Published in final edited form as:
JAMA. 2016 December 06; 316(21): 2237–2252. doi:10.1001/jama.2016.17573.

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

Lauren T. Wasson, MD, MPH[1], **Amberle Cusmano, MA**[1], **Laura Meli, MS Ed**[1], **Irene Louh, MD, PhD**[1], **Louise Falzon, PGDipInf**[1], **Meghan Hampsey, High School Diploma**[2], **Geoffrey Young, PhD**[3], **Jonathan Shaffer, PhD, MS**[4], and **Karina W. Davidson, PhD, MASc**[1,5]

[1]Center for Behavioral Cardiovascular Health, Columbia University Medical Center, New York, NY

[2]George Washington University

[3]Association of American Medical Colleges, Washington, D.C.

[4]Department of Psychology, College of Liberal Arts and Sciences, University of Colorado at Denver

[5]NewYork-Presbyterian Hospital, New York, NY

Abstract

Importance—There are concerns about the current quality of undergraduate medical education (UME) and its effect on students' well-being.

Objective—This systematic review was designed to identify best practices for UME 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 the ERIC database from the 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 5–18; higher scores indicate higher design and methods quality, and a score of ≥ 14 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 post-test only (n=10), single-group pre-/post-test (n=2), nonrandomized two-group (n=13), and randomized clinical trial (n=3); 93% were

Corresponding Author: Karina W. Davidson, Center for Behavioral Cardiovascular Health, Columbia University Medical Center, 622 W. 168th Street, PH 9-314, New York, NY 10032; Telephone: 212-342-4493; Fax: 212-342-3431; kd2124@cumc.columbia.edu.

Conflicts of Interest: No potential conflicts of interest relevant to this article were reported.

Access to Data: Drs. Davidson and Wasson had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

conducted at a single site, and the mean MERSQI score for all studies was 10.3 (range 5–13, SD=2.11). Studies encompassed a variety of types of interventions, including those focused on pass/fail grading systems (n=3, mean MERSQI=12.0), mental health programs (n=4, MERSQI=11.9), mind-body skills programs (n=7, MERSQI=11.2), curriculum structure (n=3, MERSQI=9.5), multicomponent program reform (n=5, MERSQI=9.4), wellness programs (n=4, MERSQI=9.0), and advising/mentoring programs (n=3, MERSQI=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.

Keywords

undergraduate medical education; student well-being; intervention; medical students; satisfaction

Introduction

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]

The Association of American Medical Colleges (AAMC) 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.

Methods

Search strategy

Potentially relevant articles were identified by searching the biomedical electronic databases Ovid MEDLINE, EMBASE, the Cochrane Library, and the ERIC database from the database inception dates to October 2016 (Appendix A). Additional records were identified by scanning the reference lists of relevant studies and reviews published between May 2011 and October 2016, by using the *Similar Articles* feature in PubMed and the *Cited Reference Search* in the Web of Science. We searched for grey 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, AAMC) (Appendix A).

Selection criteria

Studies must 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 methodology rating used in this review. Medical education interventions measured with open-ended responses are reviewed and appraised elsewhere.[8,9]

Methodology 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 six 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 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 cut-off values differentiating high-quality from low-quality study methods, one study used a MERSQI score of ≥ 14.0 as an a priori cutoff of high quality.[12]

Data extraction

Two review authors 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 (Table 1). Data disagreements were resolved by consultation with the third and fourth review authors. The original intention noted in the study protocol was to conduct a meta-analysis, but due to 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.[13]

Results

The literature search yielded 4207 publications, of which 28 met the eligibility criteria for this systematic review (Figure 1). 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 that contained interventions not focused on the learning environment. The studies included at least 8224 (one study did not report a sample size) student participants and encompassed a variety of designs, including single-group, cross-sectional or post-test only (n=10), single-group pre-/post-test (n=2), nonrandomized two-group (n=13), and randomized clinical trial (RCT) (n=3) designs; 96% 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 [P/F] 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 detailed in Tables 1 and 2. The included studies' methodologic rigor varied, with MERSQI scores ranging from 5.0 to 13.0 (mean 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 highest methodology studies crossed all types of interventions and all types of outcome measures. The highest scored categories tested interventions involving P/F grading, mental health programs, and mind-body skills education/training.

Pass/fail grading system (average MERSQI=12.0)

Bloodgood et al[14] (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 P/F 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,[14] and stress,[15] and better well-being[14] and group cohesion scores at various study timepoints.[15] These two studies differed, however, in the durability of improvements. Bloodgood et al[14] found no difference at 2 years between the cohort of students with a 2-year P/F system compared to a cohort of students with a 5-interval system on measures of anxiety (General Well-Being Schedule [GWB][16] anxiety subscore range 3–28; lower scores indicate more severe distress; there is no accepted minimum clinically important difference [MCID]; M=14.08 vs. 14.20; *P*=.86), depression (GWB[16] depression subscore range 2–22; lower scores indicate more severe distress; there is no accepted MCID; M=15.56 vs. 15.35; *P*=0.71), or well-being (GWB[16] well-being subscore range 3–18; lower scores indicate more severe distress; there is no accepted MCID; M=10.59 vs. 10.40; *P*=.67). Rohe et al[15] did report a persistent difference at 2 years between grading cohorts on a measure of stress (Perceived Stress Scale [PSS][17] range 0–40; higher score indicates more stress; there is no accepted MCID; M[SD]=15.8 [6.8] vs. 20.5 [7.8]; *P*=.01) and speculated this difference was due to continuing reports of elevated group cohesion (Perceived Cohesion Scale[18] range 0–36; higher scores indicate more cohesion; there is no accepted MCID; M[SD]=33.8 [8.0] vs. 29.0 [9.9]; *P*=.02).

Reed et al[19] (n=2056, MERSQI=12.5) compared well-being among students at different medical schools with grading systems that were categorized as either 3+-interval (eg, honors/P/F) or P/F and found that 3+-interval systems were associated with statistically

significantly more stress (β=1.91; 95% confidence interval [CI], 1.05–2.78; *P*<.001) and burnout (odds ratio [OR]=1.58; 95% CI, 1.24–2.01; *P*<.001), and a higher likelihood of considering withdrawing from medical school (OR=1.91; 95% CI, 1.30–2.80; *P*=.001).

Mental health programs (MERSQI=11.9)

Thompson et al[20] (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).[22] Seritan et al[21] (n=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.[21] 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.[21]

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[22] (n=1008, MERSQI = 13.0) described a program that was associated with an increase in mental health service utilization and a decrease in assessed suicide risk during the 4 years that was not statistically significant, perhaps due to low screening rates (34%). Moutier et al[23] (n=498, MERSQI = 11.5) reported that that 11% of medical students exposed to another educational program were referred to a mental health professional, though no comparison was provided and the screening rate was also low (27%).

Mind-body skills education/training programs (MERSQI=11.3)

Two RCTs evaluated mind-body programs. Erogul et al[24] (n=58, MERSQI=12.0) found that students randomized to attend a mindfulness program reported a significant reduction in stress after intervention (PSS[17] range 0–40; higher score indicates more stress; there is no accepted MCID; M_{change}=3.63; 95% CI, 0.37–6.89; *P*=.03) but not at 6-month follow-up (M_{change}=2.91; 95% CI, −0.37–6.19; *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 [SCS][25] range 0–5; higher score indicates more self-compassion; there is no accepted MCID; M_{change}=0.56; 95% CI, 0.25–0.87; *P*=.001).[24] In the study by Holtzworth-Munroe et al[26] (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 pre-/post-test design with nonrandomized control groups. Kraemer et al[27] (n=52, MERSQI=12.0) found that students undergoing mind-body skills training reported significantly improved distress tolerance (Distress Tolerance Scale-G[28] range 1–5; higher scores indicate higher levels of distress tolerance; there is no accepted MCID; M_{change}=0.53; 95% CI, .92 to .14; *P*=0.01); no

difference was found for the control group. Rosenzweig et al[29] (n=302, MERSQI=11.0) described a mindfulness-based stress reduction program associated with significant improvements in total mood disturbance (Profile of Mood States[30] range 0–200; higher scores indicate higher mood disturbance; there is no accepted MCID; intervention group pre-M[SD]=38.7 [33.3] vs. post-M=31.8 [33.89]; *P*=0.05; control group pre-M[SD]=28.0 [31.2] vs. post-M=38.6 [32.8]; *P*<0.001; interaction *P*<.001). Finkelstein et al[31] (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 pre-/post-test design without a control group also reported associations with significant improvements in well-being. Greeson et al[32] (n=44, MERSQI=11.5) reported improved stress (PSS[17] range 0–40; higher score indicates more stress; there is no accepted MCID; pre-M[SD]=29.73 [9.61]; post-M[SD]=20.25 [9.03]; t_{33}=7.90; *P*<.001; d=1.38) and mindfulness (pre-M[SD]=29.24 [5.54]; post-M[SD]=33.88 [6.13]; t_{33}=5.27; *P*<.001; d=0.92). Bond et al[33] (n=27, MERSQI=11.5) reported improved self-regulation (Self Regulation Questionnaire[34] range 1–5; higher score indicates more self-regulation; there is no accepted MCID; M_{change}[SD]= 0.13 [0.20]; *P*=.003; d=−0.41) and self-compassion (SCS[25] range 0–5; higher score indicates more self-compassion; there is no accepted MCID; M_{change}[SD]=0.28 [0.61]; *P*=.04; d= −0.55).

Curriculum structure (MERSQI=9.5)

Elements of curriculum structure targeted by studies identified in this review were varied. Reed et al[19] (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 (OR=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).[19]

Camp et al[35] (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).[36]

Multicomponent program reform (MERSQI=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[37] (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[38] (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[39] (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 Saint Louis University School of Medicine also undertook multicomponent program reform that was introduced in phases to preclinical students: (1) P/F grading for preclinical courses, reduced preclinical contact hours, extended electives, and learning communities; (1/2) addition of mind-body skills training; and (1/2/3) addition of anatomy course reform. As reported in a study by Slavin et al[40] (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. Phase 1/2 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.[40] Phase 1/2/3 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.[40]

Strayhorn[41] (n=478, MERSQI=12.0) compared one school's curriculum changes to a comparison school 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 (MERSQI=9.0)

In self-hypnosis training RCT, Whitehouse et al[42] (n=35, MERSQI=12.0) reported significant improvements in anxiety (Brief Symptom Inventory[43] range 20–80; higher scores indicate higher anxiety; there is no accepted MCID; $F[3,96]=2.96$; $P<.05$).[42] 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.[44] 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).[45] Kushner et al[46] (n=343, MERSQI=8.0) evaluated a wellness course that included a section on behavior change plans; out 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 (MERSQI=8.2)

Three studies evaluated small group-based faculty advisor/mentor programs that were formally integrated into the academic curriculum. Sastre et al[47](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 with their advisor (51% vs. 27%; $P<.001$).[47] Coates et al[48] (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[49] (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 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 the 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 P/F grading system should be considered. All of the studies reviewed here show that a preclinical P/F 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 P/F 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.[14,15] This is consistent with other literature exclusively focused on academic outcomes of P/F grading.[50–52] According to the 2014–2015 Liaison Committee on Medical Education Annual Medical School Questionnaire, 87 of the 144 participating schools used P/F grading systems for at least some portion of the preclinical courses.[53]

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.[54] 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.[55] 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 professionals seems to be an efficacious approach, and is associated with lower depression and suicidal ideation rates.[20]

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, including fear of documentation in the academic record and evaluators' knowledge of student mental health conditions with subsequent career implications.[55–59] Medical students reported preferring help from a mental health specialist, family, or friends, rather than medical school personnel,[58] and reported preferring accessing mental health services through a location other than the office of student affairs.[59] In other studies, students have reported concerns about time, convenience of office hours, location, and financial costs.[56–59] 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, mood, and distress tolerance. This association was found even when skills were taught in condensed workshops lasting only 4 weeks,[32] 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 one study tested for statistical significance.[47] However, it is important that mentors do not grade students, to keep their role as advisors separate from assessment so as to foster open communication.[49] 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.[49] 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."[60]

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.[19] 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.[40]

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.[61,62] Historically, research conducted in established educational settings and involving normal educational practices were considered exempt from institutional review board oversight.[63] However, issues of coercion and lack of informed consent about randomization of medical students when conducting learning environment interventions tests have recently been raised.[64,65] 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 a clerkship director or faculty advisor holds. Guidance is provided elsewhere for future UME educators to decide when and under what circumstances randomization is ethical and practical for learning environment interventions.[64,65]

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.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Funding: Dr. Davidson was supported by research grant K24 HL084034. Dr. Wasson was supported by research grant K08 HS024598 01A1. 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 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; and decision to submit the manuscript for publication.

References

1. Dezee KJ, Artino AR, Elnicki DM, et al. Medical education in the United States of America. Med Teach. 2012; 34(7):521–525. [PubMed: 22489971]
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. [PubMed: 25250752]
3. Dyrbye LN, Massie FS, Eacker A, et al. Relationship between burnout and professional conduct and attitudes among US medical students. JAMA. 2010; 304(11):1173–1180. [PubMed: 20841530]
4. Schwenk TL, Davis L, Wimsatt LA. Depression, stigma, and suicidal ideation in medical students. JAMA. 2010; 304(11):1181–1190. [PubMed: 20841531]
5. Dyrbye LN, Thomas MR, Massie FS, et al. Burnout and suicidal ideation among U.S. medical students. Ann Intern Med. 2008; 149(5):334–341. [PubMed: 18765703]

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. The Fourth International Conference on Grey Literature (GL '99). Washington, DC: 1999 Oct.

8. Stacy R, Spencer J. Assessing the evidence in qualitative medical education research. Med Educ. 2000; 34(7):498–500. [PubMed: 10886625]

9. Côté L, Turgeon J. Appraising qualitative research articles in medicine and medical education. Med Teach. 2005; 27(1):71–75. [PubMed: 16147774]

10. Reed DA, Cook DA, Beckman TJ, et al. Association between funding and quality of published medical education research. JAMA. 2007; 298(9):1002–1009. [PubMed: 17785645]

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. [PubMed: 26107881]

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. [PubMed: 26445081]

13. Egger M, Smith GD, Phillips AN. Meta-analysis: principles and procedures. BMJ. 1997; 315(7121):1533. [PubMed: 9432252]

14. 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. [PubMed: 19704204]

15. Rohe DE, Barrier PA, Clark MM, et al. The benefits of pass-fail grading on stress, mood, and group cohesion in medical students. Mayo Clin Proc. 2006; 81(11):1443–1448. [PubMed: 17120399]

16. Dupuy, HJ. Self-representations of general psychological well-being of American adults. Presented at: American Public Health Association Meeting; Oct. 15–19, 1978; Los Angeles, CA.

17. Cohen S, Kamarck T, Mermelstein R. A global measure of perceived stress. J Health Soc Behav. 1983; 24(4):385–396. [PubMed: 6668417]

18. Bollen KA, Hoyle RH. Perceived cohesion: a conceptual and empirical examination. Social Forces. 1990; 69(2):479–504.

19. 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. [PubMed: 21952063]

20. 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. [PubMed: 20881686]

21. 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. [PubMed: 24840666]

22. 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. [PubMed: 24705825]

23. 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. [PubMed: 22373625]

24. 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. [PubMed: 25318029]

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

26. 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. [PubMed: 3886910]

27. Kraemer KM, Luberto CM, O'Bryan EM, et al. Mind-body skills training to improve distress tolerance in medical students: a pilot study. Teach Learn Med. 2016; 28(2):219–228. [PubMed: 27064724]

28. Simons JS, Gaher RM. The Distress Tolerance Scale: development and validation of a self-report measure. Motivation and Emotion. 2005; 29(2):83–102.
29. Rosenzweig S, Reibel DK, Greeson JM, et al. Mindfulness-based stress reduction lowers psychological distress in medical students. Teach Learn Med. 2003; 15(2):88–92. [PubMed: 12708065]
30. McNair, DM.; Lorr, M.; Droppleman, LF. POMS Manual Profile of Mood States. San Diego, CA: EdiTS/Educational and Industrial Testing Service; 1992.
31. Finkelstein C, Brownstein A, Scott C, Lan YL. Anxiety and stress reduction in medical education: an intervention. Med Educ. 2007; 41(3):258–264. [PubMed: 17316210]
32. 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. [PubMed: 25792145]
33. 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.
34. University of New Mexico Center on Alcoholism, Substance Abuse, and Addictions. The self-regulation questionnaire. Albequerque, NM: Available from: http://casaa.unm.edu/inst/SelfRegulation%20Questionnaire%20(SRQ).pdf [cited 26 October 2016]
35. Camp DL, Hollingsworth MA, Zaccaro DJ, et al. Does a problem-based learning curriculum affect depression in medical students? Acad Med. 1994; 69(10 suppl):S25–S27. [PubMed: 7916817]
36. 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. [PubMed: 16184295]
37. 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. [PubMed: 20042835]
38. 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. [PubMed: 23887019]
39. 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.
40. 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. [PubMed: 24556765]
41. Strayhorn G. Effect of a major curriculum revision on students' perceptions of well-being. Acad Med. 1989; 64(1):25–29. [PubMed: 2914060]
42. 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. [PubMed: 8771625]
43. Derogatis LR, Melisaratos N. The Brief Symptom Inventory: an introductory report. Psychol Med. 1983; 13(3):595–605. [PubMed: 6622612]
44. Goetzel RZ, Croen LG, Shelov S, et al. Evaluating self-help support groups for medical students. J Med Educ. 1984; 59(4):331–340. [PubMed: 6708071]
45. Lee J, Graham AV. Students' perception of medical school stress and their evaluation of a wellness elective. Med Educ. 2001; 35(7):652–659. [PubMed: 11437967]
46. Kushner RF, Kessler S, McGaghie WC. Using behavior change plans to improve medical student self-care. Acad Med. 2011; 86(7):901–906. [PubMed: 21617509]
47. 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. [PubMed: 20854149]
48. Coates WC, Crooks K, Slavin SJ, et al. Medical school curricular reform: fourth-year colleges improve access to career mentoring and overall satisfaction. Acad Med. 2008; 83(8):754–760. [PubMed: 18667890]
49. 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. [PubMed: 6620345]

50. 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.
51. 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. [PubMed: 10353289]
52. Robins LS, Fantone JC, Oh MS, et al. The effect of pass/fail grading and weekly quizzes on first-year students' performances and satisfaction. Acad Med. 1995; 70(4):327–329. [PubMed: 7718068]
53. Number of Medical Schools Using Selected Grading Systems in Pre-Clerkship Courses (Excluding Physical Diagnosis/Clinical Skills). Association of American Medical Colleges. [Accessed October 16, 2016] website. https://www.aamc.org/initiatives/cir/406418/11.html.
54. Dyrbye LN, Thomas MR, Shanafelt TD. Systematic review of depression, anxiety, and other indicators of psychological distress among U.S. and Canadian medical students. Acad Med. 2006; 81(4):354–373. [PubMed: 16565188]
55. Givens JL, Tjia J. Depressed medical students' use of mental health services and barriers to use. Acad Med. 2002; 77(9):918–921. [PubMed: 12228091]
56. Gentile JP, Roman B. Medical student mental health services: psychiatrists treating medical students. Psychiatry (Edgmont). 2009; 6(5):38–45.
57. Schwenk TL, Davis L, Wimsatt LA. Depression, stigma, and suicidal ideation in medical students. JAMA. 2010; 304(11):1181–1190. [PubMed: 20841531]
58. 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. [PubMed: 25650824]
59. Plaut SM, Maxwell SA, Seng L, et al. Mental health services for medical students: perceptions of students, student affairs deans, and mental health providers. Acad Med. 1993; 68(5):360–365. [PubMed: 8484849]
60. Rose GL, Rukstalis MR, Schuckit MA. Informal mentoring between faculty and medical students. Acad Med. 2005; 80(4):344–348. [PubMed: 15793017]
61. Mosteller, F.; Boruch, RF. Evidence Matters: Randomized Trials in Education Research. Washington, DC: Brookings Institution Press; 2002.
62. Green, JL.; Camilli, G.; Elmore, PB. Handbook of Complementary Methods in Education Research. 3rd. Washington, DC: American Educational Research Association; 2006.
63. Federal Policy for the Protection of Human Subjects. Federal Register. US Department of Health and Human Services website; http://archive.hhs.gov/ohrp/documents/19881110.pdf [Accessed October 16, 2016]
64. Johansson AC, Durning SJ, Gruppen LD, et al. Perspective: medical education research and the institutional review board: reexamining the process. Acad Med. 2011; 86(7):809–817. [PubMed: 21617512]
65. Keune JD, Brunsvold ME, Hohmann E, et al. The ethics of conducting graduate medical education research on residents. Acad Med. 2013; 88(4):449–453. [PubMed: 23425981]
66. Cohen, S. Perceived stress in a probability sample of the United States. In: Oskamp, SSS., editor. The Social Psychology of Health. Thousand Oaks, CA: Sage Publications; 1988. p. 31-67.
67. Maslach C, Jackson SE. The measurement of experienced burnout. Journal of Organizational Behavior. 1981; 2(2):99–113.
68. Stewart, AL.; Ware, JE., editors. Measuring Functioning and Well-being: the Medical Outcome Study Approach. Durham, NC: Duke University Press; 1992.
69. Tarlov AR, Ware JE Jr, Greenfield S, et al. The Medical Outcomes Study. An application of methods for monitoring the results of medical care. JAMA. 1989; 262(7):925–930. [PubMed: 2754793]
70. Radloff LS. The CES-D Scale: a self-report depression scale for research in the general population. Applied Psychological Measurement. 1977; 1(3):385–401.
71. Graduation Questionnaire (GQ). American Association of Medical Colleges website. [Accessed October 21, 2016] https://www.aamc.org/data/gq/.

72. Spitzer RL, Kroenke K, Williams JB. Validation and utility of a self-report version of PRIME-MD: the PHQ primary care study. Primary Care Evaluation of Mental Disorders. Patient Health Questionnaire. JAMA. 1999; 282(18):1737–1744. [PubMed: 10568646]

73. 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. [PubMed: 11556941]

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

75. Thompson ER. Development and Validation of an Internationally Reliable Short-Form of the Positive and Negative Affect Schedule (PANAS). Journal of Cross-Cultural Psychology. 2007; 38(2):227–242.

76. Derogatis, LR. SCL-90-R administration, Scoring and Procedures Manual. Vol ll. Towson, MD: Clinical Psychometric Research; 1983.

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

78. 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. [PubMed: 9229283]

79. 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). Journal of Psychopathology and Behavioral Assessment. 2006; 29(3):177.

80. Hojat M, Mangione S, Nasca TJ, et al. The Jefferson Scale of Physician Empathy: development and preliminary psychometric data. Educational and Psychological Measurement. 2001; 61(2): 349–365.

81. Zung WW. A self-rating depression scale. Arch Gen Psychiatry. 1965; 12:63–70. [PubMed: 14221692]

82. 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. [PubMed: 7966923]

83. 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. [PubMed: 5456533]

84. 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: The RAND Corporation; 1979.

85. Russell DW. UCLA Loneliness Scale (Version 3): reliability, validity, and factor structure. J Pers Assess. 1996; 66(1):20–40. [PubMed: 8576833]

86. Moos RH. The assessment of the social climate of correctional institutions. Journal of Research in Crime and Delinquency. 1968; 5:174–188.

87. Moos RH, Hout P. The assessment of the social climate of the psychiatric ward. J Abnorm Psychol. 1968; 73(6):595–604. [PubMed: 5717365]

88. 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.

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. There is a need for high-quality research.

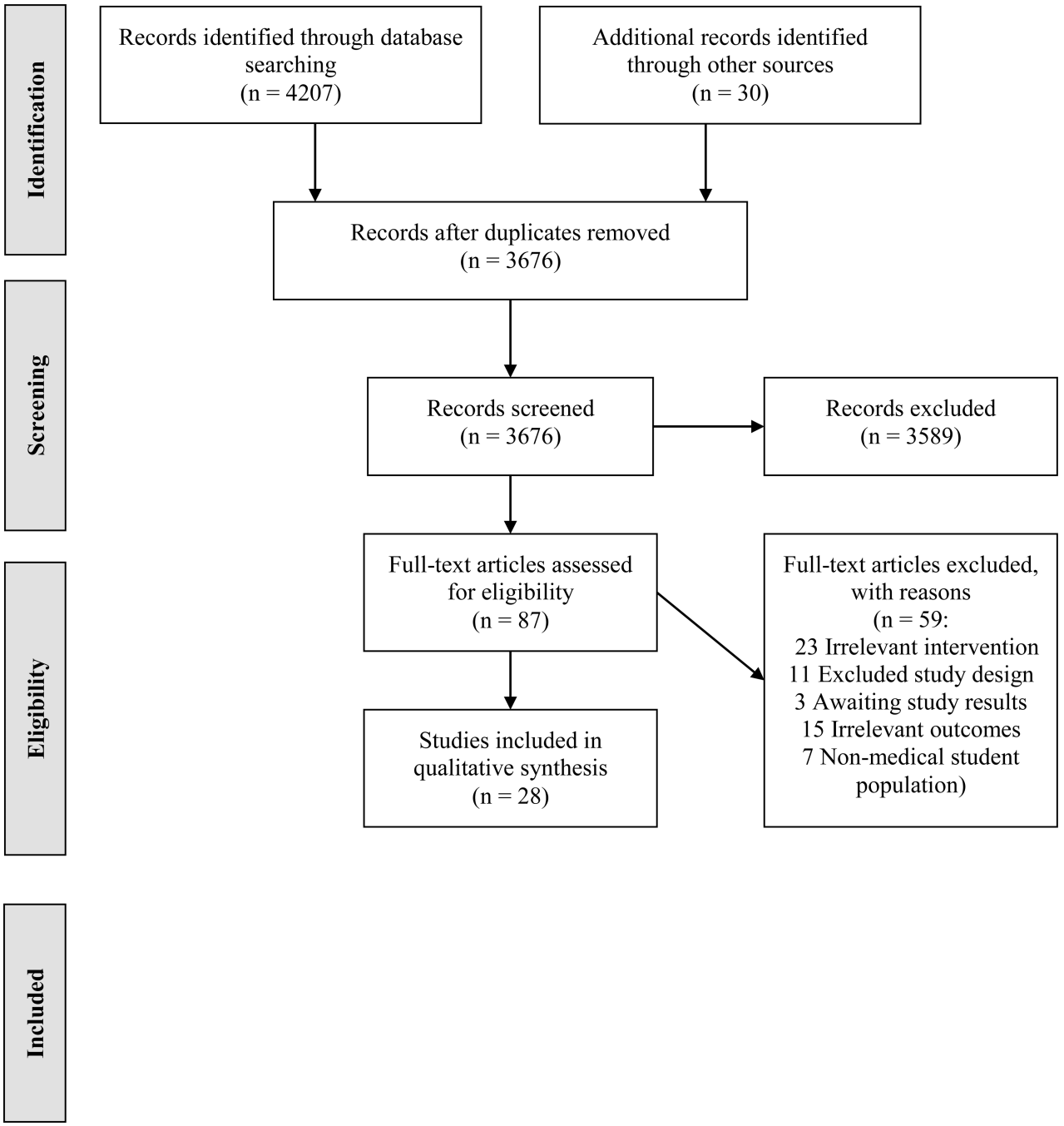

Figure 1.
PRISMA Flow Diagram for Systematic Review on the Association Between Learning Environment Interventions and Medical Student Well-being

Table 1

Study comparison and outcome measures

Study	Design	Population	Comparison(s)	Sample Size(s)	Outcome Measures	Main Results	*P* Value
Pass/Fail Grading Systems							
Bloodgood et al,[14] 2009	Nonrandomized 2 group	First- and second-year medical students	Cohort with pass/fail grading system versus 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[a]	Pass/fail versus graded after semester 1 (scores) • Anxiety: 18.14 (versus 15.98) • Depression: 17.62 (versus 15.89) • Well-being: 13.02 (versus 11.02) • Self-control: 15.51 (versus 14.12) • Vitality: 14.60 (versus 12.15) • General health: 12.56 (versus 11.48) Pass/fail versus graded after semester 2 (scores) • Anxiety: 19.01 (versus 17.65) • Depression: 17.61 (versus 16.65) • Well-being: 13.09 (versus 12.20) • Self-control: 15.10 (versus 14.45) • Vitality: 15.16 (versus 13.31) • General health: 11. 99 (versus 11.24) Pass/fail versus graded after semester 3 (scores) • Anxiety: 17.02 (versus 14.55) • Depression: 16.92 (versus 15.08) • Well-being: 12.37 (versus 10.74) • Self-control: 15.13 (versus 14.40) • Vitality: 14.10 (versus 11.95) • General health: 11.25 (versus 10.84) Pass/fail versus graded after semester 4 (scores) • Anxiety: 14.08 (versus 14.20) • Depression: 15.56 (versus 15.35)	.002 < .001 < .001 < .001 .02 .05 .05 .03 .13 .001 .15 .001 .001 < .001 .08 < .001 .47 .86 .71 .67 .63 .11 .99

Study	Design	Population	Comparison(s)	Sample Size(s)	Outcome Measures	Main Results	*P* Value
						• Well-being: 10.59 (versus 10.40) • Self-control: 14.61 (versus 14.42) • Vitality: 12.88 (versus 12.06) • General health: 11.30 (versus 11.31)	
Rohe et al,[15] 2006	Nonrandomized 2 group	First- and second-year medical students	Cohort with pass/fail grading system versus earlier cohort with 5-interval grading system (A/B/C/D/F)	n = 81 Pass/fail = 40 5-interval = 41	• Perceived Stress Scale[a] • Profile of Mood States[a] • Perceived Cohesion Scale[a]	Pass/fail versus graded at end of first year (score): • Perceived Stress Scale: 10.9 (SD 6.2) versus 13.8 (SD 6.4) • Profile of Mood States: 13.0 (SD 23.5) versus 32.0 (SD 39.0) • Perceived Cohesion Scale: 37.8 (SD 5.5) versus 32.9 (SD 8.4) Pass/fail versus graded at end of second year (score): • Perceived Stress Scale: 15.8 (SD 6.8) versus 20.5 (SD 7.8) • Profile of Mood States: 47.1 (SD 31.9) versus 64.6 (SD 40.5) • Perceived Cohesion Scale: 33.8 (SD 8.0) versus 29.0 (SD 9.9)	.02 .02 .01 .01 .07 .02
Reed et al,[19] 2011	Nonrandomized ≥ 2 groups	First- and second-year medical students	Institutions with pass/fail grading systems versus institutions with 3+ interval grading systems (eg, honors/pass/fail)	n = 2,056 n = 1,192 (responded) Pass/fail = 701 3+ interval = 491	• Maslach Burnout Inventory[a] • Perceived Stress Scale[a] • Medical Outcomes Study Short Form[a]	• Perceived stress score: β = 1.91; 95% CI, 1.05 to 2.78 • Mental quality of life: β = −2.79; 95% CI, −4.09 to −1.5 • Burnout: OR 1.85; 95% CI, 1.24 to 2.01 • Seriously considered dropping out of medical school in the past year: OR 1.91; 95% CI, 1.30 to 2.80	< .001 < .001 < .001 .001
Mental Health Programs							
Thompson et al,[20] 2010	Nonrandomized 2 group	Third-year medical students	Cohort with multi-pronged mental health program versus earlier cohort without the program	n = 120 Program cohort = 62 Earlier cohort = 58	• Center for Epidemiologic Studies Depression Scale[a] question on suicidal ideation	Pre-/post-intervention (frequency) • Depressive symptoms: 26/44 (59.1%) versus 14/58 (24.1%); $\chi^2 = 12.84$; df = 2 • Suicidal ideation: 13/43 (30.2%) versus 1/33 (3.0%); $\chi^2 = 13.05$; df = 1	< .01 < .001
Seritan et al,[21] 2013	Nonrandomized ≥ 2 group	All years	Cohort with mental health program versus earlier cohort without program	No sample size provided for number of students referred to	• American Medical Colleges Graduation Questionnaire[a]	Mental health service self-referral (percentage, no numbers provided) • Time 1: 50% • Time 2: 88% • Time 3: 91%	---

Study	Design	Population	Comparison(s)	Sample Size(s)	Outcome Measures	Main Results	*P* Value
			and versus national average	services Accreditation Council for Graduate Medical Education Graduation Survey, n = 525		Other referral • Time 1: 50% • Time 2: 12% • Time 3: 9% Satisfaction with program versus national average, pre- and most recent postintervention (score) • Personal counseling • 2009: 3.5 (3.7) • 2013: 4.4 (4.0) • Student mental health services • 2009: 3.5 (3.6) • 2013: 4.3 (4.0) • Stress-management programs (postintervention only) • 2009: 3.6 (3.8) • 2013: 4.3 (3.9)	
Downs et al,[22] 2014	Single group cross-sectional or post-test only	All years	---	n = 1008 (program) n = 343 (program and screen)	• Patient Health Questionnaire-9[a]	Among those screened, mental health service utilization (percentage, no numbers provided) • Year 1: 11.5% • Year 4: 15.0% • χ^2: 1.27, df = 3 Among those screened, suicide risk • Year 1: 8.8% • Year 4: 6.2% • χ^2 = 0.45; df = 3	--- --- NS --- NS ---
Moutier et al,[23] 2012	Single group cross-sectional or post-test only	All years	---	n = 498 n = 132 (screened)	• Items from Patient Health Questionnaire-9[a] 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,[24] 2014	Randomized clinical trial	First-year medical students	Mindfulness-based stress reduction intervention versus control (randomized)	n = 58 Intervention = 28 Control = 30	• Perceived Stress Scale[a] • Self-Compassion Scale[a]	Change in case from pre- to postintervention (change score) • Perceived Stress Scale: 3.63; 95% CI, 0.37 to 6.89 • Self-Compassion Scale: 0.58; 95% CI, 0.23 to 0.92 Change in case from pre- to 6-month follow-up (change score) • Perceived Stress Scale: 2.91; 95% CI, −0.37 to 6.19 • Self-Compassion Scale: 0.56; 95% CI, 0.25 to 0.87	.03 .002 .08 .001

Study	Design	Population	Comparison(s)	Sample Size(s)	Outcome Measures	Main Results	*P* Value
Holtzworth-Munroe et al,[26] 1985	Randomized clinical trial	First- and second-year medical students	Mind-body program versus control (randomized)	n = 40 Intervention = 20 Control = 20	• Spielberger Trait Anxiety Inventory[a] • Anxiety in test and social situation questionnaire • Tension and depression questionnaire • Self-esteem measure • Stress questionnaire	Intervention versus control at follow-up (score) • More aware of tension: $F(5, 18) = 37.16$ • Dealing better with school stress: $F(5, 18) = 5.05$ Anxiety before test (score): $F(1, 22) = 10.42$	< .001 < .04 < .005
Kraemer et al,[27] 2016	Nonrandomized 2 group	First- and second-year medical students	Mind-body program versus control (non-randomized)	n = 52 Intervention = 28 Control = 24	• Distress Tolerance Scale[a] • Perceived Stress Scale-10[a] • Positive Affect Negative Affect Schedule[a]	Changes in distress tolerance (change score) • Mind-body: 0.53; $t = -2.81$; 95% CI, 0.92 to 0.14 • Control: 0.25; $t = -1.66$; 95% CI, −0.06 to 0.55 •	.01 .11
Rosenzweig et al,[29] 2003	Nonrandomized 2 group	Second-year medical students	Mindfulness-based stress reduction program versus control (non-randomized)	n = 302 Intervention = 140 Control = 162	• Profile of Mood States[a]	Profile of Mood States total mood disturbance for intervention versus control (score) • Intervention: 38.7 (SD 33.3) versus 31.8 (SD 33.8) • Control: 28.0 (SD 31.2) versus 38.6 (SD 32.8) • Interaction: $d = -0.18$	.05 <.001 < .001
Finkelstein et al,[31] 2007	Nonrandomized 2 group	Second-year medical students	Mind-body elective versus control (nonrandomized)	n = 72 Intervention = 26 Control = 46	• Symptom Checklist-90 Anxiety Subscale[a] • Profile of Mood States[a] • Perceived Stress of Medical School Scale[a] • The 2-item Depression Index[a]	Time/group interaction for scores • Anxiety (Symptom Checklist-90): $F(1,2) = 3.95$ • The Profile of Mood States: $F(1,2) = 3.77$ • Perceived Stress of Medical School Scale: $F(1,2) = .11$	< .05 < .05 NS
Greeson et al,[32] 2015	Single group pre- and post-test	All years	Before versus after mind-body skills intervention	n = 44	• Cognitive and Affective Mindfulness Scale-Revised[a] • Perceived Stress Scale[a] • Open-ended feedback	Pre-/post-intervention (score) • Perceived Stress Scale: 29.73 (SD 9.61) versus 20.25 (SD 9.03) ○ $t(33) = 7.90$; $d = 1.38$ • Mindfulness: 29.24 (SD 5.54) versus 33.88 (SD 6.13) ○ $t(33) = 5.27$; $d = 0.92$	< .001 < .001
Bond et al,[33] 2013	Single group pre- and post-test	First- and second-year medical students	Before versus after mind-body course	n = 27	• Cohen's Perceived Stress Scale • Self-regulation questionnaire[a] • Self-Compassion Scale[a] • Jefferson Scale of Physician Empathy[a]	Pre-/post-intervention (change score): • Perceived stress: −.05 (SD 0.62); $d = .14$ • Self-regulation: 0.13 (SD 0.2); $d = -0.41$ • Self-compassion: 0.28 (SD 0.61); $d = -0.55$	.70 .003 .04 .30

Study	Design	Population	Comparison(s)	Sample Size(s)	Outcome Measures	Main Results	*P* Value
						• Empathy: 0.11 (SD 0.5); $d = -0.13$	
Curriculum Structure							
Reed et al,[19] 2011	Nonrandomized; ≥ 2 groups	First- and second-year medical students	7 institutions' curriculum structures	n = 2056 n = 1192 (responded)	• Maslach Burnout Inventory[a] • Perceived Stress Scale[a] • Medical Outcomes Study Short Form[a]	Association between clinical experiences and the following scores • Perceived stress: β .02; 95% CI [−.10 to .13] • Burnout: OR 1.01; 95% CI, 0.98 to 1.05 • Mental quality of life: β .00; 95% CI, −0.16 to 0.16 • Serious thoughts of dropping out: OR 0.96; 95% CI, 0.93 to 1.00 Association between testing experiences and the following scores • Perceived stress: β .29; 95% CI, 0.10 to 0.84 • Burnout: OR 1.10; 95% CI, 0.89 to 1.23 • Mental quality of life: b–β −.63; 95% CI, −0.29 to 0.96 • Serious thoughts of dropping out: OR 1.19; 95% CI, 1.12 to 1.27 Association between number of tests and the following scores • Perceived stress: β −0.02; 95% CI, −0.6 to 0.03 • Burnout: OR 0.99; 95% CI, 0.97 to 1.01 • Mental quality of life: β 0.03; 95% CI, −0.05 to 0.04 • Serious thoughts of dropping out: OR 1.00; 95% CI, 0.97 to 1.02	.79 .42 .98 .03 .003 .09 < .001 < .001 .48 .19 .44 .82
Camp et al,[35] 1994	Nonrandomized 2 group	First- and second-year medical students	Problem-based learning versus lecture-based learning	n = 275 Problem-based learning = 60 Lecture-based learning = 215	• Zung Self-Rating Depression Scale[a]	Depression problem-based learning versus lecture-based learning (score) • Overall OR 0.42; 95% CI, 0.14 to 1.21 • Adjustment for sex and self actualization OR 0.45; 95% CI, 0.14 to 1.42	.07 .14

Study	Design	Population	Comparison(s)	Sample Size(s)	Outcome Measures	Main Results	*P* Value
Kornitzer et al,[36] 2005	Cross-sectional post-test 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 (percentage, no numbers provided) • Gained confidence: 85.7% • Made the transition easier: 100% • Made friends:100% Humanities and Medicine group (percentage, no numbers provided) • Gained confidence: 97% • Made the transition easier: 97% • Made friends: 93.9%	--- --- --- --- --- ---
Multi-component Program Reform							
Drolet and Rodgers,[37] 2010	Single group cross-sectional or post-test only	All years	---	n = 116	• Satisfaction survey	Student Wellness Committee satisfaction (percentage, no numbers provided) • Positive experience with Student Wellness Committee: 95%	---
Fleming et al,[38] 2013	Single group cross-sectional or post-test only (for the outcome measure relevant to this review)	All years	---	n = 245	• Vanderbilt University student affairs survey	Reported that colleges design contributed meaningfully or somewhat meaningfully to their Vanderbilt University experience (percentage, no numbers provided): 91%	---
Real et al,[39] 2015	Single group cross-sectional or post-test only	All years	---	n = 450	• Maslach Burnout Inventory[a] • Primary Care Evaluation of Mental Disorders[a] • 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	--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---

Study	Design	Population	Comparison(s)	Sample Size(s)	Outcome Measures	Main Results	*P* Value
						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: $r = 0.1$	
Slavin et al,[40] 2014	Nonrandomized ≥ 2 group	First- and second-year medical students	Cohorts with different phases of multi-program reform implementation	n = 875–890	• Center for Epidemiological Studies Depression Scale[a] • Spielberger State-Trait Anxiety Inventory[a] • Perceived Stress Scale[a] • Perceived Cohesion Scale[a] • American Medical Colleges Graduation Questionnaire[a]	Moderate or severe depression according to the Center for Epidemiological Studies Depression Scale (percentage, no numbers provided) • End year 1: Cramér V = 0.16 • Phase 1 versus control: 21% (versus 27%) • Phase 1/2 versus control: 18% (versus 27%) • Phase 1/2/3 versus control: 11% (versus 27%) • End year 2: Cramér V = 0.18 • Phase 1 versus control: 17% (versus 32%) • Phase 1/2 versus control: 18% (versus 32%) • Phase 1/2/3 versus control: 16% (versus 32%) Anxiety (percentage, no numbers provided) • End year 1: Cramér V = 0.23 • Phase 1 versus control: 45%	NS NS <.05 < .05 < .05 < .0 NS < .05 < .05 NS < .05 < .05 NS < .05 < .05 < .05 < .05 NS < .05 <0.5 < .05 < .05 < .05

Study	Design	Population	Comparison(s)	Sample Size(s)	Outcome Measures	Main Results	*P* Value
						(versus 55%) • Phase 1/2 versus control: 31% (versus 55%) • Phase 1/2/3 versus control: 31% (versus 55%) • End year 2: Cramér V = 0.18 • Phase 1 versus control: 61% (versus 60%) • Phase 1/2 versus control: 39% (versus 60%) • Phase 1/2/3 versus control: 46% (versus 60%) Stress according (score) • End year 1: $\eta^2 = 0.06$ • Phase 1 versus control: 14.9 (SD 6.7) versus 16.3 (SD 7.4) • Phase 1/2 versus control: 13 (SD 6.8) versus 16.3 (SD 7.4) • Phase 1/2/3 versus control: 12.1 (SD 6.1) versus 16.3 (SD 7.4) • End year 2: partial η^2 = 0.05 • Phase 1 versus control: 14.4 (SD 5.8) versus 16.9 (SD 7.3) • Phase 1/2 versus control: 13.9 (SD 6.4) versus 16.9 (SD 7.3) • Phase 1/2/3 versus control: 13.5 (SD 6.8) versus 16.9 (SD 7.3) Cohesion (score) • End year 1: partial η^2 = 0.03 • Phase 1 versus control: 8.1 (SD 1.7) versus 7.9 (SD 2.1) • Phase 1/2 versus control: 8.5 (SD	

Study	Design	Population	Comparison(s)	Sample Size(s)	Outcome Measures	Main Results	*P* Value
						2.1) versus 7.9 (SD 2.1) • Phase 1/2/3 versus control: 8.8 (SD 1.8) versus 7.9 (SD 2.1) • End year 2: partial η^2 = 0.02 • Phase 1 versus control: 8.2 (SD 1.6) versus 7.7 (SD 2.0) • Phase 1/2 versus control: 8.5 (SD 2.2) versus 7.7 (SD 2.0) • Phase 1/2/3 versus control: 8.1 (SD 2.0) versus 7.7 (SD 2.0)	
Strayhorn,[41] 1989	Nonrandomized 2 group	First-year medical students	Cohort with multi-component program reform versus earlier cohort University of North Carolina, Chapel Hill versus comparison school	Responders n = 478 (original sample size not reported)	• Learning Environment Questionnaire[a] • Rand Health Insurance Questionnaires[a] • Environment stresses questionnaire • Social support questionnaire	New versus old curriculum stress questionnaire • Overall fewer stresses; $t(223) = -1.7$ • Less perceived stress from social and recreational sources • No reduction in financial-related stress New versus old curriculum mental well-being • Greater overall well-being $t(197) = -2.04$ • Greater sense of positive well-being • Greater sense of vitality • Less depression • Less anxiety • Social well-being $t(223) = -1.66$ New versus old curriculum social support • No perceived difference in availability of social supports $t(227) = -0.36$ • Less class advisor support • Class advisors less willing to listen • Class advisors less willing to help with personal problems • Concerned about students welfare	.09 .03 --- .04 < .001 < .001 < .001 < .001 .10 .721 .002 .003 < .001 .003 .05 .01 --- --- --- .01 .002 .02 .91

Study	Design	Population	Comparison(s)	Sample Size(s)	Outcome Measures	Main Results	*P* Value

						• Greater support from administrators
						• Could rely on administrators when things got tough
						• Perceived level of support from
						  • Fellow students
						  • Friends
						  • Significant others
						Time - control (University of North Carolina, Chapel Hill versus comparison) learning environment
						• Fewer environmental stressors $F(1467) = 6.41$
						• Greater mental well-being $F(1460) = 9.32$
						• Greater social well-being $F(1466) = 5.37$
						• No difference in social support $F(1477) = 0.01$	

Miscellaneous Wellness Programs							
Whitehouse et al,[42] 1996	Randomized clinical trial	First-year medical students	Self-hypnosis intervention versus 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: $F[3, 96] = 2.96$ At examination period self-hypnosis subjects significantly lower stressfulness scores: $t(30) = 2.11$	< .05 < .05
Goetzel et al,[44] 1984	Single group cross-sectional or post-test only	First-, second-, and third-year medical students	---	n = 26	• Group Environment Scale	Agreement with statement on scale (Likert 1–5): • “I am no longer as lonely; I feel more together with people”: 3.33 (of 5)	---
Lee and Graham,[45] 2001	Single group cross-sectional or post-test 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 (frequency) • 4/22 (18.2%) strongly agree • 17/22 (77.3%) agree Students felt that the Wellness Elective over-emphasized stress itself and devalued the worth of hard work; realistic expectations offered in this course seemed discouraging (frequency)	---

Study	Design	Population	Comparison(s)	Sample Size(s)	Outcome Measures	Main Results	*P* Value
						•1/22 (4.5%) agree	
Kushner et al,[46] 2011	Single group cross-sectional or post-test 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 (frequency) • 6/9 (66.7%) agree	---
Group-based Faculty Advisor/Mentor Programs							
Sastre et al,[47] 2010	Nonrandomized 2 group	First-, second-, and third-year medical students	Cohort with Advisory College Program versus earlier cohort with Faculty Advisory Program	n = 318 Cohort with 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 versus Faculty Advisory Program wellness advising (percentage, no numbers provided) • I feel comfortable discussing my personal stress with my advisor: 62% versus 24%; χ^2 = 40.9 • I feel comfortable discussing my mental health with my advisor: 51% versus 27%; χ^2 = 31.84 Satisfaction with how well advisors promoted wellness (percentage, no numbers provided: 27% versus 72%	< .001 < .001 <.001
Coates et al,[48] 2008	Nonrandomized 2 group	Fourth-year medical students	Cohort with mentoring program versus earlier cohort	n = 100 Cohort with program = 70 Earlier cohort = 30	• 25-item telephone survey	Cohort with mentoring program versus earlier cohort • Feels connected with faculty (frequency): 14/30 (47%) versus 49/70 (70%) • Feels connected with classmates (frequency): 11/30, (37%) versus 30/70 (43%)	--- ---
Ficklin et al,[49] 1983	Single group cross-sectional or post-test 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 the anxieties of starting school	--- --- ---

[a]Literature describing the development and validation of the various scales, scores, and questionnaires are as follows: Dupuy General Well-being Schedule,[16] Perceived Stress Scale,[17,66] Profile of Mood States,[30] Perceived Cohesion Scale,[18] Maslach Burnout Inventory,[67] Medical Outcomes Study Short Form,[68,69] Center for Epidemiologic Studies Depression Scale,[70] American Medical Colleges Graduation Questionnaire,[71] Patient Health Questionnaire,[72,73] Self-Compassion Scale,[25] Spielberger Trait Anxiety Inventory,[74] Distress Tolerance Scale,[28] Positive Affect Negative Affect Schedule,[75]

Symptom Checklist-90 Anxiety Subscale,[76] Perceived Stress of Medical School Scale,[77] 2-Item Depression Index,[78] Cognitive and Affective Mindfulness Scale- Revised,[79] Self-Regulation Questionnaire,[34] Jefferson Scale of Physician Empathy,[80] Zung Self-Rating Scale,[81] Primary Care Evaluation of Mental Disorders,[72,82] Learning Environment Questionnaire,[83] Rand Health Insurance Questionnaires,[84] Brief Symptom Inventory,[43] University of California, Los Angeles Loneliness Scale,[85] Group Environmental Scale.[86–88]

CI, confidence interval; OR, odds ratio; SD, standard deviation.

Table 2

Study methodology

Study	Number of Sites	Overall Sample Size	Women, No. (%)	Intervention	Evaluation	MERSQI	Aim
Pass/Fail Grading Systems							
Bloodgood et al,[14] 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 (A/B/C/D/F) 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 psychologic 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 (A/B/C/D/F) or 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,[19] 2011	Multisite	n = 2056 n = 1192 (responded)	550/1192 (47%)	Multisite survey of 2 different grading scales: 1) pass/fail and 2) 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,[20] 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

Study	Number of Sites	Overall Sample Size	Women, No. (%)	Intervention	Evaluation	MERSQI	Aim
Seritan et al,[21] 2013	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 preventative student wellness
Downs et al,[22] 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%)	4-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
Moutier et al,[23] 2012	Single site	n = 498 n = 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,[24] 2014	Single site	n = 58	26/58 (45.6%)	8-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,[26] 1985	Single site	n = 40	---	6 weekly meetings focused on teaching skills to reduce stress levels (progressive muscle relaxation, skills to help recognize and change maladaptive thoughts, and	Self-assessment	10	Help students acquire and develop skills to cope with stress

Study	Number of Sites	Overall Sample Size	Women, No. (%)	Intervention	Evaluation	MERSQI	Aim
				meditation techniques)			
Kraemer et al,[27] 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,[29] 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
Finkelstein et al,[31] 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,[32] 2015	Single site	n = 44	29/44 (65%)	4 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

Study	Number of Sites	Overall Sample Size	Women, No. (%)	Intervention	Evaluation	MERSQI	Aim
Bond et al,[33] 2013	Single site	n = 27	---	11-week *Embodied Health* course combining yoga meditation and neuroscience didactics	Self-assessment and semistructured interview	11.5	Evaluate the psychologic effects of an 11-week mind-body elective course
Curriculum Structure							
Reed et al,[19] 2011	See above in *Grading Systems* section						
Camp et al,[35] 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 learning or problem-based learning program
Kornitzer et al,[36] 2005	Single site	n = 92	---	6-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
Multi-component Program Reform							
Drolet and Rodgers,[37] 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	Survey	6.5	Evaluate a multicomponent wellness program

Study	Number of Sites	Overall Sample Size	Women, No. (%)	Intervention	Evaluation	MERSQI	Aim
				Committee)			
Fleming et al,[38] 2013	Single site	n = 245	---	Initiatives, activities, and resources including: 1) the 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) the student-led Student Wellness Committee focused on peer mentoring, social community, and mind/body wellness programming; 3) Vanderbilt Medical Students Careers in Medicine for career exploration, advising and planning, as well as residency application preparation; 4) VMS Live Program focused on the personal development of physicians-in-training; and 5) the 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
Real et al,[39] 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 which divides students into 1 of 4 colleges	Self-assessment and survey	10.5	The association of a wellness initiative on distress

Study	Number of Sites	Overall Sample Size	Women, No. (%)	Intervention	Evaluation	MERSQI	Aim
Slavin et al,[40] 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 consist with other courses	Self-assessment and survey	12	Discuss the utility and relevance of curricular changes and association with student mental health
Strayhorn,[41] 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,[42] 1996	Single site	n = 35	(60%)	Daily practice of self-hypnosis and diary records of sleep, mood, physical symptoms, and frequency of relaxation	Self-assessment	12	Determine the effectiveness of a self-hypnosis/relaxation intervention to

Study	Number of Sites	Overall Sample Size	Women, No. (%)	Intervention	Evaluation	MERSQI	Aim
				practice			relieve symptoms of psychologic distress and immune system reactivity to examination stress
Goetzel et al,[44] 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,[45] 2001	Single site	n = 66	40/60 (66%)	6-week wellness elective consisting of 1-hour lectures by physician presenters, discussions, and writing exercises	Survey	7	Explore students' perceptions of medical school stress and to assess their perspective on the wellness elective
Kushner et al,[46] 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,[47] 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 1-on-1 mentoring
Coates et al,[48] 2008	Single site	n = 100	---	Group-based mentoring program (the College Program) exclusively for fourth-year medical students, which divided students into academic	Survey	8	Change in fourth-year curriculum to include more mentors

Study	Number of Sites	Overall Sample Size	Women, No. (%)	Intervention	Evaluation	MERSQI	Aim
				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			
Ficklin et al,[49] 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

MERSQI, Medical Education Research Study Quality Instrument.