	Intervention Group Score(s) Pre-test	Intervention Group Score(s) During or Post-test	Control Group Score(s) Pre-test	Control Group Score(s) During or Post-test	Within Group *p*-value	Between Group *p*-value
Hart et al., [63] 2019	46	63.0	"The Happiness Program"; 6 monthly discussion sessions on resilience, being conscious, honoring feelings and learning life lessons.	yes	MBI & 5PL Satisfaction Survey	MBI (mean ± SD) EE: 24.3 ± 9.8 DP: 14.2 ± 5.4 PA: 33.1 ± 5.0	MBI (mean ± SD) EE: 26.2 ± 9.7 DP: 15.8 ± 6.6 PA: 37.9 ± 5.3 5PL Survey (mean) Satisfaction: 1.47 Lessened burnout: 0.84 Improved wellness: 1.18	N/A	N/A	NR	N/A
Bui et al., [64] 2020	369	44.0	Mindfulness training and narrative medicine, program "wellness champions" to liaise with faculty, and grant to hire more workers and reduce burden.	yes	MBI & PHQ-2	N/A	MBI ($EE^3$27 or $DP^3$10) Burnout positive: 63.2% PHQ-2 (^{3}1 yes) Depression positive: 36.7%	N/A	N/A	N/A	N/A
Mari et al., [65] 2019	23	56.5	Based on physician feedback; improve on-call experience, increase social activities, provide wellness education and resources for primary care.	yes	CBI	CBI (mean ± SD) Personal: 63 ± 27 Work: 62 ± 26 Client: 46 ± 27	CBI (mean ± SD) Personal: 54 ± 24 Work: 54 ± 27 Client: 42 ± 29	N/A	N/A	CBI Personal: >0.05 Work: >0.05 Client: >0.05	N/A

PRIME-MD (Depression) Primary Care Evaluation of Mental Disorders– Patient Health Questionnaire for Depression, *MBI* Maslach Burnout Inventory, *NR* Not Reported, *N/A* Not Applicable, *5PL* 5-Point Likert scale, *PFI* Professional Fulfillment Index, *DASS-21* Depression Anxiety Stress Scale, *FFMQ* Five Facet Mindfulness Questionnaire, *SD* Standard Deviation, *EE* Emotional Exhaustion, *DP* Depersonalization, *PA* Personal Accomplishment, *K10* Kessler Psychological Distress Scale, *MABEL* Medicine in Australia Balancing Employment and Life, *CBI* Copenhagen Burnout Inventory, *WHO-5* World Health Organization Well-Being Index, *AWCIM* Andrew Weil Center for Integrative Medicine, *DRS* Dispositional Resilience Scale, *CDRS* Connor Davidson Resilience Scale, *RSE* Response to Stressful Experiences, *GS* Gratitude Scale, *SPFI* Stanford Professional Fulfilment Index, *PF* Professional Fulfilment, *BO* Burnout, *SCS* Self Compassion Scale, *PSS* Perceived Stress Scale, *JES* Jefferson Empathy Scale, *PE* Professional Efficacy, *RWBI* Resident Well-Being Index, *PHQ-9* Patient Health Questionnaire 9 - Depression Test Questionnaire, *EWBI* Expanding Well-Being Index, *FSS-CV* Financial Stress Scale– College Version, *IQR* Inter-Quartile Range, *BOEQI* Bar-On Emotional Quotient Inventory 2.0, *EI* Emotional Intelligence, *PGWBI* Psychological General Wellbeing Index, *PHQ-2* Patient Health Questionnaire 2

*Surveys taken twice, once in 2004 and once in 2005

**Population size differed at baseline vs. follow-up 2 years later

63, 65] of the studies employed systemic programs that were administered to participant groups as a whole and included peer support groups, adjustments to workload and social group activities. The remaining three (8.3%) [45, 46, 54] studies incorporated multiple interventions, having both individual and systemic approaches. The wellness programs may also be categorized by the facet(s) of wellness that they focused on. Thirteen (36.1%) [34–38, 40, 41, 47, 52, 57, 59–61] studies provided educational curriculum interventions that informed participants of the important components of wellness, including but not limited to coping strategies, signs of burnout and how to build resilience. Five (13.9%) [30, 32, 33, 44, 58] studies incorporated physical activities in their interventions. This included walks, yoga classes, and fitness challenges in attempt to improve mental wellness. In two (5.5%) [39, 62] studies, protected personal days and reduced overtime were new administrative policies introduced in the hopes of reducing physician stress and burnout. Nine (25%) [31, 42, 48, 49, 51, 53, 55, 56, 63] studies focused on social aspects of wellness, including interventions such as peer support programs or mental health counseling services. Lastly, seven (19.4%) [43, 45, 46, 50, 54, 64, 65] studies utilized a multifactorial approach, implementing interventions belonging to two or more of these general categories. Overall, the heterogeneity of wellness program interventions implemented necessitated the variation in their evaluation.

Study outcomes

Primary outcomes of each study were determined in various ways (Tables 2 and 3). Many used multiple assessment scales; eighteen (50%) [30–37, 44, 45, 48, 52, 57–59, 62–64] studies used two or more questionnaires to evaluate the efficacy of their wellness intervention. Most used a standardized wellness or burnout scale to assess participants, however, ten (27.8%) [38, 41–43, 46, 49, 51, 53, 54, 61] studies solely gathered qualitative feedback from participants regarding their satisfaction with the wellness intervention. The most common scale utilized was the Maslach Burnout Inventory (MBI), used in eleven (30.6%) [31–33, 37, 44, 52, 56, 57, 62–64] of the reported studies. Seven (19.4%) [30, 34–36, 45, 48, 65] other studies utilized questionnaires similarly structured to the MBI, such as the Copenhagen Burnout Inventory, the Connor Davidson Resiliency Scale, and the Perceived Stress Scale. Meanwhile, the final eight (22.2%) [39, 40, 47, 50, 55, 58–60] studies focused on generally evaluating positive wellness. This includes use of the Professional Fulfilment Index, various wellness indices, and basic 5-point Likert scaled wellness surveys. In studies with quantitative data, statistical evidence of an improvement in patient outcomes was determined by a p-value of less than 0.05. Studies with qualitative data simply reported percentages of participants who provided positive feedback.

Efficacy of interventions based on study type

The reported change in physician wellness varied extensively across the studies (Tables 2 and 3). Of the studies that performed statistical analyses, 62.5% [30–34, 36, 44, 46, 48, 50, 52, 58–60, 62] found at least one significant improvement in physician well-being, while 37.5% [35, 37, 39, 45, 55–57, 63, 65] showed no significant changes or lacked statistical reporting.

Five of the RCTs reported standard deviations for their data [31, 32, 34–36], allowing analysis of effect size estimates (ETable 5). Of the wellness metrics assessed, four studies had at least one large effect size ($g > 0.8$) [31, 32, 34, 36]. All RCTs evaluated had at least one medium effect size ($g > 0.5$), though the majority of the effect sizes were interpreted as small ($g < 0.5$) in two of the analyzed studies [32, 36].

11 of the cross-sectional studies (CSS) reported standard deviations for their data pre- and post-intervention [44, 45, 50, 52, 56, 57, 59, 60, 62, 63, 65], allowing analysis of effect size estimates for wellness metrics (ETable 6). Only three studies had at least one metric with a large effect size ($g > 0.8$) [50, 62, 63], and seven total had at least a medium effect size ($g > 0.5$) [44, 45, 50, 57, 62, 63, 65]. Nine of these analyzed studies (81.1%) were deemed to have a small effect size in the majority of their wellness metrics [45, 52, 56, 57, 59, 60, 62, 63, 65]. Conversely, almost all (92.3%) [37, 38, 40, 41, 43, 45, 47–49, 54, 61, 62] of the CSS that gathered feedback reported that most of their participants stated that they found the intervention helpful.

Efficacy of interventions based on intervention type

Of the RCTs reporting statistical analyses, there were three educational interventions [34–36], three physical activity interventions [30, 32, 33] and one social intervention [31]. All, except for one educational intervention [35], demonstrated at least some statistically significant improvement in physician wellness following the intervention implementation. One RCT – which was a socially-focused intervention – provided feedback received from the participants, which showed the majority found it helpful [31].

For CCS showing statistical significance, three (33.3%) [39, 44, 56] were educational interventions, two (22.2%) [36, 51] encouraged physical activity, one (11.1%) [62] was an administrative intervention, one (11.1%) [42] was a socially focused intervention and two (22.2.%) [46, 50] were combined interventions. With respect to the interventions in these studies, five (38.5%) [38, 40, 41, 47, 61] were educational, four (38.5%) [42, 49, 51, 53] were social, and three (23.1%) [43, 46, 54] were multifactorial.

Overall, these results indicate that some degree of improvement was shown in approximately three-quarters (76.9%) [34, 36, 38, 40, 41, 47, 52, 59–61] of the educational interventions, all (100.0%) [30, 32, 33, 44, 58] of the physical interventions, half (50.0%) [62] of the administrative interventions, two-thirds (66.7%) [31, 42, 48, 49, 51, 53] of the social interventions and just over half (57.1%) [43, 46, 50, 54] of multifactorial interventions.

Differences in interventions among physicians by sex, age, and wellness goals

Four (11.1%) [51, 55, 58, 64] studies compared the effects of wellness interventions on physicians by sex, while only one (2.8%) [64] study compared outcomes by age groups (Table 4). With respect to the wellness goals, female physicians were more likely to prioritize gratitude and mental health than compared to their male counterparts (29% vs. 7%, respectfully) [51]. However, male physicians were more likely to focus on lifestyle and exercise wellness goals [51]. Female physicians also showed improvements in professional fulfilment, burnout, work exhaustion and interpersonal disengagement score, while males saw deteriorations in the same scores [55]. Another study described a significantly significant relation between female participants and increased burnout rates [58]. The one study evaluating on both sex and age reported that 71.4% and 40.9% of female participants experienced burnout and depression, respectively [64]. Furthermore, this study also found that 52.9% and 31.5% of participants over age 30 screened positive for symptoms of burnout and depression, respectively [64]. This relationship was deemed statistically significant among those over the age of 30 ($p< 0.05$) [64].

Differences in interventions by medical specialty

Three (8.3%) [58, 60, 64] studies compared the effects of wellness interventions on physicians based on their medical specialty (Table 5). One study generalized specialities based on their qualities as "nonprocedural", "procedural" and "surgical". Reported mean burnout scores were based on the Resident Well-Being Index in each specialty, where procedural residents experienced the greatest burnout, followed by surgical residents and finally nonprocedural residents [58]. However, none of the three mean scores met the criteria of significant burnout in the population, as characterized by a value of 5 of higher on the index [58].

Another study evaluated two specialties, being pediatrics and Med-Peds. Pediatric residents saw no statistically significant changes overall in their emotional intelligence scores from the Bar-On Emotional Quotient Inventory [60]. However, Med-Peds residents had statistically significant improvements in their self-perception and interpersonal scores, as well as total emotional intelligence [60]. The final study divided surgical residents divided into groups of neurosurgery, obstetrics and gynecology, orthopedic surgery, otolaryngology, surgery (general, plastics, vascular, surgical oncology, colorectal, thoracic, transplant) and urology. They reported burnout rates ranging from 50 to 73.3% and depression rates from 20 to 45.8% across specialties. Furthermore, they found that orthopedic surgery residents showed the highest burnout, and while surgical residents represented the highest rates of depression [64].

Confidence in cumulative evidence

Based on the assessment utilizing the Grading of Recommendations, Assessment, Development and Evaluation (GRADE) approach, the total evidence from these studies was determined to have a 'low' certainty (Supplementary Table 4). Despite the number of RCTs summarized and the lack of consistent bias identified in our quality appraisal, the studies vary considerably in implementation and evaluation with inconsistent findings of improvement. The cumulative evidence is relevant and legitimate in the assessment of effective interventions for physician wellness.

Discussion

In this systematic review of wellness program interventions to improve physician wellness, we identified 36 studies, seven of which were randomized, including 6,708 physicians with highly diverse interventions, methods of outcome assessment and outcome measurement. Study outcomes were self-reported with low response rates or non-representative samples [31, 32, 35, 36, 43, 44, 54–56, 60]. Interventions led to statistical improvement in at least one metric in most studies [30–34, 36, 44, 46, 48, 50, 52, 58–60, 62], and when assessed, participant satisfaction was high.

Study interventions, although highly variable, were thematically categorizable (educational curriculum, physical activities, administrative policies, social aspects). Of these studies, interventions concerning either wellness education or physical activity seemed to be the most effective, with most of those groups' respective studies showing at least some degree of efficacy. By this definition, administrative interventions were the least effective. Physical activity interventions may improve physician wellness by boosting energy levels and lowering the risk of chronic diseases [20, 66], while educational interventions equip physicians with knowledge and skills to manage stress and foster better work-life balance [3, 67]. However, it is important to note the sample size of studies conducting strictly administrative interventions for wellness were limited, rendering comparisons difficult. Effect size estimates indicate that the RCT studies had a greater proportion of medium-to-large effect sizes when

Table 4 Studies assessing interventions for physician wellness by age and/or sex

Author, Year	Sex *n* (%) female	Outcome Assessment	Definition of Outcome	Outcome	Age Distribution *n* (%)	Outcome
Awadallah et al., [51] 2023	32 (34.4)	Basic Satisfaction Survey	Development of professional goals following intervention; types of goals	Commitment to exercise M: 10 (67%) F: 4 (57%) Healthy lifestyle goals M: 4 (27%) F: 1 (14%) Gratitude goals M: 1 (7%) F: 2 (29%)	NR	NR
Chang et al., [55] 2023	42 (50)	SPFI	Increase in professional fulfillment scores and decrease in overall burnout, work exhaustion and interpersonal disengagement scores	Mean Score Changes Professional fulfillment M: −0.22 F: 0.47 Overall burnout M: 0.16 F: −0.30 Work exhaustion M: 0.13 F: −0.40 Interpersonal disengagement M: 0.18 F: −0.23	NR	NR
Johnson et al., [58] 2022	116 (57.7)	RWBI & PHQ-9	Self-reported burnout in the last month	Burnout F: Odds ratio = 3.53*	NR	NR
Bui et al., [64] 2020	71 (44.1)	MBI & PHQ-2	Burnout High EE (^{3}27) OR High DP (^{3}10) Depression Yes to ^{3}1 PHQ-2 item	Burnout F: 50 (71.4%) Depression F: 29 (40.9%)	^{3}30: 73 (45.3%)	Burnout: ^{3}30: 37 (52.9%) * Depression: ^{3}30: 23 (31.5%)

M Male, *F* Female, *NR* Not Reported, *SPFI* Stanford Professional Fulfilment Index, *RWBI* Resident Wellbeing Index, *PHQ-9* Patient Health Questionnaire 9 - Depression Test Questionnaire, *MBI* Maslach Burnout Inventory, *PHQ-2* Patient Health Questionnaire 2, *EE* Emotional Exhaustion, *DP* Depersonalization

*Statistically significant value *($p < 0.05$)*

compared to the CSS, which had a majority of small effect sizes identified in their metrics. Our results suggest that wellness interventions have the potential to be effective in improving physician wellness and reducing burnout. However, the lack of standardization and consistency in both intervention implementation and the number of studies that conducted each type of intervention makes it difficult to ascertain how much of these improvements can be attributed to the intervention itself. Despite the heterogeneity of the studies, these findings improve our understanding of how we may compare the efficacy of different interventions for physician wellness, which may inform how they are implemented in the future.

Among the few studies that assessed factors such as sex, age, and medical speciality, the efficacy of physician wellness program interventions varied. The general finding most consistent within these studies was that higher burnout was reported in female physicians above 30 years of age, however, there were no clear differences in how effective the interventions were in improving wellness by these demographic factors. Further, the studies comparing medical specialties had narrow scopes in their study design, only comparing a few similar specialties to each other. Therefore, due to the limited number of studies concerning this topic, key information on targeted interventions is still unknown. Recent reviews identify similar difficulties in developing clear recommendations for physician burnout. Significant promise in organization-directed rather than individual-directed interventions has been reported [20, 68], though we have found that more recent studies may instead combine both approaches to provide a broader improvement in overall wellness. Furthermore, our categorization of the interventions within these studies into more specific groups may elucidate the aspects of general wellness they aim to improve.

The design of these studies and the reporting practices within the medical community need to be considered for the improvement of wellness program interventions for physicians. According to our critical appraisal, the overall quality of evidence was moderate to low. Most studies had small sample sizes or lacked control groups, which limits causal inference. Additionally, the use of self-reported measures introduces the potential for response bias, while small sample sizes reduce statistical power and limit the generalizability of observed effects [69, 70]. Studies with low response rates likely introduce bias and limitations that undermine the validity and generalizability of the findings, thereby hindering our ability to draw meaningful conclusions about intervention efficacy [38, 45, 58, 71, 72]. Further, the lack of RCTs in this review highlights the need for more rigorous testing of interventions in multi-centre studies across all specialties to improve the internal validity of findings. These factors collectively diminish the strength of the conclusions that can be drawn regarding the true efficacy of physician wellness interventions. The study design and assessment methods of physician wellness interventions and their outcomes should be considered for improved replicability and reliability [73–75].

However, evaluating these interventions presents significant challenges. Physician wellness is complex and subjective, making it challenging to assess using traditional quantitative metrics [3, 8, 15, 19, 73]. Various indices and scales exist, but no standardized metric allows for consistent comparisons across individuals [73]. Additionally, interventions targeting physician wellness may vary widely in their approach, duration, and effectiveness, further complicating the evaluation process [17, 18, 20]. Taking these challenges into consideration, researchers need to enhance the design and consistency of these studies by improving communication and developing standardized protocols, which will maximize the quality and reliability of study data [72, 76–78].

Nevertheless, the variability in reported prevalence rates of physician burnout distinguishes the complexity and multifactorial nature of the problem [7, 79]. Burnout prevalence may vary widely depending on factors related to study design like measurement tools and multiple definitions, making it challenging to obtain a clear and accurate understanding of the true prevalence of burnout among physicians [7, 79–81]. Further, physicians may be hesitant to acknowledge or report burnout due to fear of professional consequences or perceived lack of support, leading to underreporting and underestimation of the true prevalence [82, 83]. Thus, groups at potentially greater risk, based on specific age, gender, and/or medical specialty, are not clearly identifiable. The identification of such groups within the physician population can help inform the development and refinement of wellness interventions aimed at mitigating burnout and enhancing wellness [84, 85]. By recognizing specific subpopulations who may require additional support for their wellness, interventions can be tailored to address their unique needs effectively, thereby increasing the likelihood of success [86, 87]. Therefore, the findings of this review underscore the importance of standardized and replicable study designs to address the multifaceted aspects of physician wellness and the need for more rigorous research to better identify effective strategies for improving physician wellness.

Limitations

Firstly, the implementation of the interventions, the primary outcomes and their self-reported assessments and questionnaires were all unique, making it difficult to evaluate comparisons between studies. Our inclusion criteria aimed to solely review studies which characterized

Table 5 Studies assessing interventions for physician wellness by speciality type

Author, Year	Outcome Assessment	Definition of Outcome	Specialty Distribution *n* (%)	Outcome
Johnson et al., [58] 2022	RWBI & PHQ-9	Burnout RWBI score ^{3}5	Nonsurgical: 70 (34.8) Procedural: 48 (23.9) Surgical: 83 (41.3)	Mean RWBI Score Nonsurgical: 3.0 Procedural: 3.9 Surgical: 3.2
Shahid et al., [60] 2018	BOEQI	Increase in emotional intelligence score assessed by BOEQI	Pediatric: 20 (64.5) Med-Peds: 11 (35.5)	Median Change in EI (IQR) Total Peds: 3 (−2–6) Med-Peds: 8 (2–14) * Self-perception Peds: 1 (−4–6) Med-Peds: 6 (3–15) * Self-expression Peds: 4 (−1–7) Med-Peds: 6 (−2–13) Interpersonal Peds: −1 (−7-2) Med-Peds: 5 (−1–10) *Decision makingPeds: 1 (−4–8) Med-Peds: 8 (1–15) Stress management Peds: 6 (3–10) Med-Peds: 3 (2–13) Wellness Peds: 6 (−2–13) Med-Peds: 3 (0–10)
Bui et al., [64] 2020	MBI & PHQ-2	Burnout High EE (^{3}27) or High DP (^{3}10) Depression Yes to ^{3}1 PHQ-2 item	Neurosurgery: 12 (7.5) Obstetrics and Gynecology: 39 (24.2) Orthopedic Surgery: 15 (9.3) Otolaryngology: 20 (12.4) Surgery: 59 (36.7) Urology: 16 (9.9)	Burnout Neurosurgery: 5 (50) Obstetrics and Gynecology: 27 (73) Orthopedic Surgery: 11 (73.3) Otolaryngology: 11 (57.9) Surgery: 34 (58.6) Urology: 10 (62.5) Depression Neurosurgery: 5 (41.7) Obstetrics and Gynecology: 15 (38.5) Orthopedic Surgery: 3 (20) Otolaryngology: 4 (20) Surgery: 27 (45.8) Urology: 5 (31.3)

RWBI Resident Wellbeing Index, *PHQ-9* Patient Health Questionnaire 9 - Depression Test Questionnaire, *BOEQI* Bar-On Emotional Quotient Inventory, *EI* Emotional Intelligence, *IQR* Inter-Quartile Range, *MBI* Maslach Burnout Inventory, *PHQ-2* Patient Health Questionnaire 2, *EE* Emotional Exhaustion, *DP* Depersonalization

*Statistically significant value ($p < 0.05$)

the efficacy of their established interventions, however the definition of efficacy varied across the included studies. Some studies utilized statistical analysis to determine significant results, while others solely relied on optional and qualitative participant feedback. We did not limit studies to a certain style of intervention or to a specific, standardized evaluation method, which would have made it more feasible to compare results. This review may be subject to publication bias, as the included studies were limited to peer-reviewed literature and did not incorporate grey literature or unpublished studies. We may also have an overrepresentation of positive results in the evidence base, as studies reporting non-significant findings can be less likely to be published. This could affect the generalizability of our conclusions.

Additionally, this review had limited access to the inclusion of demographic factors in the evaluation of the wellness interventions. Although many studies reported on the demographics of their participants, very few assessed and compared how different groups separated by factors like sex, age and specialty were impacted by the interventions. This limited our ability to evaluate and characterize potential at-risk groups for whom tailored wellness interventions could be investigated. Finally, nearly all the studies were conducted in Canada and the United States, providing a narrow global perspective of the state of physician wellness program interventions. This may restrict the applicability of our findings for diverse healthcare systems and cultural contexts. That said, it may be that such studies are lacking in other geographical regions, and bringing this research to light may inform further investigation into how we may improve the wellness on physicians worldwide.

Conclusions

Through this systematic review, we identified various physician wellness program interventions and the standardized assessment tools used to evaluate their efficacy in physicians. All studies were self-reported surveys with vast heterogeneity in the outcomes of their participants' wellness, as well as limited information concerning differences related to sex, age, and medical specialty. Clear evidence identifying specific wellness interventions that are most effective is limited. However, when separating studies by intervention type, educational programs and physical activities were found to have the greatest proportion of studies showing at least some degree of improvement. Wellness interventions have been shown to have the potential to benefit physicians, but the complexity of individual wellness is an obstacle to optimizing and refining these interventions further. Future population-based longitudinal studies with assessments of demographic factors such as sex, age and specialty could possibly better define the prevalence of physician burnout, which will inform the improvement of more tailored and effective physician wellness interventions.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12913-025-12934-z.

Supplementary Material 1.

Acknowledgements
N/A.

Authors' contributions
Author Contributions: NK had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis. Concept and design: ES, MMS -Acquisition, analysis, or interpretation of data: all authors -Drafting of the manuscript: NK, MMS- Critical revision of the manuscript for important intellectual content: all authors-Statistical analysis: NK, MMS-Obtained funding: MMS-Administrative, technical, or material support: MG, GW-Supervision: MMS.

Funding
The Dept. of Medicine Research Chair in Physicians Health and Wellness, The Ottawa Hospital, University of Ottawa.

Data availability
All data analyzed during this study is referenced and included in this article and supplementary information file.

Declarations

Ethics approval and consent to participate
N/A.

Consent for publication
N/A.

Competing interests
MMS has received speaker fees/has been a consultant for AstraZeneca, Bayer, GlaxoSmithKline, Boehringer, Otsuka.

Received: 27 January 2025 / Accepted: 20 May 2025
Published online: 03 July 2025

References

1. Farr S, Berry JA, Berry DK, et al. The impact of the COVID-19 pandemic on resident physicians Well-Being in the surgical and primary care specialties in the united States and Canada. Cureus. 2021;13(11):e19677. https://doi.org/10.7759/cureus.19677.
2. Amanullah S, Ramesh Shankar R. The impact of COVID-19 on physician burnout globally: A review. Healthc Basel Switz. 2020;8(4):421. https://doi.org/10.3390/healthcare8040421.
3. West CP, Dyrbye LN, Erwin PJ, Shanafelt TD. Interventions to prevent and reduce physician burnout: a systematic review and meta-analysis. Lancet Lond Engl. 2016;388(10057):2272–81. https://doi.org/10.1016/S0140-6736(16)31279-X.
4. Grover S, Adarsh H, Naskar C, Varadharajan N. Physician burnout: A review. J Ment Health Hum Behav. 2018;23(2):78. https://doi.org/10.4103/jmhhb.jmhhb_47_19.
5. McCray LW, Cronholm PF, Bogner HR, Gallo JJ, Neill RA. Resident physician burnout: is there hope?? Fam Med. 2008;40(9):626–32.
6. De Hert S. Burnout in healthcare workers: prevalence, impact and preventative strategies. Local Reg Anesth. 2020;13:171–83. https://doi.org/10.2147/LRA.S240564.

7. Rotenstein LS, Torre M, Ramos MA, et al. Prevalence of burnout among physicians: A systematic review. JAMA. 2018;320(11):1131–50. https://doi.org/10.1001/jama.2018.12777.
8. Azam K, Khan A, Alam MT. Causes and adverse impact of physician burnout: A systematic review. J Coll Physicians Surg–Pak JCPSP. 2017;27(8):495–501.
9. Patel RS, Bachu R, Adikey A, Malik M, Shah M. Factors related to physician burnout and its consequences: A review. Behav Sci Basel Switz. 2018;8(11):98. https://doi.org/10.3390/bs8110098.
10. Dyrbye L, Shanafelt T. A narrative review on burnout experienced by medical students and residents. Med Educ. 2016;50(1):132–49. https://doi.org/10.1111/medu.12927.
11. Picton A. Work-life balance in medical students: self-care in a culture of self-sacrifice. BMC Med Educ. 2021;21(1):8. https://doi.org/10.1186/s12909-020-02434-5.
12. Wong AMF. Beyond burnout: looking deeply into physician distress. Can J Ophthalmol. 2020;55(3, Supplement 1):7–16. https://doi.org/10.1016/j.jcjo.2020.01.014.
13. Wallace JE, Lemaire JB, Ghali WA. Physician wellness: a missing quality indicator. Lancet. 2009;374(9702):1714–21. https://doi.org/10.1016/S0140-6736(09)61424-0.
14. Rothenberger DA. Physician burnout and Well-Being: A systematic review and framework for action. Dis Colon Rectum. 2017;60(6):567. https://doi.org/10.1097/DCR.0000000000000844.
15. Sinskey JL, Margolis RD, Vinson AE. The wicked problem of physician Well-Being. Anesthesiol Clin. 2022;40(2):213–23. https://doi.org/10.1016/j.anclin.2022.01.001.
16. Eskander J, Rajaguru PP, Greenberg PB. Evaluating wellness interventions for resident physicians: A systematic review. J Grad Med Educ. 2021;13(1):58–69. https://doi.org/10.4300/JGME-D-20-00359.1.
17. Thomas G, Lynch M, Spencer LH. A systematic review to examine the evidence in developing social prescribing interventions that apply a Co-Productive, Co-Designed approach to improve Well-Being outcomes in a community setting. Int J Environ Res Public Health. 2021;18(8):3896. https://doi.org/10.3390/ijerph18083896.
18. Melnyk BM, Kelly SA, Stephens J, et al. Interventions to improve mental health, Well-Being, physical health, and lifestyle behaviors in physicians and nurses: A systematic review. Am J Health Promot AJHP. 2020;34(8):929–41. https://doi.org/10.1177/0890117120920451.
19. Scheepers RA, Boerebach BCM, Arah OA, Heineman MJ, Lombarts KMJMH. A systematic review of the impact of physicians' occupational Well-Being on the quality of patient care. Int J Behav Med. 2015;22(6):683–98. https://doi.org/10.1007/s12529-015-9473-3.
20. 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. https://doi.org/10.1001/jamainternmed.2016.7674.
21. Woo BFY, Lee JXY, Tam WWS. The impact of the advanced practice nursing role on quality of care, clinical outcomes, patient satisfaction, and cost in the emergency and critical care settings: a systematic review. Hum Resour Health. 2017;15(1):63. https://doi.org/10.1186/s12960-017-0237-9.
22. Romani M, Ashkar K. Burnout among physicians. Libyan J Med. 2014;9(1):23556. https://doi.org/10.3402/ljm.v9.23556.
23. Lemaire JB, Ewashina D, Polachek AJ, Dixit J, Yiu V. Understanding how patients perceive physician wellness and its links to patient care: A qualitative study. PLoS ONE. 2018;13(5):e0196888. https://doi.org/10.1371/journal.pone.0196888.
24. Mathura P, Pascheto I, Ringrose J, Ramsay G. What about physician wellness? Impact of a quality improvement intervention. BMJ Open Qual. 2025;14(1):e003078. https://doi.org/10.1136/bmjoq-2024-003078.
25. Parsons M, Bailitz J, Chung AS, et al. Evidence-Based interventions that promote resident wellness from the Council of emergency residency directors. West J Emerg Med. 2020;21(2):412–22. https://doi.org/10.5811/westjem.2019.11.42961.
26. Study Quality Assessment Tools| NHLBI, NIH, Accessed. 25 Apr 2025. https://www.nhlbi.nih.gov/health-topics/study-quality-assessment-tools.
27. Lockwood C, Munn Z, Porritt K. Qualitative research synthesis: methodological guidance for systematic reviewers utilizing meta-aggregation. JBI Evid Implement. 2015;13(3):179. https://doi.org/10.1097/XEB.0000000000000062.
28. Guyatt GH, Oxman AD, Vist GE, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. BMJ. 2008;336(7650):924–6. https://doi.org/10.1136/bmj.39489.470347.AD.
29. Lewin S, Booth A, Glenton C, et al. Applying GRADE-CERQual to qualitative evidence synthesis findings: introduction to the series. Implement Sci. 2018;13(1):2. https://doi.org/10.1186/s13012-017-0688-3.
30. Kavanaugh J, Hardison ME, Rogers HH, White C, Gross J. Assessing the impact of a Shinrin-Yoku (Forest Bathing) intervention on physician/healthcare professional burnout: A randomized, controlled trial. Int J Environ Res Public Health. 2022;19(21):14505. https://doi.org/10.3390/ijerph192114505.
31. Huang L, Harsh J, Cui H, et al. A randomized controlled trial of balint groups to prevent burnout among residents in China. Front Psychiatry. 2019;10:957. https://doi.org/10.3389/fpsyt.2019.00957.
32. Taylor J, McLean L, Richards B, Glozier N. Personalised yoga for burnout and traumatic stress in junior Doctors. Postgrad Med J. 2020;96(1136):349–57. https://doi.org/10.1136/postgradmedj-2019-137413.
33. Loewenthal J, Dyer NL, Lipsyc-Sharf M, et al. Evaluation of a Yoga-Based Mind-Body intervention for resident physicians: A randomized clinical trial. Glob Adv Health Med. 2021;10:21649561211001038. https://doi.org/10.1177/21649561211001038.
34. Sood A, Prasad K, Schroeder D, Varkey P. Stress management and resilience training among department of medicine faculty: a pilot randomized clinical trial. J Gen Intern Med. 2011;26(8):858–61. https://doi.org/10.1007/s11606-011-1640-x.
35. Spilg EG, Kuk H, Ananny L, et al. The impact of stress management and resilience training (SMART) on academic physicians during the implementation of a new health information system: an exploratory randomized controlled trial. PLoS ONE. 2022;17(4):e0267240. https://doi.org/10.1371/journal.pone.0267240.
36. Saadat H, Snow DL, Ottenheimer S, Dai F, Kain ZN. Wellness program for anesthesiology residents: a randomized, controlled trial. Acta Anaesthesiol Scand. 2012;56(9):1130–8. https://doi.org/10.1111/j.1399-6576.2012.02705.x.
37. Williamson K, Lank PM, Olson A, Cheema N, Lovell E. A positive depression screen is associated with emergency medicine resident burnout and is not affected by the implementation of a wellness curriculum. West J Emerg Med. 2021;22(6):1341–6. https://doi.org/10.5811/westjem.2021.9.52016.
38. Jacob KM, Lambert N. Trauma exposure response: how secondary trauma affects personal and professional life. MedEdPORTAL J Teach Learn Resour. 2021;17:11192. https://doi.org/10.15766/mep_2374-8265.11192.
39. Garcia DI, Pannuccio A, Gallegos J, et al. Resident-Driven wellness initiatives improve resident wellness and perception of work environment. J Surg Res. 2021;258:8–16. https://doi.org/10.1016/j.jss.2020.06.028.
40. Jung MJ, Libaw JS, Manuel SP, Kathiriya IS, Srejic U, Gandhi S. Interactive anesthesiology educational program improves wellness for anesthesiologists and their children. J Clin Anesth. 2021;70:110192. https://doi.org/10.1016/j.jclinane.2021.110192.
41. Mohamed BA, Fowler WK, Thakkar M, Fahy BG. The BUDDYS system: A unique peer support strategy among anaesthesiology residents during the COVID-19 pandemic. Turk J Anaesthesiol Reanim. 2022;50(Supp1):S62–7. https://doi.org/10.5152/TJAR.2021.21247.
42. Dabrow S, Russell S, Ackley K, Anderson E, Fabri PJ. Combating the stress of residency: one school's approach. Acad Med J Assoc Am Med Coll. 2006;81(5):436–9. https://doi.org/10.1097/01.ACM.0000222261.47643.d2.
43. Hategan A, Riddell T. Bridging the Gap: responding to resident burnout and restoring well-being. Perspect Med Educ. 2020;9(2):117–22. https://doi.org/10.1007/s40037-020-00567-3.
44. Babbar S, Renner K, Williams K. Addressing obstetrics and gynecology trainee burnout using a Yoga-Based wellness initiative during dedicated education time. Obstet Gynecol. 2019;133(5):994–1001. https://doi.org/10.1097/AOG.0000000000003229.
45. Petrie K, Stanton K, Gill A, Simmons J, Harvey SB. Effectiveness of a multimodal hospital-wide Doctor mental health and wellness intervention. BMC Psychiatry. 2022;22(1):244. https://doi.org/10.1186/s12888-022-03908-0.
46. Wothe J, Bosacker L, Nalluri H, Cullen MJ, Brunsvold ME. Improving mental health support, career transitions and access to health care for surgical residents. J Surg Educ. 2022;79(2):286–9. https://doi.org/10.1016/j.jsurg.2021.09.014.
47. Baumann N, Faulk C, Vanderlan J, Chen J, Bhayani RK. Small-Group discussion sessions on imposter syndrome. MedEdPORTAL J Teach Learn Resour. 2020;16:11004. https://doi.org/10.15766/mep_2374-8265.11004.
48. Ward MC, Crinall K, McDonald R, et al. The kindness COVID-19 toolkit: a mixed-methods evaluation of a programme designed by Doctors in training for Doctors in training. BMJ Open. 2022;12(11):e060575. https://doi.org/10.1136/bmjopen-2021-060575.

49. Keyser EA, Weir LF, Valdez MM, Aden JK, Matos RI. Extending peer support across the military health system to decrease clinician burnout. Mil Med. 2021;186(Suppl 1):153–9. https://doi.org/10.1093/milmed/usaa225.
50. Fabreau G, Elliott M, Khanna S, et al. Shifting perceptions: a pre-post study to assess the impact of a senior resident rotation bundle. BMC Med Educ. 2013;13:115. https://doi.org/10.1186/1472-6920-13-115.
51. Awadallah NS, Jones TS, Christian N, et al. Proactive coaching in general surgery internship: incorporating Well-being practices into resident professional life. J Surg Educ. 2023;80(2):177–84. https://doi.org/10.1016/j.jsurg.2022.09.021.
52. Ricker M, Brooks AJ, Bodine S, Lebensohn P, Maizes V. Well-being in residency: impact of an online physician Well-being course on resiliency and burnout in incoming residents. Fam Med. 2021;53(2):123–8. https://doi.org/10.22454/FamMed.2021.314886.
53. Fournier M, Tourian L. A peer mentoring initiative across medical residency programs. Ment Health Soc Incl. 2019;24(1):33–40. https://doi.org/10.1108/MHSI-09-2019-0027.
54. Ey S, Moffit M, Kinzie JM, Brunett PH. Feasibility of a comprehensive wellness and suicide prevention program: A decade of caring for physicians in training and practice. J Grad Med Educ. 2016;8(5):747–53. https://doi.org/10.4300/JGME-D-16-00034.1.
55. Chang J, Saggar V, Cortijo-Brown A, et al. Improving physician well-being and reducing burnout using a peer-to-peer recognition program. AEM Educ Train. 2023;7(2):e10861. https://doi.org/10.1002/aet2.10861.
56. Letica-Kriegel AS, Connolly MR, Westfal ML, et al. Implementation, feasibility, and perception of facilitated process groups in surgical residency. J Surg Educ. 2023;80(7):987–93. https://doi.org/10.1016/j.jsurg.2023.04.002.
57. Runyan C, Savageau JA, Potts S, Weinreb L. Impact of a family medicine resident wellness curriculum: a feasibility study. Med Educ Online. 2016;21:30648. https://doi.org/10.3402/meo.v21.30648.
58. Johnson WR, Davidson M, Nagler A, Terhune KP. Take 10: A resident Well-Being initiative and burnout mitigation strategy. J Surg Educ. 2022;79(2):322–9. https://doi.org/10.1016/j.jsurg.2021.09.022.
59. Cawyer CR, Blanchard C, Kim KH. Financial literacy and physician wellness: can a financial curriculum improve an obstetrician/gynecologist resident and fellow's Well-Being? AJP Rep. 2022;12(1):e64–8. https://doi.org/10.1055/s-0041-1742268.
60. Shahid R, Stirling J, Adams W. Promoting wellness and stress management in residents through emotional intelligence training. Adv Med Educ Pract. 2018;9:681–6. https://doi.org/10.2147/AMEP.S175299.
61. Aggarwal R, Deutsch JK, Medina J, Kothari N. Resident wellness: an intervention to decrease burnout and increase resiliency and happiness. MedEdPORTAL J Teach Learn Resour. 2017;13:10651. https://doi.org/10.15766/mep_2374-8265.10651.
62. Bisgaard E, Clark A, Hester C, et al. Resident engagement in a wellness program in a large academic residency: A Follow-Up after two years of wellness. J Surg Educ. 2021;78(5):1430–7. https://doi.org/10.1016/j.jsurg.2021.01.013.
63. Hart D, Paetow G, Zarzar R. Does implementation of a corporate wellness initiative improve burnout?? West J Emerg Med. 2019;20(1):138–44. https://doi.org/10.5811/westjem.2018.10.39677.
64. Bui AH, Ripp JA, Oh KY, et al. The impact of program-driven wellness initiatives on burnout and depression among surgical trainees. Am J Surg. 2020;219(2):316–21. https://doi.org/10.1016/j.amjsurg.2019.10.027.
65. Mari S, Meyen R, Kim B. Resident-led organizational initiatives to reduce burnout and improve wellness. BMC Med Educ. 2019;19(1):437. https://doi.org/10.1186/s12909-019-1756-y.
66. Seward MW, Marso CC, Soled DR, Briggs LG. Medicine in motion: addressing physician burnout through fitness, philanthropy, and interdisciplinary community Building. Am J Lifestyle Med. 2020;16(4):462–8. https://doi.org/10.1177/1559827620983782.
67. Krasner MS, Epstein RM, Beckman H, et al. Association of an educational program in mindful communication with burnout, empathy, and attitudes among primary care physicians. JAMA. 2009;302(12):1284–93. https://doi.org/10.1001/jama.2009.1384.
68. De Simone S, Vargas M, Servillo G. Organizational strategies to reduce physician burnout: a systematic review and meta-analysis. Aging Clin Exp Res. 2021;33(4):883–94. https://doi.org/10.1007/s40520-019-01368-3.
69. Faber J, Fonseca LM. How sample size influences research outcomes. Dent Press J Orthod. 2014;19(4):27–9. https://doi.org/10.1590/2176-9451.19.4.027-029.ebo.
70. Serdar CC, Cihan M, Yücel D, Serdar MA. Sample size, power and effect size revisited: simplified and practical approaches in pre-clinical, clinical and laboratory studies. Biochem Med. 2021;31(1):010502. https://doi.org/10.11613/BM.2021.010502.
71. Mazor KM, Clauser BE, Field T, Yood RA, Gurwitz JH. A demonstration of the impact of response bias on the results of patient satisfaction surveys. Health Serv Res. 2002;37(5):1403–17. https://doi.org/10.1111/1475-6773.11194.
72. Barriball KL, While AE. Non-response in survey research: a methodological discussion and development of an explanatory model. J Adv Nurs. 1999;30(3):677–86. https://doi.org/10.1046/j.1365-2648.1999.01117.x.
73. Lall MD, Gaeta TJ, Chung AS, et al. Assessment of physician Well-being, part two: beyond burnout. West J Emerg Med. 2019;20(2):291–304. https://doi.org/10.5811/westjem.2019.1.39666.
74. Brady KJS, Kazis LE, Sheldrick RC, Ni P, Trockel MT. Selecting physician well-being measures to assess health system performance and screen for distress: conceptual and methodological considerations. Curr Probl Pediatr Adolesc Health Care. 2019;49(12):100662. https://doi.org/10.1016/j.cppeds.2019.100662.
75. Bart R, Ishak WW, Ganjian S, et al. The assessment and measurement of wellness in the clinical medical setting: A systematic review. Innov Clin Neurosci. 2018;15(09–10):14–23.
76. Sammut R, Griscti O, Norman IJ. Strategies to improve response rates to web surveys: A literature review. Int J Nurs Stud. 2021;123:104058. https://doi.org/10.1016/j.ijnurstu.2021.104058.
77. Phillips AW, Reddy S, Durning SJ. Improving response rates and evaluating nonresponse bias in surveys: AMEE guide 102. Med Teach. 2016;38(3):217–28. https://doi.org/10.3109/0142159X.2015.1105945.
78. VanGeest JB, Johnson TP, Welch VL. Methodologies for improving response rates in surveys of physicians: a systematic review. Eval Health Prof. 2007;30(4):303–21. https://doi.org/10.1177/0163278707307899.
79. Hewitt DB, Ellis RJ, Hu YY, et al. Evaluating the association of multiple burnout definitions and thresholds with prevalence and outcomes. JAMA Surg. 2020;155(11):1043–9. https://doi.org/10.1001/jamasurg.2020.3351.
80. Yates SW. Physician stress and burnout. Am J Med. 2020;133(2):160–4. https://doi.org/10.1016/j.amjmed.2019.08.034.
81. Ortega MV, Hidrue MK, Lehrhoff SR, et al. Patterns in physician burnout in a Stable-Linked cohort. JAMA Netw Open. 2023;6(10):e2336745. https://doi.org/10.1001/jamanetworkopen.2023.36745.
82. Cook AF, Arora VM, Rasinski KA, Curlin FA, Yoon JD. The prevalence of medical student mistreatment and its association with burnout. Acad Med J Assoc Am Med Coll. 2014;89(5):749–54. https://doi.org/10.1097/ACM.0000000000000204.
83. Cheng MY, Neves SL, Rainwater J, et al. Exploration of mistreatment and burnout among resident physicians: a Cross-Specialty observational study. Med Sci Educ. 2020;30(1):315–21. https://doi.org/10.1007/s40670-019-00905-z.
84. Knight A, Shakeshaft A, Havard A, Maple M, Foley C, Shakeshaft B. The quality and effectiveness of interventions that target multiple risk factors among young people: a systematic review. Aust N Z J Public Health. 2017;41(1):54–60. https://doi.org/10.1111/1753-6405.12573.
85. Steffens NK, LaRue CJ, Haslam C, et al. Social identification-building interventions to improve health: a systematic review and meta-analysis. Health Psychol Rev. 2021;15(1):85–112. https://doi.org/10.1080/17437199.2019.1669481.
86. Whitlock EP, Eder M, Thompson JH, et al. An approach to addressing subpopulation considerations in systematic reviews: the experience of reviewers supporting the U.S. Preventive services task force. Syst Rev. 2017;6(1):41. https://doi.org/10.1186/s13643-017-0437-3.
87. Shahzad M, Upshur R, Donnelly P, et al. A population-based approach to integrated healthcare delivery: a scoping review of clinical care and public health collaboration. BMC Public Health. 2019;19(1):708. https://doi.org/10.1186/s12889-019-7002-z.

Publisher's Note