JAMA | Original Investigation

Association Between Learning Environment Interventions and Medical Student Well-being
A Systematic Review

Lauren T. Wasson, MD, MPH; Amberle Cusmano, MA; Laura Meli, MSEd; Irene Louh, MD, PhD; Louise Falzon, PGDipInf; Meghan Hampsey; Geoffrey Young, PhD; Jonathan Shaffer, PhD, MS; Karina W. Davidson, PhD, MASc

IMPORTANCE Concerns exist about the current quality of undergraduate medical education and its effect on students' well-being.

OBJECTIVE To identify best practices for undergraduate medical education learning environment interventions that are associated with improved emotional well-being of students.

DATA SOURCES Learning environment interventions were identified by searching the biomedical electronic databases Ovid MEDLINE, EMBASE, the Cochrane Library, and ERIC from database inception dates to October 2016. Studies examined any intervention designed to promote medical students' emotional well-being in the setting of a US academic medical school, with an outcome defined as students' reports of well-being as assessed by surveys, semistructured interviews, or other quantitative methods.

DATA EXTRACTION AND SYNTHESIS Two investigators independently reviewed abstracts and full-text articles. Data were extracted into tables to summarize results. Study quality was assessed by the Medical Education Research Study Quality Instrument (MERQSI), which has a possible range of 5 to 18; higher scores indicate higher design and methods quality and a score of 14 or higher indicates a high-quality study.

FINDINGS Twenty-eight articles including at least 8224 participants met eligibility criteria. Study designs included single-group cross-sectional or posttest only (n = 10), single-group pretest/posttest (n = 2), nonrandomized 2-group (n = 13), and randomized clinical trial (n = 3); 89.2% were conducted at a single site, and the mean MERSQI score for all studies was 10.3 (SD, 2.11; range, 5-13). Studies encompassed a variety of interventions, including those focused on pass/fail grading systems (n = 3; mean MERSQI score, 12.0), mental health programs (n = 4; mean MERSQI score, 11.9), mind-body skills programs (n = 7; mean MERSQI score, 11.3), curriculum structure (n = 3; mean MERSQI score, 9.5), multicomponent program reform (n = 5; mean MERSQI score, 9.4), wellness programs (n = 4; mean MERSQI score, 9.0), and advising/mentoring programs (n = 3; mean MERSQI score, 8.2).

CONCLUSIONS AND RELEVANCE In this systematic review, limited evidence suggested that some specific learning environment interventions were associated with improved emotional well-being among medical students. However, the overall quality of the evidence was low, highlighting the need for high-quality medical education research.

Editorial page 2195

Related article page 2214

Supplemental content

CME Quiz at jamanetworkcme.com

Author Affiliations: Center for Behavioral Cardiovascular Health, Columbia University Medical Center, New York, New York (Wasson, Cusmano, Meli, Louh, Falzon, Davidson); George Washington University, Washington, DC (Hampsey); Association of American Medical Colleges, Washington, DC (Young); Department of Psychology, College of Liberal Arts and Sciences, University of Colorado at Denver (Shaffer); NewYork-Presbyterian Hospital, New York, New York (Davidson).

Corresponding Author: Karina W. Davidson, PhD, MASc, Center for Behavioral Cardiovascular Health, Columbia University Medical Center, 622 W 168th St, PH 9-314, New York, NY 10032 (kd2124@cumc.columbia.edu).

JAMA. 2016;316(21):2237-2252. doi:10.1001/jama.2016.17573

Medical schools strive to educate knowledgeable, caring, and professional physicians and pay particular attention to opportunities for improving the undergraduate medical education (UME) learning environment as they realize its influence on the education of future physicians.[1]

A critical element of the learning environment is its effect on student well-being. Although matriculating US medical students begin training with significantly lower rates of depression and burnout and report better mental and emotional quality of life than other college-educated young adults,[2] their reported well-being decreases during the UME years. The reported rate of moderate to severe depression is approximately 14% and of burnout symptoms is 52%—higher than reported by other graduate students or population control samples.[3,4] Studies indicate that up to 11% of medical students report suicidal ideation.[5]

GWB *General Well-Being Schedule*

MERSQI *Medical Education Research Study Quality Instrument*

NOS-E *Newcastle-Ottawa Scale–Education*

UME *undergraduate medical education*

The Association of American Medical Colleges includes in its vision for improving medical education "the health and well-being of learners."[6] This systematic review evaluated the association between UME learning environment interventions and the emotional well-being of students.

Key Points

Question What undergraduate medical education learning environment interventions are associated with improved emotional well-being among medical students?

Findings In a systematic review of the medical literature, only 28 articles described empirically evaluated interventions and only 3 included randomization, so methodologic rigor was limited. However, some data support preclinical pass/fail grading, mental health programs, wellness programs, mentoring programs, curricular restructuring, and multicomponent program reform.

Meaning There is limited evidence to support learning environment interventions for improvement of emotional well-being among medical students. High-quality research is needed.

Methods

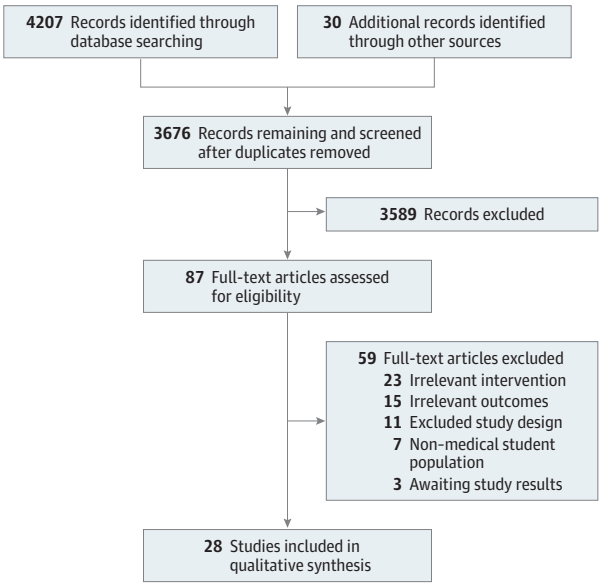
Search Strategy

Potentially relevant articles were identified (**Figure**) by searching the biomedical electronic databases Ovid MEDLINE, EMBASE, the Cochrane Library, and ERIC from database inception dates to October 2016 (eAppendix in the Supplement). Additional records were identified by scanning the reference lists of relevant studies and reviews published between May 2011 and October 2016 and by using the "similar articles" feature in PubMed and the "cited reference search" in Web of Science. We searched for gray literature ("that which is produced on all levels of government, academics, business and industry in print and electronic formats, but which is not controlled by commercial publishers")[7] through ongoing trial registries, academic dissertations, and websites of relevant organizations (eg, Association of American Medical Colleges) (eAppendix in the Supplement).

Selection Criteria

Studies had to have examined the outcomes associated with any intervention aiming to promote students' emotional well-being in the setting of an academic US medical school. The well-being outcome had to be obtained through surveys, semistructured interviews, or other quantitative methods. Open-ended response formats were excluded because their methodologic quality could not be appraised with the instrument used in this review. Medical education interventions measured with open-ended responses have been reviewed and appraised elsewhere.[8,9]

Figure. Review and Selection of Articles on the Association Between Learning Environment Interventions and Medical Student Well-being



Methodologic Quality Rating

Study quality was assessed using the Medical Education Research Study Quality Instrument (MERSQI), which was developed to appraise the methodologic quality of quantitative medical education research.[10] MERSQI scores have been positively correlated with editorial decisions to publish and with the presence of external funding for the research conducted.[10] The instrument is based on 10 design and methods criteria: study design, number of institutions studied, response rate, data type, internal structure, content validity, criterion validity, appropriateness of data analysis, complexity of analysis, and outcome level. These criteria form 6 domains, each with a maximum score of 3 and a minimum of 0 or 1, that sum to produce a total score that ranges from 5 to 18.

The MERSQI was preferred to the Newcastle-Ottawa Scale-Education (NOS-E), another assessment tool for medical education research quality, because it was found to have

Table 1. Study Comparison and Outcome Measures

Source	Design	Population	Comparisons	Sample Sizes	Outcome Measures	Main Results
Pass/Fail Grading Systems						
Bloodgood et al,[13] 2009	Nonrandomized 2-group	First- and second-year medical students	Cohort with pass/fail grading system vs earlier cohort with 5-interval grading system (A/B/C/D/F)	N = 281 Pass/fail = 140 5-interval = 141	Questions regarding satisfaction with school, satisfaction with personal life Dupuy General Well-being Schedule[14]	Pass/fail vs graded after semester 1 (scores): Anxiety: 18.14 (vs 15.98) (P = .002); depression: 17.62 (vs 15.89) (P < .001); well-being: 13.02 (vs 11.02) (P < .001); self-control: 15.51 (vs 14.12) (P < .001); vitality: 14.60 (vs 12.15) (P < .001); general health: 12.56 (vs 11.48) (P = .02) Pass/fail vs graded after semester 2 (scores): Anxiety: 19.01 (vs 17.65) (P = .05); depression: 17.61 (vs 16.65) (P = .05); well-being: 13.09 (vs 12.20) (P = .03); self-control: 15.10 (vs 14.45) (P = .13); vitality: 15.16 (vs 13.31) (P = .001); general health: 11. 99 (vs 11.24) (P = .15) Pass/fail vs graded after semester 3 (scores): Anxiety: 17.02 (vs 14.55) (P = .001); depression: 16.92 (vs 15.08) (P = .001); well-being: 12.37 (vs 10.74) (P < .001); self-control: 15.13 (vs 14.40) (P = .08); vitality: 14.10 (vs 11.95) (P < .001); general health: 11.25 (vs 10.84) (P = .47) Pass/fail vs graded after semester 4 (scores): Anxiety: 14.08 (vs 14.20) (P = .86); depression: 15.56 (vs 15.35) (P = .71); well-being: 10.59 (vs 10.40) (P = .67); self-control: 14.61 (vs 14.42) (P = .63); vitality: 12.88 (vs 12.06) (P = .11); general health: 11.30 (vs 11.31) (P = .99)
Rohe et al,[15] 2006	Nonrandomized 2-group	First- and second-year medical students	Cohort with pass/fail grading system vs earlier cohort with 5-interval grading system (A/B/C/D/F)	N = 81 Pass/fail = 40 5-interval = 41	Perceived Stress Scale[16,17] Profile of Mood States[18] Perceived Cohesion Scale[19]	Pass/fail vs graded at end of first year (score): Perceived Stress Scale: 10.9 (SD, 6.2) vs 13.8 (SD, 6.4) (P = .02); Profile of Mood States: 13.0 (SD, 23.5) vs 32.0 (SD, 39.0) (P = .02); Perceived Cohesion Scale: 37.8 (SD, 5.5) vs 32.9 (SD, 8.4) (P = .01) Pass/fail vs graded at end of second year (score): Perceived Stress Scale: 15.8 (SD, 6.8) vs 20.5 (SD, 7.8) (P = .01); Profile of Mood States: 47.1 (SD, 31.9) vs 64.6 (SD, 40.5) (P = .07); Perceived Cohesion Scale: 33.8 (SD, 8.0) vs 29.0 (SD, 9.9) (P = .02)
Reed et al,[20] 2011	Nonrandomized ≥2-group	First- and second-year medical students	Institutions with pass/fail grading systems vs institutions with ≥3-interval grading systems (eg, honors/pass/fail)	N = 2056 n = 1192 responded Pass/fail = 701 ≥3-interval = 491	Maslach Burnout Inventory[21] Perceived Stress Scale[16,17] Medical Outcomes Study Short Form[22,23]	Perceived Stress Scale: β = 1.91; 95% CI, 1.05-2.78 (P < .001) Mental quality of life: β = −2.79; 95% CI, −4.09 to 1.5 (P < .001) Burnout: OR, 1.85; 95% CI, 1.24-2.01 (P < .001) Seriously considered dropping out of medical school in the past year: OR, 1.91; 95% CI, 1.30-2.80 (P = .001)
Mental Health Programs						
Thompson et al,[24] 2010	Nonrandomized 2-group	Third-year medical students	Cohort with multipronged mental health program vs earlier cohort without program	N = 120 Program cohort = 62 Earlier cohort = 58	Center for Epidemiologic Studies Depression Scale[25] question on suicidal ideation	Pre-/postintervention: Depressive symptoms, 26/44 (59.1%) vs 14/58 (24.1%); χ^2_2 = 12.84 (P < .01); suicidal ideation, 13/43 (30.2%) vs 1/33 (3.0%); χ^2_1 = 13.05 (P < .001)
Downs et al,[26] 2014	Single-group cross-sectional or posttest only	All years		N = 1008 (program) n = 343 (program and screen)	Patient Health Questionnaire 9[27,28]	No. not provided; all P values reported as "nonsignificant" Among those screened, mental health service utilization: Year 1: 11.5%; year 4: 15.0%; χ^2_3: 1.27 Among those screened, suicide risk: Year 1: 8.8%; year 4: 6.2%; χ^2_3 = 0.45

(continued)

Table 1. Study Comparison and Outcome Measures (continued)

Source	Design	Population	Comparisons	Sample Sizes	Outcome Measures	Main Results
Mental Health Programs (continued)						
Seritan et al,[29] 2015	Nonrandomized ≥2-group	All years	Cohort with mental health program vs earlier cohort without program and vs national average	No sample size for No. of students referred to services ACGME Graduation Survey, n = 525	American Medical Colleges Graduation Questionnaire[30]	Mental health service self-referral (No. not provided): Time 1, 50%; time 2, 88%; time 3, 91% Other referral: Time 1: 50%; time 2: 12%; time 3: 9% Pre- vs last postintervention satisfaction scores are shown for the program (national scores from that year in parentheses): Personal counseling: 2009: 3.5 (3.7) vs 2013: 4.4 (4.0) Student mental health services: 2009: 3.5 (3.6) vs 2013: 4.3 (4.0) Stress management programs (postintervention only): 2009: 3.6 (3.8) vs 2013: 4.3 (3.9)
Moutier et al,[31] 2012	Single group cross-sectional or post-test only	All years		N = 498 (132 screened)	Items from Patient Health Questionnaire 9[27,28] measuring mental health service referral rate	Referred to mental health professional based in part on Patient Health Questionnaire 9, of those screened: 15/132 (11%)
Mind-Body Skills Education/Training Programs						
Erogul et al,[32] 2014	Randomized clinical trial	First-year medical students	Mindfulness-based stress reduction intervention vs control (randomized)	N = 58 Intervention = 28 Control = 30	Perceived Stress Scale[16,17] Self-Compassion Scale[33]	Change in case from pre- to postintervention (change score): Perceived Stress Scale, 3.63; 95% CI, 0.37-6.89 (P = .03); Self-Compassion Scale, 0.58; 95% CI, 0.23-0.92 (P = .002) Change in case from preintervention to 6-mo follow-up (change score): Perceived Stress Scale, 2.91; 95% CI, −0.37 to 6.19 (P = .08); Self-Compassion Scale, 0.56; 95% CI, 0.25-0.87 (P = .001)
Holtzworth-Munroe et al,[34] 1985	Randomized clinical trial	First- and second-year medical students	Mind-body program vs control (randomized)	N = 40 Intervention = 20 Control = 20	Spielberger State-Trait Anxiety Inventory[35] Anxiety in test and social situation questionnaire Tension and depression questionnaire Self-esteem measure Stress questionnaire	Intervention vs control at follow-up (score) More aware of tension: $F_{5,18}$ = 37.16 (P < .001); dealing better with school stress: $F_{5,18}$ = 5.05 (P < .04); anxiety before test: $F_{1,22}$ = 10.42 (P < .005)
Kraemer et al,[36] 2016	Nonrandomized 2-group	First- and second-year medical students	Mind-body program vs control (nonrandomized)	N = 52 Intervention = 28 Control = 24	Distress Tolerance Scale[37] Perceived Stress Scale 10[16,17] Positive Affect Negative Affect Schedule[38]	Changes in distress tolerance (change score): Mind-body, 0.53; t = −2.81; 95% CI, 0.92-0.14 (P = .01); control: 0.25; t = −1.6695% CI, −0.06 to 0.55 (P = .11)
Rosenzweig et al,[39] 2003	Nonrandomized 2-group	Second-year medical students	Mindfulness-based stress reduction program vs control (nonrandomized)	N = 302 Intervention = 140 Control = 162	Profile of Mood States[18]	Profile of Mood States total mood disturbance for intervention vs control (score): Intervention, 38.7 (SD, 33.3) vs 31.8 (SD, 33.8) (P = .05); control: 28.0 (SD, 31.2) vs 38.6 (SD, 32.8) (P < .001); interaction: d = −0.18 (P < .001)
Finkelstein et al,[40] 2007	Nonrandomized 2-group	Second-year medical students	Mind-body elective vs control (nonrandomized)	N = 72 Intervention = 26 Control = 46	Symptom Checklist 90 Anxiety Subscale[41] Profile of Mood States[18] Perceived Stress of Medical School Scale[42] 2-Item Depression Index[43]	Time/group interaction for scores: Anxiety (Symptom Checklist-90): $F_{1,2}$ = 3.95 (P < .05); Profile of Mood States: $F_{1,2}$ = 3.77 (P < .05); Perceived Stress of Medical School Scale: $F_{1,2}$ = .11 (P value reported as "nonsignificant")
Greeson et al,[44] 2015	Single-group pretest/posttest	All years	Before vs after mind-body skills intervention	N = 44	Cognitive and Affective Mindfulness Scale-Revised[45] Perceived Stress Scale[16,17] Open-ended feedback	Pre-/postintervention (score): Perceived Stress Scale, 29.73 (SD, 9.61) vs 20.25 (SD, 9.03); t (33) = 7.90; d = 1.38 (P < .001) Mindfulness: 29.24 (SD, 5.54) vs 33.88 (SD, 6.13); t (33) = 5.27; d = 0.92 (P < .001)
Bond et al,[46] 2013	Single group pretest/posttest	First- and second-year medical students	Before vs after mind-body course	N = 27	Cohen's Perceived Stress Scale[47,48] Self-Regulation Questionnaire[49] Self-Compassion Scale[33] Jefferson Scale of Physician Empathy[50]	Pre-/postintervention (change score): Perceived stress: −0.05 (SD, 0.62); d = .14 (P = .70); self-regulation: 0.13 (SD, 0.2); d = −0.41 (P = .003); self-compassion: 0.28 (SD, 0.61); d = −0.55 (P = .04); empathy: 0.11 (SD, 0.5); d = −0.13 (P = .30)

(continued)

Table 1. Study Comparison and Outcome Measures (continued)

Source	Design	Population	Comparisons	Sample Sizes	Outcome Measures	Main Results
Curriculum Structure						
Reed et al,[20] 2011	Nonrandomized ≥2-group	First- and second-year medical students	7 Institutions' curriculum structures	N = 2056 (1192 responded)	Maslach Burnout Inventory[21] Perceived Stress Scale[16,17] Medical Outcomes Study Short Form[22,23]	Association between clinical experiences and the following scores: Perceived stress: β = 0.02; 95% CI, −0.10 to 0.13 (*P* = .79); burnout: OR, 1.01; 95% CI, 0.98-1.05 (*P* = .42; mental quality of life: β = 0.00; 95% CI, −0.16 to 0.16 (*P* = .98); serious thoughts of dropping out: OR, 0.96; 95% CI, 0.93-1.00 (*P* = .03) Association between testing experiences and the following scores: Perceived stress: β = 0.29; 95% CI, 0.10-0.84 (*P* = .003); burnout: OR, 1.10; 95% CI, 0.89-1.23 (*P* = .09); mental quality of life: b - β = −0.63; 95% CI, −0.291 to 0.96 (*P* < .001); serious thoughts of dropping out: OR, 1.19; 95% CI, 1.12-1.27 (*P* < .001) Association between No. of tests and the following scores: Perceived stress: β = −0.02; 95% CI, −0.6 to 0.03 (*P* = .48); burnout: OR, 0.99; 95% CI, 0.97-1.01 (*P* = .19); mental quality of life: β = 0.03; 95% CI, −0.05 to 0.04 (*P* = 44); serious thoughts of dropping out: OR, 1.00; 95% CI, 0.97-1.02 (*P* = .82)
Camp et al,[51] 1994	Nonrandomized 2-group	First- and second-year medical students	Problem-based vs lecture-based learning	N = 275 Problem-based learning = 60 Lecture-based learning = 215	Zung Self-Rating Depression Scale[52]	Depression problem-based learning vs lecture-based learning (score): Overall: OR, 0.42; 95% CI, 0.14-1.21 (*P* = .07); adjustment for sex and self-actualization: OR, 0.45; 95% CI, 0.14-1.42 (*P* = .14)
Kornitzer et al,[53] 2005	Cross-sectional posttest only	All cohorts		N = 92	Questions regarding program attendance factors, subjective medical school transition factors, program ratings and student perceptions, and academic benefits of program	Underrepresented in medicine group (No. not provided): Gained confidence: 85.7%; made transition easier: 100%; made friends: 100% Humanities and medicine group (No. not provided): Gained confidence: 97%; made transition easier: 97%; made friends: 93.9%
Multicomponent Program Reform						
Drolet and Rodgers,[54] 2010	Single-group cross-sectional or posttest only	All years		N = 116	Satisfaction survey	Positive experience with Student wellness committee: 95% (No. not provided)
Fleming et al,[55] 2013	Single-group cross-sectional or posttest only (for the outcome measure relevant to this review)	All years		N = 245	Vanderbilt University student affairs survey	Reported that college's design contributed meaningfully or somewhat meaningfully to university experience: 91% (No. not provided)
Real et al,[56] 2015	Single-group cross-sectional or posttest only	All years		N = 450	Maslach Burnout Inventory[21] Primary Care Evaluation of Mental Disorders[27,57] Participation survey Perception of burnout survey	Level of burnout within aspects of program (score: 0 = more burnout; 100 = less burnout): Faculty mentors: 70; annual retreats: 58.6; student-led programming committee: 64; overall wellness program: 69.2 Faculty mentors correlation with the following scores: emotional exhaustion: *r* = −0.27; depersonalization: *r* = −0.22; personal accomplishment: *r* = 0.19 Annual retreats correlation with the following scores: Emotional exhaustion: *r* = −0.32; depersonalization: *r* = −0.32; personal accomplishment: *r* = 0.16 Student-led programming committee correlation with the following scores: Emotional exhaustion: *r* = −0.31; depersonalization: *r* = −0.3; personal accomplishment: *r* = 0.23 Overall wellness program correlation with the following scores: Emotional exhaustion: *r* = −0.32; depersonalization: *r* = −0.23; personal accomplishment: *r* = 0.1

(continued)

Table 1. Study Comparison and Outcome Measures (continued)

Source	Design	Population	Comparisons	Sample Sizes	Outcome Measures	Main Results
Multicomponent Program Reform (continued)						
Slavin et al,[58] 2014	Nonrandomized >2-group	First- and second-year medical students	Cohorts with different phases of multiprogram reform implementation	N = 875-890	Center for Epidemiological Studies Depression Scale[25] Spielberger State-Trait Anxiety Inventory[35] Perceived Stress Scale[16,17] Perceived Cohesion Scale[19] American Medical Colleges Graduation Questionnaire[30]	Moderate or severe depression according to the Center for Epidemiological Studies Depression Scale (No. not provided): End year 1: Cramér V = 0.16; phase 1 vs control: 21% vs 27% (*P* value reported as "nonsignificant"); phase 1/2 vs control: 18% vs 27% (*P* value reported as "nonsignificant"); phase 1/2/3 vs control: 11% vs 27% (*P* < .05) End year 2: Cramér V = 0.18; phase 1 vs control: 17% vs 32% (*P* < .05); phase 1/2 vs control: 18% vs 32 (*P* < .05); phase 1/2/3 vs control: 16% vs 32% 32 (*P* < .05) Anxiety (No. not provided): End year 1: Cramér V = 0.23; phase 1 vs control: 45% vs 55% (*P* value reported as "nonsignificant"); phase 1/2 vs control: 31% vs 55% (*P* < .05); phase 1/2/3 vs control: 31% vs 55% (*P* < .05) End year 2: Cramér V = 0.18; phase 1 vs control: 61% vs 60% (*P* value reported as "nonsignificant"); phase 1/2 vs control: 39% vs 60% (*P* < .05); phase 1/2/3 vs control: 46% vs 60% (*P* < .05) Stress (score): End year 1: η^2 = 0.06; phase 1 vs control: 14.9 (SD, 6.7) vs 16.3 (SD, 7.4) (*P* value reported as "nonsignificant"); phase 1/2 vs control: 13 (SD, 6.8) vs 16.3 (SD, 7.4) (*P* < .05); phase 1/2/3 vs control: 12.1 (SD, 6.1) vs 16.3 (SD, 7.4) (*P* < .05) End year 2: partial η^2 = 0.05; phase 1 vs control: 14.4 (SD, 5.8) vs 16.9 (SD, 7.3) (*P* < .05); phase 1/2 vs control: 13.9 (SD, 6.4) vs 16.9 (SD, 7.3 (*P* < .05)); phase 1/2/3 vs control: 13.5 (SD, 6.8) vs 16.9 (SD, 7.3) (*P* < .05) Cohesion (score): End year 1: partial η^2 = 0.03; phase 1 vs control: 8.1 (SD, 1.7) vs 7.9 (SD, 2.1) (*P* value reported as "nonsignificant"); phase 1/2 vs control: 8.5 (SD, 2.1) vs 7.9 (SD, 2.1) (*P* < .05); phase 1/2/3 vs control: 8.8 (SD, 1.8) vs 7.9 (SD, 2.1) (*P* < .05) End year 2: partial η^2 = 0.02; phase 1 vs control: 8.2 (SD, 1.6) vs 7.7 (SD, 2.0) (*P* < .05); phase 1/2 vs control: 8.5 (SD, 2.2) vs 7.7 (SD, 2.0) (*P* < .05); phase 1/2/3 vs control: 8.1 (SD, 2.0) vs 7.7 (SD, 2.0) (*P* < .05)
Strayhorn,[59] 1989	Nonrandomized 2-group	First-year medical students	Cohort with multicomponent program reform vs earlier cohort University of North Carolina, Chapel Hill vs comparison school	Responders = 478 (original sample size not reported)	Learning Environment Questionnaire[60] Rand Health Insurance Questionnaires[61] Environment stresses questionnaire Social support questionnaire	New vs old curriculum stress questionnaire: Overall fewer stresses: t_{223} = −1.7 (*P* = .09); less perceived stress from social and recreational sources (*P* = .03); no reduction in financial-related stress (*P* value not reported) New vs old curriculum mental well-being: Greater overall well-being: t_{197} = −2.04 (*P* = .04); greater sense of positive well-being (*P* < .001); greater sense of vitality (*P* < .001); less depression (*P* < .001); less anxiety (*P* < .001); social well-being: t_{223} = −1.66 (*P* = .10) New vs old curriculum social support: No perceived difference in availability of social supports: t_{227} = −0.36 (*P* = .72); less class advisor support (*P* = .002); class advisors less willing to listen (*P* = 003); class advisors less willing to help with personal problems (*P* < .001); concerned about students' welfare (*P* = .003); greater support from administrators (*P* = .05); could rely on administrators when things got tough (*P* = .01); perceived level of support from fellow students, friends, significant others (*P* value not reported) Time-control (University of North Carolina, Chapel Hill vs comparison) learning environment: Fewer environmental stressors: F_{1467} = 6.41 (*P* = .01); greater mental well-being: F_{1460} = 9.32 (*P* = .002); greater social well-being: F_{1466} = 5.37 (*P* = .02); no difference in social support: F_{1477} = 0.01 (*P* = .91)

(continued)

Table 1. Study Comparison and Outcome Measures (continued)

Source	Design	Population	Comparisons	Sample Sizes	Outcome Measures	Main Results
Miscellaneous Wellness Programs						
Whitehouse et al,[62] 1996	Randomized clinical trial	First-year medical students	Self-hypnosis intervention vs control (randomized)	N = 35 Intervention = 21 Control = 14	Medical history Profile of Mood States Brief Symptom Inventory University of California, Los Angeles Loneliness Scale	Time-group intervention analysis of score: Brief Symptom Inventory Anxiety: $F_{3,96} = 2.96$ ($P < .05$) At examination period, self-hypnosis participants had significantly lower stressfulness scores: $t_{30} = 2.11$ ($P < .05$)
Goetzel et al,[63] 1984	Single-group cross-sectional or posttest only	First-, second-, and third-year medical students		N = 26	Group Environment Scale	Agreement with statement on 1- to 5-point Likert scale: "I am no longer as lonely; I feel more together with people": 3.33 of 5 (*P* value not reported)
Lee and Graham,[64] 2001	Single-group cross-sectional or posttest only	First- and second-year medical students		N = 66	Questionnaire related to the wellness elective	Students appreciated that the wellness elective helped them realize the importance of personal well-being, gave permission for self-care and an opportunity to find collegiality, and provided various coping strategies: 4/22 (18.2%) strongly agree; 17/22 (77.3%) agree (*P* value not reported) Students felt that the wellness elective overemphasized stress itself and devalued the worth of hard work; realistic expectations offered in this course seemed discouraging: 1/22 (4.5%) agree (*P* value not reported)
Kushner et al,[65] 2011	Single-group cross-sectional or posttest only	Second-year medical students		N = 343 (9 related to mental and emotional health)	Form relating to goal and achievement	Self-reported achievement of mental/emotional health behavior change goals: 6/9 (66.7%) agree (*P* value not reported)
Group-Based Faculty Advisor/Mentor Programs						
Sastre et al,[66] 2010	Nonrandomized 2-group	First-, second-, and third-year medical students	Cohort in an advisory college program vs earlier cohort in a faculty advisory program	N = 318 Cohort in program = 103 Earlier cohort = 215	Questionnaires on perceived effectiveness of the system and role of advisor in promoting wellness and career counseling	Advisory college program vs faculty advisory program wellness advising (No. not provided) I feel comfortable discussing my personal stress with my advisor: 62% vs 24%; $\chi^2 = 40.9$ ($P < .001$) I feel comfortable discussing my mental health with my advisor: 51% vs 27%; $\chi^2 = 31.84$ ($P < .001$) Satisfaction with how well advisors promoted wellness (No. not provided) 27% vs 72% ($P < .001$)
Coates et al,[67] 2008	Nonrandomized 2-group	Fourth-year medical students	Cohort in a mentoring program vs earlier cohort	N = 100 Cohort in program = 70 Earlier cohort = 30	25-Item telephone survey	Cohort with mentoring program vs earlier cohort: Feels connected with faculty: 14/30 (47%) vs 49/70 (70%) Feels connected with classmates: 11/30 (37%) vs 30/70 (43%)
Ficklin et al,[68] 1983	Single-group cross-sectional or posttest only	First-year medical students		N = 151	Survey assessing 12 personal needs of first-year medical students	Program helpfulness (only descriptive summary of results provided): Becoming better acquainted with peers Becoming close to some classmates Helping students with anxieties of starting school

Abbreviations: ACGME, Accreditation Council for Graduate Medical Education; OR, odds ratio.

generally higher interrater reliability (0.68-0.89)[11] than the NOS-E. This may be due to its more objective assessments of design strengths and weaknesses, although it omits items on the comparability of groups and blinding.[11] Although there are no defined cutoff values differentiating high-quality from low-quality study methods, 1 study used a MERSQI score of 14.0 or higher as an a priori cutoff of high quality.[12]

Data Extraction

Two review authors (L.M. and L.F.) independently scanned the title or abstract of all search results to determine which studies required further assessment, investigated all potentially relevant articles as full text, selected studies to include in this review, assigned a MERSQI score for each, and calculated a mean quality score across studies. Data disagreements were resolved by consultation with the third and fourth review authors (L.T.W. and I.L.). The original intention noted in the study protocol was to conduct a meta-analysis, but because of the considerable variation in the interventions, study designs, and outcomes, we did not pool the studies quantitatively, as they were judged to not be combinable.[69]

Results

The literature search yielded 4207 publications, of which 28 met the eligibility criteria for this systematic review (Figure). Publications were excluded if they were irrelevant or did not meet the inclusion criteria; for example, we excluded publications that focused on medical residents rather than medical students, measured academic rather than well-being outcomes, or contained interventions not focused on the learning environment. The studies included at least 8224 student participants (1 study did not report a sample size) and encompassed a variety of designs, including single-group cross-sectional or posttest only (n = 10), single-group pretest/posttest (n = 2), nonrandomized 2-group (n = 13), and randomized clinical trial (RCT; n = 3) designs; 89.2% were conducted at a single site. They had a wide range of approaches to improving students' well-being that are categorized and described below (pass-fail grading systems [n = 3], mental health programs [n = 4], mind-body skills education/training [n = 7], curriculum structure [n = 3], multicomponent program reform [n = 5], wellness programs [n = 4], and group-based faculty advisor/mentor programs [n = 3]). Individual study results are described below and statistical details are provided for many key findings; additional results and methods are shown in **Table 1** and **Table 2**. The included studies' methodologic rigor varied, with MERSQI scores ranging from 5.0 to 13.0 (mean score, 10.3; SD, 2.11 [n = 28]). The mean MERSQI score in published medical education studies, as assessed in another review, was 10.0.[10] The studies with the highest-quality methods crossed all types of interventions and all types of outcome measures. The highest-scored categories tested interventions involving pass/fail grading, mental health programs, and mind-body skills education/training.

Pass/Fail Grading System (Mean MERSQI Score, 12.0)

Bloodgood et al[13] (n = 281; MERSQI, 11.5) and Rohe et al[15] (n = 81; MERSQI, 12.0) each described that a cohort of preclinical students graded according to a pass/fail grading system, compared with an earlier student cohort evaluated according to a 5-interval grading system (A/B/C/D/F), reported statistically significantly better well-being. They reported less anxiety, depression,[13] and stress[15] and better well-being[13] and group cohesion scores at various study time points.[15] These 2 studies differed, however, in the durability of improvements. Bloodgood et al[13] found no difference at 2 years between the cohort of students with a 2-year pass/fail system compared with a cohort of students with a 5-interval system on measures of anxiety (General Well-Being Schedule [GWB][14] anxiety subscore [range, 3-28; lower scores indicate more severe distress]; mean, 14.08 vs 14.20; $P = .86$), depression (GWB[14] depression subscore [range, 2-22; lower scores indicate more severe distress]; mean, 15.56 vs 15.35; $P = .71$), or well-being (GWB[14] well-being subscore [range, 3-18; lower scores indicate more severe distress]; mean, 10.59 vs 10.40; $P = .67$). Rohe et al[15] reported a persistent difference at 2 years between grading cohorts on a measure of stress (Perceived Stress Scale[16] [range, 0-40; higher score indicates more stress]; mean, 15.8 [SD, 6.8] vs 20.5 [SD, 7.8]; $P = .01$) and speculated that this difference was due to continuing reports of elevated group cohesion (Perceived Cohesion Scale[19] [range, 0-36; higher scores indicate more cohesion]; mean, 33.8 [SD, 8.0] vs 29.0 [SD, 9.9]; $P = .02$).

Reed et al[20] (n = 2056; MERSQI, 12.5) compared well-being among students at different medical schools with grading systems that were categorized as either having 3 or more intervals (eg, honors/pass/fail) or pass/fail and found that systems with 3 or more intervals were associated with statistically significantly more stress (β = 1.91; 95% CI, 1.05-2.78; $P < .001$) and burnout (odds ratio, 1.58; 95% CI, 1.24 to 2.01; $P < .001$), and a higher likelihood of considering withdrawing from medical school (odds ratio, 1.91; 95% CI, 1.30-2.80; $P = .001$).

Mental Health Programs (Mean MERSQI Score, 11.9)

Thompson et al[24] (n = 120; MERSQI, 11.5) evaluated a multipronged program aimed at reducing mental health stigma and making services more accessible. The study found that significantly smaller proportions of the student cohort exposed to the program compared with the prior student cohort reported symptoms of mild or probable depression (14/58 [24.1%] vs 26/44 [59.1%]; $P < .01$) and suicidal ideation (1/33 [3.0%] vs 13/43 [30.2%]; $P < .001$).[26] Seritan et al[29] (number of participants not reported; MERSQI, 11.5) examined a different multipronged mental health/wellness program offering prevention, support, and enhanced clinical services, which was associated with improved student ratings of personal counseling, mental health, and stress management services.[29] Percentages of self-referral to mental health services increased from a baseline rate of 50% to a postintervention rate of 91%. For both findings, statistical significance was not reported.[29]

Table 2. Methods of Included Studies

Source	No. of Sites	Overall Sample Size	Women, No. (%)	Intervention	Evaluation	MERSQI Score	Aim
Pass/Fail Grading Systems							
Bloodgood et al,[13] 2009	Single site	N = 281	5-Interval (A/B/C/D/F): 62% Pass/fail: 46%	Changed first- and second-year grading system from 5-interval letter grades to pass/fail grading system in first 2 preclinical years	Self-assessment	11.5	Measure the association of change in grading systems on medical student satisfaction and psychological well-being
Rohe et al,[15] 2006	Single site	N = 81	5-Interval (A/B/C/D/F): 26/41 (63%) Pass/fail: 20/40 (50%)	Replaced 5-interval grading system for first preclinical year with a modified pass/fail system (grading included pass/marginal pass requiring student action for remediation/fail) during first preclinical year	Self-assessment	12	Measure the sustained and immediate effects of a pass/fail grading system on stress, mood, group cohesion, and test anxiety
Reed et al,[20] 2011	Multisite	N = 2056 (1192 responded)	550/1192 (47%)	Multisite survey of 2 different grading scales: pass/fail and ≥3-interval (eg, honors/pass/fail, honors/high pass/pass/marginal pass/fail)	Self-assessment	12.5	Examine the relationship among curriculum structure, grading scales, and student well-being
Mental Health Programs							
Thompson et al,[24] 2010	Single site	N = 120		Multipronged intervention for third-year students aimed at (1) reducing barriers to mental health treatment by reducing stigma via faculty education, mental health curriculum, including lectures and a student handbook and (2) fully confidential and reduced/no-cost counseling services	Self-assessment	11.5	Test the effectiveness of an intervention meant to reduce depressive symptoms and suicidal ideation
Downs et al,[26] 2014	Single site	N = 1008 (program) n = 343 (program and screen)	Year 1: 93/148 (63%) Year 2: 34/65 (52%) Year 3: 27/49 (55%) Year 4: 49/79 (62%)	Four-year intervention including an educational group program (lectures, workshops, trainings) and a web-based mental health screening survey	Self-assessment and survey	13	Educate, destigmatize, identify, refer, and treat individuals with depression and increased suicide risk
Seritan et al,[29] 2015	Single site			Multipronged mental health/wellness program offering prevention, support, and enhanced clinical services (ie, hiring a psychiatrist to offer medication management) through development of a new Office of Student Wellness with evening hours and strict confidentiality	Survey	11.5	Presentation of a model for effective preventive student wellness
Moutier et al,[31] 2012	Single site	N = 498 (132 screened)		Two-pronged intervention consisting of grand rounds lecture on mental health and a web-based mental health screening survey	Self-assessment	11.5	Develop a mental health program to address physician and medical student depression and suicide
Mind-Body Skills Education/Training Programs							
Erogul et al,[32] 2014	Single site	N = 58	26/58 (45.6%)	Eight-week mindfulness-based stress reduction intervention for first-year medical students	Self-assessment	12	Assess whether an abridged mindfulness-based stress reduction intervention can improve wellness
Holtzworth-Munroe et al,[34] 1985	Single site	N = 40		Six weekly meetings focused on teaching skills to reduce stress levels (progressive muscle relaxation, skills to help recognize and change maladaptive thoughts, and meditation techniques)	Self-assessment	10	Help students acquire and develop skills to cope with stress
Kraemer et al,[36] 2016	Single site	N = 52	62.7%	Mind-body program consisting of 11 weekly skill training groups focusing on mind-body skills (biofeedback, guided imagery, relaxation, breathing exercises, autogenic training, and meditation)	Self-assessment and survey	12	Describe changes in distress tolerance after completing a mind-body skills training group
Rosenzweig et al,[39] 2003	Single site	N = 302		Mindfulness-based stress reduction including 10 weekly 90-minute sessions teaching mindfulness meditation practices and daily, independent meditation	Self-assessment	11	Examine the effectiveness of mindfulness-based stress reduction seminar

(continued)

Table 2. Methods of Included Studies (continued)

Source	No. of Sites	Overall Sample Size	Women, No. (%)	Intervention	Evaluation	MERSQI Score	Aim
Mind-Body Skills Education/Training Programs (continued)							
Finkelstein et al,[40] 2007	Single site	N = 72	Time 1: intervention, 17/26 (77.3%); control, 22/46 (61.1%) Time 2: intervention, 17/26 (77.3%); control, 16/25 (80%) Time 3: intervention, 15/23 (75%); control, 20/40 (62.5%)	"Mind-Body Medicine: An Experiential Elective," including 10 sessions of didactic and small-group instruction	Self-assessment	11	Assess the effectiveness of a stress reduction elective on second-year medical students
Greeson et al,[44] 2015	Single site	N = 44	29/44 (65)	Four weekly 1.5-hour small-group sessions and home practice of mind-body skills in addition to monitoring a weekly self-care goal	Self-assessment and semistructured interview	11.5	Evaluate the feasibility, acceptability, and effectiveness of a stress-management and self-case workshop
Bond et al,[46] 2013	Single site	N = 27		Eleven-week embodied health course combining yoga meditation and neuroscience didactics	Self-assessment and semistructured interview	11.5	Evaluate the psychological effects of an 11-week mind-body elective course
Curriculum Structure							
Reed et al,[20] 2011	Multisite	N = 2056 (1192 responded)	550/1192 (47%)	Multisite survey of 2 different grading scales: pass/fail and ≥3-interval (eg, honors/pass/fail, honors/high pass/pass/marginal pass/fail)	Self-assessment	12.5	Examine the relationship among curriculum structure, grading scales, and student well-being
Camp et al,[51] 1994	Single site	N = 275	93/275 (33.8%)	Student-directed, project-based learning approach featuring small-group, problem-based sessions in which both basic and clinical science learning issues are generated; lecture-based learning is an instructor-directed, didactic approach	Self-assessment	12	Assess changes in depression among medical students enrolled in a lecture-based vs problem-based learning program
Kornitzer et al,[53] 2005	Single site	N = 92		Six-week prematriculation enrichment program targeting educationally disadvantaged students (didactic sessions and laboratory component)	Survey	7	Determine whether educationally disadvantaged students participating in a summer enrichment program were reported to have had an easier time adjusting to medical school
Multicomponent Program Reform							
Drolet and Rodgers,[54] 2010	Single site	N = 116		Vanderbilt medical student wellness program to promote student health and well-being through changes, including faculty mentoring (advisory college program and Vanderbilt Medical Student Careers in Medicine), curriculum (VMS Live Program), and student well-being (student wellness committee)	Survey	6.5	Evaluate a multicomponent wellness program
Fleming et al,[55] 2013	Single site	N = 245		Initiatives, activities, and resources including (1) advisory college program for student well-being and career mentoring/advising with an additional aim of establishing relationships between students and faculty serving as both teachers and role models; (2) student-led student wellness committee focused on peer mentoring, social community, and mind/body wellness programming; (3) Vanderbilt Medical Student Careers in Medicine for career exploration, advising, and planning as well as residency application preparation; (4) VMS Live Program focused on personal development of physicians in training; and (5) 4-year College Colloquium Course focused on medical humanities and formally addressing professionalism, ethics, and leadership skills	Survey	6	Reflect on and describe learning community system and effect on student satisfaction

(continued)

Table 2. Methods of Included Studies (continued)

Source	No. of Sites	Overall Sample Size	Women, No. (%)	Intervention	Evaluation	MERSQI Score	Aim
Multicomponent Program Reform (continued)							
Real et al,[56] 2015	Single site	N = 450	55%	Vanderbilt wellness program including a faculty-led mentoring system, annual retreat series, and student-led programming committee, all organized around a college system that divides students into 1 of 4 colleges	Self-assessment and survey	10.5	Association of a wellness initiative on distress
Slavin et al,[58] 2014	Single site	N = 875-890		Phase 1: pass/fail replaced 4-interval (honors/near honors/pass/fail) grading system, reduction in contact hours by approximately 10%, longitudinal electives, established 5 learning communities of medical students and faculty with common interests beyond the classroom Phase 2: refined pass/fail grading system by eliminating norm-referenced performance data, resilience/mindfulness program spanning 6 hours Phase 3: modified human anatomy course to occur later in the first year and to have examinations with mean scores consistent with other courses	Self-assessment and survey	12	Discuss the utility and relevance of curricular changes and association with student mental health
Strayhorn,[59] 1989	Single site	n = 478 responders (original sample size not provided)		Major curriculum revision, including seminar- and small group-based learning, analytical and problem-solving skill building, increased free time for student learning, formal instruction in social and behavioral sciences, increased mentoring, and development of a new student/faculty/curriculum evaluation system	Self-assessment	12	Assess student well-being and perceptions on medical school learning environment after curriculum change
Miscellaneous Wellness Programs							
Whitehouse et al,[62] 1996	Single site	N = 35	60%	Daily practice of self-hypnosis and diary records of sleep, mood, physical symptoms, and frequency of relaxation practice	Self-assessment	12	Determine the effectiveness of a self-hypnosis/relaxation intervention to relieve symptoms of psychological distress and immune system reactivity to examination stress
Goetzel et al,[63] 1984	Single site	N = 26	45%	Human Dimensions Program: biweekly self-help support group	Survey	9	Assess the quality of support groups at Albert Einstein College of Medicine
Lee and Graham,[64] 2001	Single site	N = 66	40/60 (66%)	Six-week wellness elective consisting of one-hour lectures by physician presenters, discussions, and writing exercises	Survey	7	Explore students' perceptions of medical school stress and assess their perspective on the wellness elective
Kushner et al,[65] 2011	Single site	N = 343	171/343 (49.8%)	Behavior change plan in which students attempt to change one of their own health behaviors, including a mental/emotional health personal goal	Self-assessment	8	Teach medical students the principles and practice of behavior change using a behavior change plan
Group-Based Faculty Advisor/Mentor Programs							
Sastre et al,[66] 2010	Single site	N = 318		Faculty advisory program, advisory college program, consisting of 4 advisory colleges each co-led by 2 faculty members nominated and competitively selected by a student committee; advisory college program faculty focus on advising by promoting wellness and providing career counseling	Survey	9.5	Determine if advisory college program is more effective than one-on-one mentoring
Coates et al,[67] 2008	Single site	N = 100		Group-based mentoring program (the "College Program") exclusively for fourth-year medical students, which divided students into academic interest-based groups led by a faculty chair and included a team of both faculty and student mentors/advisors/role models; the College Program provided mentoring, career advising, and curricular support	Survey	8	Change in fourth-year curriculum to include more mentors
Ficklin et al,[68] 1983	Single site	N = 151		Small group-based faculty advisor program exclusively advising first-year medical students with goals of increased student/faculty communication, informal student/faculty activities, increased student-to-student communication and support, and decreased anonymity; advisory groups were maintained as sections of larger courses	Survey	6	Provide advice and support in areas of documented stress

Abbreviation: MERSQI, Medical Education Research Study Quality Instrument.

Two studies evaluated programs consisting of education and a web-based mental health screening survey to facilitate students' use of mental health services. Downs et al[26] (n = 1008; MERSQI, 13.0) described a program that was associated with an increase in mental health service utilization and a non-statistically significant decrease in assessed suicide risk during the 4 years, perhaps due to low screening rates (34%). Moutier et al[31] (n = 498; MERSQI, 11.5) reported that that 11% of medical students exposed to another educational program were referred to a mental health care professional, although no comparison was provided and the screening rate was also low (27%).

Mind-Body Skills Education/Training Programs (Mean MERSQI Score, 11.3)

Two RCTs evaluated mind-body programs. Erogul et al[32] (n = 58; MERSQI, 12.0) found that students randomized to attend a mindfulness program reported a significant reduction in stress after intervention (Perceived Stress Scale[16]; mean change, 3.63; 95% CI, 0.37-6.89; $P = .03$) but not at 6-month follow-up (mean change, 2.91; 95% CI, −0.37 to 6.19; $P = .08$). However, students in the mind-body program reported a significant increase in self-compassion that persisted at 6-month follow-up (Self-Compassion Scale[33] [range, 0-5; higher score indicates more self-compassion]; mean change, 0.56; 95% CI, 0.25-0.87; $P = .001$).[32] In the study by Holtzworth-Munroe et al[34] (n = 40; MERSQI, 10.0), students randomized to a mind-body program were reported to have significantly more awareness of tension ($F_{5,18} = 37.16$; $P < .001$), better ability to deal with school stress ($F_{5,18} = 5.05$; $P < .04$), and less test anxiety at 10-week follow-up ($F_{1,22} = 10.42$; $P < .005$).

Three studies evaluated mind-body programs using a pretest/posttest design with nonrandomized control groups. Kraemer et al[36] (n = 52; MERSQI, 12.0) found that students undergoing mind-body skills training reported significantly improved distress tolerance (Distress Tolerance Scale G[37] [range, 1-5; higher scores indicate higher levels of distress tolerance]; mean change, 0.53; 95% CI, 0.92-0.14; $P = .01$); no difference was found for the control group. Rosenzweig et al[39] (n = 302; MERSQI, 11.0) described a mindfulness-based stress reduction program associated with significant improvements in total mood disturbance (Profile of Mood States[18] [range, 0-200; higher scores indicate higher mood disturbance]; intervention group pretest mean, 38.7 [SD, 33.3] vs posttest mean, 31.8 [SD, 33.89]; $P = .05$; control group pretest mean, 28.0 [SD, 31.2] vs posttest mean, 38.6 [SD, 32.8]; $P < .001$; interaction $P < .001$). Finkelstein et al[40] (n = 72; MERSQI, 11.0) found a significant group × time interaction association with improved anxiety ($F_{1,2} = 3.95$; $P < .05$).

Two studies evaluating medical student mind-body programs with a pretest/posttest design without a control group also reported associations with significant improvements in well-being. Greeson et al[44] (n = 44; MERSQI, 11.5) reported improved stress (Perceived Stress Scale[16]; pretest mean, 29.73 [SD, 9.61]; posttest mean, 20.25 [SD, 9.03]; $P < .001$; $d = 1.38$) and mindfulness (pretest mean, 29.24 [SD, 5.54]; posttest mean, 33.88 [SD, 6.13]; $P < .001$; $d = 0.92$). Bond et al[46] (n = 27; MERSQI, 11.5) reported improved self-regulation (Self-Regulation Questionnaire[49] [range, 1-5; higher score indicates more self-regulation]; mean change, 0.13 [SD, 0.20]; $P = .003$; $d = −0.41$) and self-compassion (Self-Compassion Scale[33]; mean change, 0.28 [SD, 0.61]; $P = .04$; $d = −0.55$).

Curriculum Structure (Mean MERSQI Score, 9.5)

Elements of curriculum structure targeted by studies identified in this review were varied. Reed et al[20] (n = 2056; MERSQI, 12.5) compared elements of curriculum structure at different medical schools. Students who reported more clinical contact hours were significantly less likely to report serious thoughts of dropping out (odds ratio, 0.96; 95% CI, 0.93-1.00; $P = .03$). Although the number of tests was not associated with any difference in well-being, spending more time taking tests was associated with significantly higher perceived stress (β = 0.29; 95% CI, 0.10-0.84; $P = .003$) and lower mental quality of life (β = 2.79; 95% CI, 4.09-1.50; $P < .001$).[20]

Camp et al[51] (n = 275; MERSQI, 12.0) found that students in a new problem-based learning curriculum, compared with a lecture-based one, had similar reports of depression with covariate adjustment. A prematriculation summer enrichment program for medicine and nonscience undergraduate majors from underrepresented groups described reports of gaining confidence, making friends, and perceiving an easier transition to medical school (n = 92; MERSQI, 7.0).[53]

Multicomponent Program Reform (Mean MERSQI Score, 9.4)

Vanderbilt University restructured its medical school learning environment, which, after multiple iterations, ultimately took the form of "learning communities" or colleges within the school. These intentionally developed groups of faculty and students work together longitudinally, with functions that include mentoring, wellness programming (including mind-body skill training, career advising, and personal and professional development), and formal medical humanities coursework. Several different studies evaluated the multicomponent program at various stages of its development and implementation. Drolet and Rodgers[54] (n = 116; MERSQI, 6.5) evaluated the faculty advisor/mentor program after the addition of several components and found that 95% of students reported a positive experience with the wellness program. Fleming et al[55] (n = 245; MERSQI, 6.0) assessed the association of the most recent program iteration, including colleges, and found that more than 91% of students reported that colleges contributed at least somewhat meaningfully to their medical school experience. Real et al[56] (n = 450; MERSQI, 10.5) reported that students credited the program in general (and more specifically, faculty mentors), the student-led programming committee, and annual retreats with lowering reported rates of burnout.

The St Louis University School of Medicine also undertook multicomponent program reform that was introduced in phases to preclinical students: (1 component) pass/fail grad-

ing for preclinical courses, reduced preclinical contact hours, extended electives, and learning communities; (2 components) addition of mind-body skills training; and (3 components) addition of anatomy course reform. As reported in a study by Slavin et al[58] (n = 890; MERSQI, 12.0), phase 1 was significantly associated with improved depression, stress, and cohesion by the end of the second year of UME. The 2-component phase was associated with significantly improved anxiety, stress, and cohesion by the end of the first year of UME; depression was reported to be improved by the end of the second year of UME.[58] The 3-component phase was associated with statistically significant improvements in all measures of well-being by the end of the first year, persisting through the second year of UME.[58]

Strayhorn[59] (n = 478; MERSQI, 12.0) compared one school's curriculum changes with a comparison school's curriculum and found significant time × school interactions that favored the changes with regard to reported stressors (F_{1467} = 6.41; P = .01), mental well-being (F_{1460} = 9.32; P = .002), and social well-being (F_{1466} = 5.37; P = .02).

Miscellaneous Wellness Programs (Mean MERSQI Score, 9.0)

In a self-hypnosis training RCT, Whitehouse et al[62] (n = 35; MERSQI, 12.0) reported significant improvements in anxiety (Brief Symptom Inventory[70] [range, 20-80; higher scores indicate higher anxiety]; orientation mean, 59.23 [SD, 9.41]; late semester mean, 56.31 [SD, 9.29]; examination stressor mean, 58.59 [SD, 10.43]; recovery mean, 52.64 [SD, 9.66]; interaction $F_{3,96}$ = 2.96; P < .05). A cross-sectional survey (n = 26; MERSQI, 9.0) about access to student support groups reported that a majority of students felt less lonely and unique with their problems.[63] An evaluation of a wellness elective (n = 66; MERSQI, 7.0) reported that only a minority of students agreed or strongly agreed that it altered their report of the importance of well-being or permission for self-care or provided coping strategies (no significance values reported).[64] Kushner et al[65] (n = 343; MERSQI, 8.0) evaluated a wellness course that included a section on behavior change plans; of the 9 students who set mental/emotional health goals, 6 reported achieving their goals (no significance values reported).

Group-Based Faculty Advisor/Mentor Programs (Mean MERSQI Score, 8.2)

Three studies evaluated small group-based faculty advisor/mentor programs that were formally integrated into the academic curriculum. Sastre et al[66] (n = 318; MERSQI, 9.5) evaluated a program in which competitively selected faculty had protected time for advising groups of students. Compared with students with traditional one-on-one volunteer faculty advisors, intervention students were significantly more likely to report that they agreed or strongly agreed that they were satisfied with how faculty advisors promoted wellness (72% vs 27%; P < .001) and that they agreed or strongly agreed that they would feel comfortable discussing their personal stress (62% vs 24%; P < .001) or mental health (51% vs 27%; P < .001) with their advisor.[66] Coates et al[67] (n = 100; MERSQI, 8.0) reported that fourth-year medical students involved in an intervention said they felt connected with faculty and with classmates (no significance values reported).

The evaluation of a program exclusively for first-year students by Ficklin et al[68] (n = 151; MERSQI, 7.0) reported that students stated they were better acquainted with their peers, became close with some classmates, and were helped with anxiety related to starting medical school as a result of the program, but there was no comparison group and no significance values were reported.

Discussion

This systematic review identified hundreds of articles on the UME learning environment, but only a small subset contained empirically evaluated interventions. No studies included in this systematic review met a quality cutoff of 14.0.[12] Improving the content and context of the delivery of UME will benefit from studies with rigorous design, objective data collection, and appropriate intervention comparators, as used in other scientific and educational fields. Despite these limitations in the evidence, there are a number of key findings from this review that may be relevant for US medical schools.

First, implementation of a preclinical pass/fail grading system should be considered. All of the studies reviewed here show that a preclinical pass/fail grading system improves medical student well-being. The duration of benefit can be finite, with any positive effect perhaps more likely to persist in the context of good medical school class cohesion.[15] It is also important to consider educational repercussions of changing grading systems to ensure that rigorous mastery of educational material and professional preparedness is balanced with student well-being. Two studies in this review addressed this concern by showing that pass/fail grading systems can be associated with improved well-being without any significant change in course test scores, including United States Medical Licensing Examination Step 1 and 2 scores and subsequent postresidency specialty board certification scores.[13,15] This is consistent with other literature exclusively focused on academic outcomes of pass/fail grading.[71-73] According to the 2014-2015 Liaison Committee on Medical Education Annual Medical School Questionnaire, 87 of the 144 participating schools used pass/fail grading systems for at least some portion of the preclinical courses.[74]

Second, the accessibility and quality of mental health programs for medical students, as well as any stigma associated with these programs, should be taken into account.[75] Students with mental health problems may be undertreated; in one study, fewer than half of the students who reported having contemplated suicide during medical school received counseling for their depression.[76] Addressing mental health conditions with a formal program that includes treatment services is essential, and a multipronged program aimed at improving awareness, reducing stigma, and improving access to mental health care professionals seems to be an efficacious approach and is associated with lower depression and suicidal ideation rates.[24]

There are specific components of mental health programs that can be critical to improving students' well-being. Barriers to medical students' mental health treatment reported elsewhere include concern about stigma and lack of confidentiality, such as fear of documentation in the academic record and evaluators' knowledge of student mental health conditions with subsequent career implications.[4,47,48,76,77] Medical students reported preferring help from a mental health specialist, family, or friends rather than medical school personnel[47] and reported preferring accessing mental health services through a location other than the office of student affairs.[48] In other studies, students have reported concerns about time, convenience of office hours, location, and financial costs.[4,47,48,77] Although these are small studies of implementation issues, they are worth considering for the introduction of student mental health programs.

Third, introducing wellness programs that teach mind-body-based stress-reduction skills should be considered. The majority of studies in this category, including 2 RCTs, indicate that such programs are associated with reduced stress, anxiety, better mood, and higher distress tolerance. This association was found even when skills were taught in condensed workshops lasting only 4 weeks,[44] which is an important factor because programs must balance benefit derived from wellness programs with time investment.

Fourth, implementation of formal faculty advisor/mentor programs based in small groups and linked with curricular content should be examined. All 3 studies in this review that evaluated faulty advisor/mentor programs were highly regarded by students as a method of promoting wellness, although only 1 study tested for statistical significance.[66] However, it is important that mentors do not grade students to keep their role as advisors separate from assessment to foster open communication.[68] A small group-based mentoring model—rather than a one-on-one mentoring system—reduces the number of required faculty mentoring positions, allowing medical schools to have competitive selection for a subset of excellent faculty and may even enable financial support for this function.[68] Outstanding faculty mentors are critical to the success of any mentoring program because they both relay explicit academic knowledge and exemplify implicit knowledge on professionalism, ethics, and values—the "hidden curriculum."[78]

Fifth, the curriculum should be structured to balance clinical and nonclinical learning environments. Medical students report less burnout and stress when clinical time is increased.[20] Many recent changes to curriculum have decreased clinical learning exposures, so consideration of where this movement can be reversed will be useful.

Sixth, comprehensive reform of the learning environment that incorporates many of these interventions is likely required. A detailed evaluation of the sequential implementation phases indicates that there may have been synergies among program components that were associated with improvements in medical student well-being.[58]

This study has a number of limitations. First, the primary studies varied widely in design, intervention content, and outcomes collected, precluding meta-analytic pooling. Second, the scope of the review was restricted to studies evaluating the quantitative effect of learning environment interventions on medical student well-being, although there are other aspects of the learning environment that deserve attention in a comprehensive redesign of the learning offered to medical students. Third, qualitative research was not included in this systematic review. Fourth, there are concerns about the ethics of randomization of education research.[79,80] Historically, research conducted in established educational settings and involving normal educational practices were considered exempt from institutional review board oversight.[81] However, issues of coercion and lack of informed consent about randomization of medical students when conducting learning environment interventions tests have recently been raised.[82,83] These issues are complex and include whether there is a research component to the investigation of the education practice; whether there is an intent to publish; whether empirically established practices already exist; and whether the investigator has a hierarchical relationship to the participants, such as that held by a clerkship director or faculty advisor. Guidance is provided elsewhere for future UME educators to decide when and under what circumstances randomization is ethical and practical for learning environment interventions.[82,83]

Conclusions

In this systematic review, limited evidence suggested that some specific learning environment interventions were associated with improved medical student emotional well-being. However, the overall quality of these studies was low, highlighting the need for high-quality medical education research.

ARTICLE INFORMATION

Author Contributions: Drs Davidson and Wasson had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.
Concept and design: Wasson, Cusmano, Meli, Louh, Young, Davidson.
Acquisition, analysis, or interpretation of data: Wasson, Cusmano, Meli, Louh, Falzon, Hampsey, Shaffer.
Drafting of the manuscript: Wasson, Cusmano, Meli, Louh, Falzon, Hampsey, Davidson.
Critical revision of the manuscript for important intellectual content: Wasson, Meli, Louh, Young, Shaffer, Davidson.
Obtained funding: Davidson.
Administrative, technical, or material support: Cusmano, Meli, Falzon, Hampsey, Davidson.
Supervision: Wasson.

Conflict of Interest Disclosures: All authors have completed and submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest and none were reported.

Funding/Support: Dr Davidson was supported by research grant K24 HL084034 from the National Heart, Lung, and Blood Institute. Dr Wasson was supported by research grant K08 HS024598 from the Agency for Healthcare Research and Quality. Ms Falzon was supported by research grant U24AG052175 and contract S15-0142, both from the National Institutes of Health, and by Columbia University Medical Center.

Role of the Funder/Sponsor: The funders had no role in the design and conduct of the study; collection, management, analysis, and

interpretation of the data; preparation, review, or approval of the manuscript; or decision to submit the manuscript for publication.

REFERENCES

1. Dezee KJ, Artino AR, Elnicki DM, Hemmer PA, Durning SJ. Medical education in the United States of America. *Med Teach*. 2012;34(7):521-525.

2. Brazeau CM, Shanafelt T, Durning SJ, et al. Distress among matriculating medical students relative to the general population. *Acad Med*. 2014;89(11):1520-1525.

3. Dyrbye LN, Massie FS Jr, Eacker A, et al. Relationship between burnout and professional conduct and attitudes among US medical students. *JAMA*. 2010;304(11):1173-1180.

4. Schwenk TL, Davis L, Wimsatt LA. Depression, stigma, and suicidal ideation in medical students. *JAMA*. 2010;304(11):1181-1190.

5. Dyrbye LN, Thomas MR, Massie FS, et al. Burnout and suicidal ideation among US medical students. *Ann Intern Med*. 2008;149(5):334-341.

6. Association of American Medical Colleges. *Educating Doctors to Provide High Quality Medical Care: A Vision for Medical Education in the United States*. Washington, DC: Association of American Medical Colleges; 2004.

7. Gelfand J. Grey literature: plausible applications for distance education practices. In: *GL'99 Conference Proceedings: Fourth International Conference on Grey Literature: "New Frontiers in Grey Literature."* Amsterdam, the Netherlands: GreyNet; January 2000:1-11. GL Conference Series No. 4.

8. Stacy R, Spencer J. Assessing the evidence in qualitative medical education research. *Med Educ*. 2000;34(7):498-500.

9. Côté L, Turgeon J. Appraising qualitative research articles in medicine and medical education. *Med Teach*. 2005;27(1):71-75.

10. Reed DA, Cook DA, Beckman TJ, Levine RB, Kern DE, Wright SM. Association between funding and quality of published medical education research. *JAMA*. 2007;298(9):1002-1009.

11. Cook DA, Reed DA. Appraising the quality of medical education research methods: the Medical Education Research Study Quality Instrument and the Newcastle-Ottawa Scale–Education. *Acad Med*. 2015;90(8):1067-1076.

12. Lin H, Lin E, Auditore S, Fanning J. A narrative review of high-quality literature on the effects of resident duty hours reforms. *Acad Med*. 2016;91(1):140-150.

13. Bloodgood RA, Short JG, Jackson JM, Martindale JR. A change to pass/fail grading in the first two years at one medical school results in improved psychological well-being. *Acad Med*. 2009;84(5):655-662.

14. DuPuy HJ. Self-representations of general psychological well-being of American adults. Presented at: American Public Health Association Meeting; Los Angeles, CA; October 15-19, 1978.

15. Rohe DE, Barrier PA, Clark MM, Cook DA, Vickers KS, Decker PA. The benefits of pass-fail grading on stress, mood, and group cohesion in medical students. *Mayo Clin Proc*. 2006;81(11):1443-1448.

16. Cohen S, Kamarck T, Mermelstein R. A global measure of perceived stress. *J Health Soc Behav*. 1983;24(4):385-396.

17. Cohen S. Perceived stress in a probability sample of the United States. In: Oskamp SSS, ed. *The Social Psychology of Health*. Thousand Oaks, CA: Sage Publications; 1988:31-67.

18. McNair DM, Lorr M, Droppleman LF. *POMS Manual Profile of Mood States*. San Diego, CA: EdiTS/Educational and Industrial Testing Service; 1992.

19. Bollen KA, Hoyle RH. Perceived cohesion: a conceptual and empirical examination. *Soc Forces*. 1990;69(2):479-504.

20. Reed DA, Shanafelt TD, Satele DW, et al. Relationship of pass/fail grading and curriculum structure with well-being among preclinical medical students: a multi-institutional study. *Acad Med*. 2011;86(11):1367-1373.

21. Maslach C, Jackson SE. The measurement of experienced burnout. *J Organ Behav*. 1981;2(2):99-113.

22. Stewart AL, Ware JE, eds. *Measuring Functioning and Well-being: the Medical Outcome Study Approach*. Durham, NC: Duke University Press; 1992.

23. Tarlov AR, Ware JE Jr, Greenfield S, Nelson EC, Perrin E, Zubkoff M. The Medical Outcomes Study: an application of methods for monitoring the results of medical care. *JAMA*. 1989;262(7):925-930.

24. Thompson D, Goebert D, Takeshita J. A program for reducing depressive symptoms and suicidal ideation in medical students. *Acad Med*. 2010;85(10):1635-1639.

25. Radloff LS. The CES-D Scale: a self-report depression scale for research in the general population. *Appl Psychol Meas*. 1977;1(3):385-401.

26. Downs N, Feng W, Kirby B, et al. Listening to depression and suicide risk in medical students: the Healer Education Assessment and Referral (HEAR) Program. *Acad Psychiatry*. 2014;38(5):547-553.

27. Spitzer RL, Kroenke K, Williams JB. Validation and utility of a self-report version of PRIME-MD: the PHQ primary care study. *JAMA*. 1999;282(18):1737-1744.

28. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. *J Gen Intern Med*. 2001;16(9):606-613.

29. Seritan AL, Rai G, Servis M, Pomeroy C. The office of student wellness: innovating to improve student mental health. *Acad Psychiatry*. 2015;39(1):80-84.

30. American Association of Medical Colleges. Graduation Questionnaire (GQ). https://www.aamc.org/data/gq/. Accessed October 21, 2016.

31. Moutier C, Norcross W, Jong P, et al. The suicide prevention and depression awareness program at the University of California, San Diego School of Medicine. *Acad Med*. 2012;87(3):320-326.

32. Erogul M, Singer G, McIntyre T, Stefanov DG. Abridged mindfulness intervention to support wellness in first-year medical students. *Teach Learn Med*. 2014;26(4):350-356.

33. Neff KD. The development and validation of a scale to measure self-compassion. *Self Identity*. 2003;2(3):223-250.

34. Holtzworth-Munroe A, Munroe MS, Smith RE. Effects of a stress-management training program on first- and second-year medical students. *J Med Educ*. 1985;60(5):417-419.

35. Spielberger CD, Gorsuch RL, Lushene RE, Vagg PR, Jacobs GA. *Manual for the State-Trait Anxiety Inventory*. Palo Alto, CA: Consulting Psychologists Press; 1983.

36. Kraemer KM, Luberto CM, O'Bryan EM, Mysinger E, Cotton S. Mind-body skills training to improve distress tolerance in medical students: a pilot study. *Teach Learn Med*. 2016;28(2):219-228.

37. Simons JS, Gaher RM. The Distress Tolerance Scale: development and validation of a self-report measure. *Motiv Emot*. 2005;29(2):83-102.

38. Thompson ER. Development and validation of an internationally reliable short-form of the Positive and Negative Affect Schedule (PANAS). *J Cross Cult Psychol*. 2007;38(2):227-242.

39. Rosenzweig S, Reibel DK, Greeson JM, Brainard GC, Hojat M. Mindfulness-based stress reduction lowers psychological distress in medical students. *Teach Learn Med*. 2003;15(2):88-92.

40. Finkelstein C, Brownstein A, Scott C, Lan YL. Anxiety and stress reduction in medical education: an intervention. *Med Educ*. 2007;41(3):258-264.

41. Derogatis LR. *SCL-90-R Administration, Scoring and Procedures Manual, Vol II*. Towson, MD: Clinical Psychometric Research; 1983.

42. Vitaliano PP, Russo J, Carr JE, Heerwagen JH. Medical school pressures and their relationship to anxiety. *J Nerv Ment Dis*. 1984;172(12):730-736.

43. Whooley MA, Avins AL, Miranda J, Browner WS. Case-finding instruments for depression: two questions are as good as many. *J Gen Intern Med*. 1997;12(7):439-445.

44. Greeson JM, Toohey MJ, Pearce MJ. An adapted, four-week mind-body skills group for medical students: reducing stress, increasing mindfulness, and enhancing self-care. *Explore (NY)*. 2015;11(3):186-192.

45. Feldman G, Hayes A, Kumar S, et al. Mindfulness and emotion regulation: the development and initial validation of the Cognitive and Affective Mindfulness Scale-Revised (CAMS-R). *J Psychopathol Behav Assess*. 2006;29(3):177.

46. Bond AR, Mason HF, Lemaster CM, et al. Embodied health: the effects of a mind-body course for medical students. *Med Educ Online*. 2013;18:1-8.

47. Dyrbye LN, Eacker A, Durning SJ, et al. The impact of stigma and personal experiences on the help-seeking behaviors of medical students with burnout. *Acad Med*. 2015;90(7):961-969.

48. Plaut SM, Maxwell SA, Seng L, O'Brien JJ, Fairclough GF Jr. Mental health services for medical students: perceptions of students, student affairs deans, and mental health providers. *Acad Med*. 1993;68(5):360-365.

49. University of New Mexico Center on Alcoholism, Substance Abuse, and Addictions. The Self-Regulation Questionnaire. http://casaa.unm.edu/inst/SelfRegulation%20Questionnaire%20(SRQ).pdf. Accessed October 26, 2016.

50. Hojat M, Mangione S, Nasca TJ, et al. The Jefferson Scale of Physician Empathy: development and preliminary psychometric data. *Educ Psychol Meas*. 2001;61(2):349-365.

51. Camp DL, Hollingsworth MA, Zaccaro DJ, Cariaga-Lo LD, Richards BF. Does a problem-based learning curriculum affect depression in medical students? *Acad Med*. 1994;69(10)(suppl):S25-S27.

52. Zung WW. A self-rating depression scale. *Arch Gen Psychiatry*. 1965;12:63-70.

53. Kornitzer B, Ronan E, Rifkin MR. Improving the adjustment of educationally disadvantaged students to medical school: the Summer Enrichment Program. *Mt Sinai J Med*. 2005;72(5):317-321.

54. Drolet BC, Rodgers S. A comprehensive medical student wellness program—design and implementation at Vanderbilt School of Medicine. *Acad Med*. 2010;85(1):103-110.

55. Fleming A, Cutrer W, Moutsios S, et al. Building learning communities: evolution of the colleges at Vanderbilt University School of Medicine. *Acad Med*. 2013;88(9):1246-1251.

56. Real FJ, Zackoff MW, Davidson MA, Yakes EA. Medical student distress and the impact of a school-sponsored wellness initiative. *Med Sci Educ*. 2015;25(4):397-406.

57. Spitzer RL, Williams JB, Kroenke K, et al. Utility of a new procedure for diagnosing mental disorders in primary care: the PRIME-MD 1000 study. *JAMA*. 1994;272(22):1749-1756.

58. Slavin SJ, Schindler DL, Chibnall JT. Medical student mental health 3.0: improving student wellness through curricular changes. *Acad Med*. 2014;89(4):573-577.

59. Strayhorn G. Effect of a major curriculum revision on students' perceptions of well-being. *Acad Med*. 1989;64(1):25-29.

60. Rothman AI, Ayoade F. The development of a learning environment: a questionnaire for use in curriculum evaluation. *J Med Educ*. 1970;45(10):754-759.

61. Brook RH, Ware JE, Davies AR, et al. *Conceptualization and Measurement of Health for Adults in the Health Insurance Study*. Vol VIII. Santa Monica, CA: RAND Corp; 1979.

62. Whitehouse WG, Dinges DF, Orne EC, et al. Psychosocial and immune effects of self-hypnosis training for stress management throughout the first semester of medical school. *Psychosom Med*. 1996;58(3):249-263.

63. Goetzel RZ, Croen LG, Shelov S, Boufford JI, Levin G. Evaluating self-help support groups for medical students. *J Med Educ*. 1984;59(4):331-340.

64. Lee J, Graham AV. Students' perception of medical school stress and their evaluation of a wellness elective. *Med Educ*. 2001;35(7):652-659.

65. Kushner RF, Kessler S, McGaghie WC. Using behavior change plans to improve medical student self-care. *Acad Med*. 2011;86(7):901-906.

66. Sastre EA, Burke EE, Silverstein E, et al. Improvements in medical school wellness and career counseling: a comparison of one-on-one advising to an advisory college program. *Med Teach*. 2010;32(10):e429-e435.

67. Coates WC, Crooks K, Slavin SJ, Guiton G, Wilkerson L. Medical school curricular reform: fourth-year colleges improve access to career mentoring and overall satisfaction. *Acad Med*. 2008;83(8):754-760.

68. Ficklin FL, Hazelwood JD, Carter JE, Shellhamer RH. Evaluation of a small-group support program for first-year medical students. *J Med Educ*. 1983;58(10):817-819.

69. Egger M, Smith GD, Phillips AN. Meta-analysis: principles and procedures. *BMJ*. 1997;315(7121):1533-1537.

70. Derogatis LR, Melisaratos N. The Brief Symptom Inventory: an introductory report. *Psychol Med*. 1983;13(3):595-605.

71. Jones KG, Pedersen RL, Carmichael SW, Pawlina W. Effects of pass/fail grading system on academic performance of first year medical students in gross anatomy course. *FASEB J*. 2003;17:A385.

72. Vosti KL, Jacobs CD. Outcome measurement in postgraduate year one of graduates from a medical school with a pass/fail grading system. *Acad Med*. 1999;74(5):547-549.

73. Robins LS, Fantone JC, Oh MS, Alexander GL, Shlafer M, Davis WK. The effect of pass/fail grading and weekly quizzes on first-year students' performances and satisfaction. *Acad Med*. 1995;70(4):327-329.

74. Association of American Medical Colleges. Number of Medical Schools Using Selected Grading Systems in Pre-Clerkship Courses (Excluding Physical Diagnosis/Clinical Skills). https://www.aamc.org/initiatives/cir/406418/11.html. Accessed October 16, 2016.

75. Dyrbye LN, Thomas MR, Shanafelt TD. Systematic review of depression, anxiety, and other indicators of psychological distress among US and Canadian medical students. *Acad Med*. 2006;81(4):354-373.

76. Givens JL, Tjia J. Depressed medical students' use of mental health services and barriers to use. *Acad Med*. 2002;77(9):918-921.

77. Gentile JP, Roman B. Medical student mental health services: psychiatrists treating medical students. *Psychiatry (Edgmont)*. 2009;6(5):38-45.

78. Rose GL, Rukstalis MR, Schuckit MA. Informal mentoring between faculty and medical students. *Acad Med*. 2005;80(4):344-348.

79. Mosteller F, Boruch RF. *Evidence Matters: Randomized Trials in Education Research*. Washington, DC: Brookings Institution Press; 2002.

80. Green JL, Camilli G, Elmore PB. *Handbook of Complementary Methods in Education Research*. 3rd ed. Washington, DC: American Educational Research Association; 2006.

81. US Department of Health and Human Services. Federal Policy for the Protection of Human Subjects. In: *Federal Register*. http://archive.hhs.gov/ohrp/documents/19881110.pdf. Accessed October 16, 2016.

82. Johansson AC, Durning SJ, Gruppen LD, Olson ME, Schwartzstein RM, Higgins PA. Perspective: medical education research and the institutional review board: reexamining the process. *Acad Med*. 2011;86(7):809-817.

83. Keune JD, Brunsvold ME, Hohmann E, Korndorffer JR Jr, Weinstein DF, Smink DS. The ethics of conducting graduate medical education research on residents. *Acad Med*. 2013;88(4):449-453.

84. Russell DW. UCLA Loneliness Scale (Version 3): reliability, validity, and factor structure. *J Pers Assess*. 1996;66(1):20-40.

85. Moos RH. The assessment of the social climate of correctional institutions. *J Res Crime Delinq*. 1968;5:174-188.

86. Moos RH, Houts PS. Assessment of the social atmospheres of psychiatric wards. *J Abnorm Psychol*. 1968;73(6):595-604.

87. Moos RH, Insel PM, Humphrey B. *Preliminary Manual for Family Environment Scale, Work Environment Scale, Group Environment Scale*. Palo Alto, CA: Consulting Psychologists Press; 1974.

Supplementary Online Content

Wasson LT, Cusmano A, Meli L, et al. Association between learning environment interventions and medical student well-being: a systematic review. *JAMA*. doi:10.1001/jama.2016.17573

eAppendix. Search Strategies

This supplementary material has been provided by the authors to give readers additional information about their work.

eAppendix. Search Strategies

All bibliographic databases were searched May 13, 2016 and updated on October 6, 2016. Clinicaltrials.gov and websites were searched June 10, 2013 and again on October 6, 2013.

Ovid MEDLINE
Searched from 1946 to May Week 1, 2016 and In-Process & Other Non-Indexed Citations May 11, 2016. Updated search 2016 to September Week 4 2016 and In-Process & Other Non-Indexed Citations October 3, 2016.

MEDLINE

1. Education, Medical, Undergraduate/
2. Students, Medical/
3. (medical adj (student$ or undergrad$)).tw.
4. or/1-3
5. exp Emotions/
6. personal satisfaction/
7. (feel$ or emotion$ or affect or mood$ or anxiet$ or anxious$ or fear$ or frustrat$ or happie$ or happy or sad or lonely or loneli$ or please$ or pleasure or hate$ or anger$ or guilt$ or shame$ or hope$ or hostil$ or jealous$ or satisf$).tw.
8. cultur$.tw.
9. (burnout or stress$ or wellness or well-being or depress$).tw.
10. (abus$ or mistreat$ or harass$ or hostil$ or punish$ or professional$ or unprofessional$ or support$ or unsupport$ or humiliat$ or disparag$ or ignor$ or unsafe or safe$ or harm$ or personal service or appropriat$ or innappropriat$ or respect$ or dispresct$).tw.
11. empathy/
12. (empath$ or compassion or collegial$ or resilien$ or cooperat$ or collaborat$ or kind$ or integrity).tw.
13. exp prejudice/
14. (discriminat$ or sexism or sexist or sexual or racis$ or race or ethnic$ or bias$).tw.
15. or/5-14
16. exp "Surveys and Questionnaires"/
17. interview/
18. Interview, Psychological/
19. feedback/
20. (survey$ or questionnaire$).tw.
21. ((program$ or course or curricul$) adj2 (evaluat$ or feedback or view$ or opinion$)).tw.
22. (structure$ adj interview$).tw.
23. or/16-22
24. exp United States/

25. (united states or usa or america$).tw.

26. 24 or 25

27. and/4,15,23,26

28. limit 27 to english language

The Cochrane Library
All databases were searched from inception to May 13, 2016. Updated search year 2016 only.

#1 MeSH descriptor: [Education, Medical, Undergraduate] this term only
#2 MeSH descriptor: [Students, Medical] this term only
#3 (medical next (student* or undergrad*)):ti,ab
#4 #1 or #2 or #3
#5 MeSH descriptor: [Emotions] explode all trees
#6 MeSH descriptor: [Personal Satisfaction] this term only
#7 (feel* or emotion* or affect or mood* or anxiet* or anxious* or fear* or frustrat* or happi* or happy or sad or lonely or loneli* or please* or pleasure or hate* or anger* or guilt* or shame* or hope* or hostil* or jealous* or satisf*):ti,ab
#8 cultur*:ti,ab
#9 (burnout or stress* or wellness or well-being or depress*):ti,ab
#10 (abus* or mistreat* or harass* or hostil* or punish* or professional* or unprofessional* or support* or unsupport* or humiliat* or disparag* or ignor* or unsafe or safe* or harm* or "personal service" or appropriat* or innappropriat* or respect* or dispresct*):ti,ab
#11 MeSH descriptor: [Empathy] this term only
#12 (empath* or compassion or collegial* or resilien* or cooperat* or collaborat* or kind* or integrity):ti,ab
#13 MeSH descriptor: [Prejudice] explode all trees
#14 (discriminat* or sexism or sexist or sexual or racis* or race or ethnic* or bias*):ti,ab
#15 #5 or #6 or #7 or #8 or #9 or #10 or #11 or #12 or #13 or #14
#16 MeSH descriptor: [Surveys and Questionnaires] explode all trees
#17 MeSH descriptor: [Interview] this term only
#18 MeSH descriptor: [Interview, Psychological] this term only
#19 MeSH descriptor: [Feedback] this term only
#20 (survey* or questionnaire*):ti,ab
#21 ((program* or course or curricul*) near/2 (evaluat* or feedback or view* or opinion*)):ti,ab
#22 (structure* next interview*):ti,ab
#23 #16 or #17 or #18 or #19 or #20 or #21 or #22
#24 MeSH descriptor: [United States] explode all trees
#25 ("united states" or usa or america*):ti,ab
#26 #24 or #25
#27 #4 and #15 and #23 and #26

EMBASE (EMBASE.com)
Searched from 1980 to May 13, 2016. Updated search 2016 to October 6, 2016.

#27. #5 AND #18 AND #23 AND #27 AND [english]/lim AND [embase]/lim
#26. #24 OR #25

#25. 'united states':ab,ti OR usa:ab,ti OR america*:ab,ti
#24. 'united states'/exp
#23. #19 OR #20 OR #21 OR #22
#22. 'structure* interview*':ab,ti
#21. ((program* OR course OR curricul*) NEAR/2 (evaluat* OR feedback OR view* OR opinion*)):ab,ti
#20. survey*:ab,ti OR questionnaire*:ab,ti
#19. 'interview'/de OR 'semi structured interview'/de 163,741 13 May 2016
#18. #6 OR #7 OR #8 OR #9 OR #10 OR #11 OR #12 OR #13 OR #14 OR #15 OR #16 OR #17
#17. discriminat*:ab,ti OR sexism:ab,ti OR sexist:ab,ti OR sexual:ab,ti OR racis*:ab,ti OR race:ab,ti OR ethnic*:ab,ti OR bias*:ab,ti
#16. 'sexism'/exp
#15. 'racism'/de
#14. 'prejudice'/de
#13. empath*:ab,ti OR compassion:ab,ti OR collegial*:ab,ti OR resilien*:ab,ti OR cooperat*:ab,ti OR collaborat*:ab,ti OR kind*:ab,ti OR integrity:ab,ti
#12. 'empathy'/de
#11. abus*:ab,ti OR mistreat*:ab,ti OR harass*:ab,ti OR hostil*:ab,ti OR punish*:ab,ti OR professional*:ab,ti OR unprofessional*:ab,ti OR support*:ab,ti OR unsupport*:ab,ti OR humiliat*:ab,ti OR disparag*:ab,ti OR ignor*:ab,ti OR unsafe:ab,ti OR safe*:ab,ti OR harm*:ab,ti OR 'personal service':ab,ti OR appropriat*:ab,ti OR innappropriat*:ab,ti OR respect*:ab,ti OR dispresct*:ab,ti
#10. burnout:ab,ti OR stress*:ab,ti OR wellness:ab,ti OR 'well being':ab,ti OR depress*:ab,ti
#9. cultur*:ab,ti
#8. feel*:ab,ti OR emotion*:ab,ti OR affect:ab,ti OR mood*:ab,ti OR anxiet*:ab,ti OR anxious*:ab,ti OR
fear*:ab,ti OR frustrat*:ab,ti OR happi*:ab,ti OR happy:ab,ti OR sad:ab,ti OR lonely:ab,ti OR loneli*:ab,ti OR please*:ab,ti OR pleasure:ab,ti OR hate*:ab,ti OR anger*:ab,ti OR guilt*:ab,ti OR
shame*:ab,ti OR hope*:ab,ti OR hostil*:ab,ti OR jealous*:ab,ti OR satisf*:ab,ti
#7. 'student satisfaction'/exp
#6. 'emotion'/exp
#5. #1 OR #2 OR #3 OR #4
#4. 'medical undergrad*':ab,ti
#3. 'medical student*':ab,ti
#2. 'medical student'/de
#1. 'medical school'/exp

ERIC (EBSCOhost)
Searched from 1966 to May 13, 2016. Updated search 2016 to October 6, 2016.

S1 DE "Medical Students"
S2 DE "Medical Schools"
S3 TI ((medical next (student* or undergrad*))) OR AB ((medical next (student* or undergrad*)))
S4 TI medical student* OR AB medical undergrad* OR TI medical student* OR AB medical undergrad*
S5 S1 OR S2 OR S3 OR S4
S6 DE "Satisfaction"

S7 TI (feel* OR emotion* OR affect OR mood* OR anxiet* OR anxious* OR fear* OR frustrat* OR happi* OR happy OR sad OR lonely OR loneli* OR please* OR pleasure OR hate* OR anger* OR guilt* OR shame* OR hope* OR hostil* OR jealous* OR satisf*) OR AB (feel* OR emotion* OR affect OR mood* OR anxiet* OR anxious* OR fear* OR frustrat* OR happi* OR happy OR sad OR lonely OR loneli* OR please* OR pleasure OR hate* OR anger* OR guilt* OR shame* OR hope* OR hostil* OR jealous* OR satisf*)
S8 TI cultur* OR AB cultur*
S9 TI (burnout OR stress* OR wellness OR well-being OR depress*) OR AB (burnout OR stress* OR wellness OR well-being OR depress*)
S10 TI (abus* OR mistreat* OR harass* OR hostil* OR punish* OR professional* OR unprofessional* OR support* OR unsupport* OR humiliat* OR disparag* OR ignor* OR unsafe OR safe* OR harm* OR "personal service" OR appropriat* OR innappropriat* OR respect* OR dispresct*) OR AB (abus* OR mistreat* OR harass* OR hostil* OR punish* OR professional* OR unprofessional* OR support* OR unsupport* TI (empath* OR compassion OR collegial* OR resilien* OR cooperat* OR collaborat* OR kind* OR integrity) OR AB (empath* OR compassion OR collegial* OR resilien* OR cooperat* OR collaborat* OR kind* OR integrity)humiliat* OR disparag* OR ignor* OR unsafe OR safe* OR harm* OR "personal service" OR appropriat* OR innappropriat* OR respect* OR dispresct*)
S11 DE "Empathy"
S12 DE "Bias"
S13 DE "Gender Bias"
S14 DE "Racial Bias"
S15 TI (discriminat* OR sexism OR sexist OR sexual OR racis* OR race OR ethnic* OR bias*) OR AB (discriminat* OR sexism OR sexist OR sexual OR racis* OR race OR ethnic* OR bias*)
S16 S6 OR S7 OR S8 OR S9 OR S10 OR S11 OR S12 OR S13 OR S14 OR S15
S17 DE "Student Surveys"
S18 DE "Semi Structured Interviews" OR DE "Structured Interviews"
S19 DE "Feedback (Response)"
S20 TI (survey* OR questionnaire*) OR AB (survey* OR questionnaire*)
S21 TI (((program* OR course OR curricul*) N2 (evaluat* OR feedback OR view* OR opinion*))) OR AB (((program* OR course OR curricul*) N2 (evaluat* OR feedback OR view* OR opinion*)))
S22 TI structur* interview* OR AB structur* interview*
S23 S17 OR S18 OR S19 OR S20 OR S21
S24 TI (("united states" OR usa OR america*)) OR AB (("united states" OR usa OR america*))
S25 S5 AND S16 AND S23 AND S24

Clinicaltrials.gov

student | Closed Studies | Studies With Results
Closed Studies | Studies With Results | stress | United States | Adult | Other
Closed Studies | Studies With Results | anxiety OR depression | United States | Adult | Other

Websites Searched

- AAMC (Association of American Medical Colleges)
- AMA (American Medical Association)
- APSA (American Physician Scientists Association)

- AMSA (American Medical Student Association)
- LCME (Liaison Committee on Medical Education)
- ACGME (Accreditation Council for Graduate Medical Education)