



Association Between Physician Depressive Symptoms and Medical Errors A Systematic Review and Meta-analysis

Karina Pereira-Lima, PhD; Douglas A. Mata, MD, MPH; Sonia R. Loureiro, PhD; José A. Crippa, MD, PhD; Lívia M. Bolsoni, MSc; Srijan Sen, MD, PhD

Abstract

IMPORTANCE Depression is highly prevalent among physicians and has been associated with increased risk of medical errors. However, questions regarding the magnitude and temporal direction of these associations remain open in recent literature.

OBJECTIVE To provide summary relative risk (RR) estimates for the associations between physician depressive symptoms and medical errors.

DATA SOURCES A systematic search of Embase, ERIC, PubMed, PsycINFO, Scopus, and Web of Science was performed from database inception to December 31, 2018.

STUDY SELECTION Peer-reviewed empirical studies that reported on a valid measure of physician depressive symptoms associated with perceived or observed medical errors were included. No language restrictions were applied.

DATA EXTRACTION AND SYNTHESIS Study characteristics and RR estimates were extracted from each article. Estimates were pooled using random-effects meta-analysis. Differences by study-level characteristics were estimated using subgroup meta-analysis and metaregression. The Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) guideline was followed.

MAIN OUTCOMES AND MEASURES Relative risk estimates for the associations between physician depressive symptoms and medical errors.

RESULTS In total, 11 studies involving 21 517 physicians were included. Data were extracted from 7 longitudinal studies (64%; with 5595 individuals) and 4 cross-sectional studies (36%; with 15 922 individuals). The overall RR for medical errors among physicians with a positive screening for depression was 1.95 (95% CI, 1.63-2.33), with high heterogeneity across the studies ($\chi^2 = 49.91$; $P < .001$; $I^2 = 82\%$; $\tau^2 = 0.06$). Among the variables assessed, study design explained the most heterogeneity across studies, with lower RR estimates associated with medical errors in longitudinal studies (RR, 1.62; 95% CI, 1.43-1.84; $\chi^2 = 5.77$; $P = .33$; $I^2 = 13\%$; $\tau^2 < 0.01$) and higher RR estimates in cross-sectional studies (RR, 2.51; 95% CI, 2.20-2.83; $\chi^2 = 5.44$; $P = .14$; $I^2 = 45\%$; $\tau^2 < 0.01$). Similar to the results for the meta-analysis of physician depressive symptoms associated with subsequent medical errors, the meta-analysis of 4 longitudinal studies (involving 4462 individuals) found that medical errors associated with subsequent depressive symptoms had a pooled RR of 1.67 (95% CI, 1.48-1.87; $\chi^2 = 1.85$; $P = .60$; $I^2 = 0\%$; $\tau^2 = 0$), suggesting that the association between physician depressive symptoms and medical errors is bidirectional.

CONCLUSIONS AND RELEVANCE Results of this study suggest that physicians with a positive screening for depressive symptoms are at higher risk for medical errors. Further research is needed

(continued)

Key Points

Question What are the magnitude and direction of associations between physician depressive symptoms and medical errors?

Findings In this systematic review and meta-analysis of 11 studies involving 21 517 physicians, physicians with a positive screening for depression were highly likely to report medical errors. Examination of longitudinal studies demonstrated that the association between physician depressive symptoms and medical errors is bidirectional.

Meaning This study found that physician depressive symptoms were associated with medical errors, highlighting the relevance of physician well-being to health care quality and underscoring the need for systematic efforts to prevent or reduce depressive symptoms among physicians.

+ Supplemental content

Author affiliations and article information are listed at the end of this article.

Open Access. This is an open access article distributed under the terms of the CC-BY License.

Abstract (continued)

to evaluate whether interventions to reduce physician depressive symptoms could play a role in mitigating medical errors and thus improving physician well-being and patient care.

JAMA Network Open. 2019;2(11):e1916097. doi:10.1001/jamanetworkopen.2019.16097

Introduction

Medical errors are a major source of patient harm. Studies estimate that, in the United States, as many as 98 000 to 251 000 hospitalized patients die each year as result of a preventable adverse event.¹⁻⁴ In addition, medical errors are a major source of morbidity⁵ and account for billions of dollars in financial losses to health care systems every year.⁶⁻⁹

Depressive symptoms are highly prevalent among physicians,^{10,11} and several studies have investigated the associations between physician depressive symptoms and medical errors.¹²⁻¹⁶ Although most studies on physician depressive symptoms and medical errors have identified a substantial association, their results are not unanimous, and questions regarding the direction of these associations remain open in recent literature.¹⁷

Depressive symptoms have well-established clinical criteria, and a large body of work has demonstrated that depression is a preventable and treatable condition.¹⁸⁻²⁰ Several studies with physicians have identified potential individual and work environment sources of interventions to prevent the development of depressive symptoms among these professionals,²¹⁻²⁴ and although scarce, research on the efficacy of interventions to reduce depressive symptoms in physicians has shown positive results.²⁵

Given that depression is preventable and treatable, a reliable estimate of the degree to which physicians with a positive screening for depression are at higher risk for medical errors would be useful. Such an estimate would inform public health decision-making on strategies to improve patient safety and physician well-being. In this systematic review and meta-analysis, we investigated whether physician depressive symptoms were associated with medical errors. We also examined longitudinal studies to investigate the temporal associations between depressive symptoms and medical errors.

Methods

Search Strategy and Study Eligibility

Two of us (K.P.-L. and L.M.B.) independently identified cross-sectional and longitudinal studies published before December 31, 2018, that reported on the associations between physician depressive symptoms and perceived or objectively assessed medical errors. We systematically searched Embase, ERIC, PubMed, PsycINFO, Scopus, and Web of Science. In addition, guided by the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA),²⁶ we screened the reference lists of the articles and corresponded with study investigators. The search strategy we used was initially designed by the corresponding author (K.P.-L.), and critical revisions and edits to this design were provided by a multiprofessional team of researchers with expertise in conducting systematic reviews and meta-analyses on physician depression (D.A.M., S.S.) and mental health (S.R.L., J.A.C., S.S.) research. The Ribeirão Preto Medical School Institutional Review Board deemed this study exempt from approval and informed consent because it collected and synthesized nonidentifiable data from previously published studies.

For the database searches, terms related to physicians and depressive symptoms were combined with terms related to medical errors, without language restriction; full details of the search strategy are provided in the eMethods in the [Supplement](#). References identified from database searches were exported to EndNote (Clarivate Analytics). After removal of duplicates, full-text

articles were obtained if their abstracts were considered to be eligible by at least 1 of us. Each full-text article was assessed independently for final inclusion in this systematic review and meta-analysis, and disagreements were resolved by consensus (we reached 97% overall agreement [113 of 116 articles; $\kappa = 0.87$]). Peer-reviewed studies that reported data on perceived or observed medical errors associated with a valid measure of depressive symptoms in practicing and resident physicians (ie, excluding medical students and other health care professionals) were included. Studies that involved both physicians and other health care professionals were included only if they provided separate data for physicians. To be included, studies did not have to consider the association between physician depressive symptoms and medical errors as their primary outcome of interest.

Data Extraction and Quality Assessment

Two of us (K.P.-L., L.M.B.) independently extracted the following data from each article using a standardized study form: (1) study information, including geographic location, survey years, research design, sample size, percentage of respondents among eligible participants, and number of institutions included; (2) characteristics of participants, including mean age, percentage of women, specialties, and career level; and (3) outcomes, including depressive symptoms measure, medical errors question interval, method of medical errors assessment, and data for calculating effect size (eg, relative risk [RR], CIs, *P* values). The approach recommended by Zhang and Yu²⁷ for converting adjusted odds ratio for RR was used for studies that reported only the results of logistic regression for the associations between physician depressive symptoms and medical errors. Corresponding authors were contacted at least twice when studies did not report enough data to compute the effect size. When studies involved the same population of physicians, only the most comprehensive articles (ie, including those with a greater number of participants or a longer follow-up period) were included.

The methodological quality of the studies was assessed using adapted criteria from the Cochrane Library guidelines.²⁸ Studies were considered methodologically strong or weak on the basis of (1) study design (eg, longitudinal indicated strong; cross-sectional, weak), (2) sample size (≥ 200 participants indicated strong; < 200 participants, weak), (3) ascertainment of depressive symptoms measure (sensitivity and specificity $> 75\%$ indicated strong; sensitivity and specificity $\leq 75\%$, weak), (4) representativeness of the sample (≥ 2 institutions indicated strong; < 2 institutions, weak), and (5) descriptive characteristics of participants (reported data on sex, age, specialties, and career level indicated strong; missing information on sex, age, specialties, or career level, weak). Cutoff scores for sample size, representativeness, and descriptive characteristics were based on thresholds used in previous meta-analyses on physician depression,^{10,11} whereas cutoff scores for ascertainment of depressive symptoms were based on well-established psychometric quality criteria for depression questionnaires.²⁹ Disagreements regarding quality assessment scores for each individual study were resolved by consensus (with an overall agreement of 98%; $\kappa = 0.96$).

Statistical Analysis

Relative risk estimates of physician depressive symptoms associated with medical errors were calculated by pooling study-specific estimates using random-effects models with generic invariance method to incorporate the heterogeneity of the differences across the studies.

Between-study heterogeneity was measured using standard χ^2 tests and I^2 statistics (values $< 25\%$ indicate low; 25%-75%, moderate; and $> 75\%$, considerable heterogeneity).^{30,31} Sensitivity analyses were performed by serially excluding each study to determine the implications of individual studies for the pooled RR estimates.

Results from studies grouped according to prespecified study-level characteristics were compared using stratified meta-analysis (for physician career level, specialties included, medical errors question interval, geographic region, depressive symptoms measure, and quality assessment indicators [ie, study design, sample size, ascertainment of the depressive symptoms measure, representativeness of the sample, and descriptive data]) or random-effects metaregression (for year of baseline survey and percentage of women).^{32,33} To gain insight into the direction of the

association between depressive symptoms and medical errors, we calculated pooled RR estimates for longitudinal studies that reported (1) results of physician depressive symptoms associated with subsequent medical errors and (2) RR estimates of medical errors associated with subsequent physician depressive symptoms.

Bias secondary to small study effects was investigated using funnel plots and the Egger test.^{34,35} We used R, version 3.2.3 (R Project for Statistical Computing),³⁶ with meta³⁷ and metafor³⁸ packages for all analyses. Statistical tests were 2-sided and used a significance threshold of $P < .05$.

Results

Study Characteristics

Eleven studies involving a total of 21 517 physicians were included in this systematic review and meta-analysis (**Figure 1**). The characteristics of the included studies are summarized in the **Table**. A total of 7 studies (64%) were longitudinal (involving 5595 individuals)^{12-15,39,40,44} and 4 (36%) were cross-sectional (involving 15 922 individuals).^{16,41-43} Nine studies (82%) took place in the United States,^{12-16,40,42-44} 1 (9%) in Japan,³⁹ and 1 (9%) in South Korea.⁴¹ Eight studies (73%) included only training physicians (interns and/or residents),^{12-16,40,41,44} and 3 (27%) recruited physicians from any career level.^{39,42,43} Seven studies (64%) recruited physicians from multiple specialties,^{14,15,39-43} whereas 4 (36%) recruited physicians from a single specialty.^{12,13,16,44} Among these 4 studies, 1 focused on pediatric residents,¹² 1 on anesthesiology residents,¹⁶ and 2 on internal medicine residents.^{13,44} The median (interquartile range [IQR]) number of participants per study was 836 (2139). Five studies (46%) assessed depressive symptoms with the 2-item Primary Care Evaluation of Mental Disorders (PRIME-MD-2) questionnaire^{13,41-44}; 3 (27%) used the 9-item Patient Health Questionnaire (PHQ-9)^{14,15,40}; 2 (18%) used the Harvard National Depression Screening Day Scale (HANDS)^{12,16}; and 1 (9%) used the 5-item World Health Organization Well-being Index (WHO-5).³⁹ Sensitivity and specificity commonly reported for these depression instruments are available in eTable 1 in the **Supplement**.

All but 1 study¹² (9%) used self-report measures of medical errors. Eight studies (73%) inquired about medical errors in the past 3 months,^{13-15,40-44} 2 (18%) inquired about medical errors in the past year,^{16,39} and 1 (9%) actively surveyed medical errors in a 1-month interval.¹² Assessment measures and definitions of medical errors adopted by individual studies are available in eTable 2 in the **Supplement**. Although most studies inquired about major or harmful medical errors,^{13-16,39,40,42-44} 1 study (9%) inquired whether physicians were concerned about errors of any type,⁴¹ and 1 study (9%) trained a team of nurses and physicians to collect daily reports of all medication errors occurring on

Figure 1. PRISMA Flow Diagram

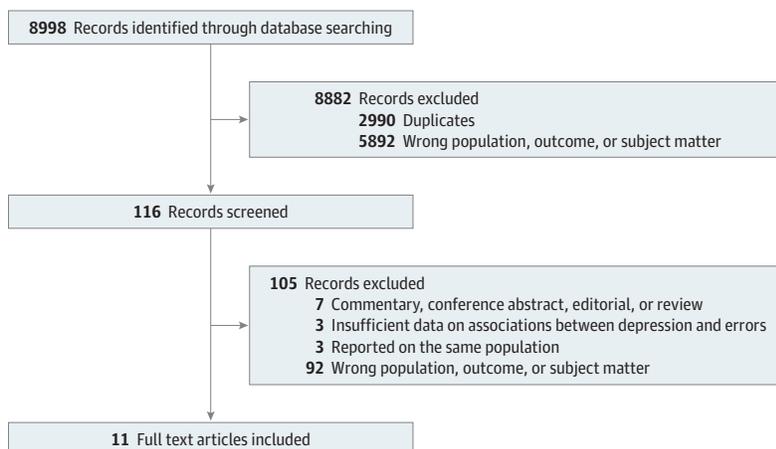


Table. Selected Characteristics of the 11 Included Studies

Source	Country	Baseline Year	Specialties	Participants, No. (%)	Physician Career Level	Age, y (%)	Women, No. (%)	Study Design	Depression Measure (Cutoff Score)	Depression Assessment	Errors	
											Measure	Assessment
Fahrenkopf et al, ¹² 2008	United States	2003	Pediatrics, medicine-pediatrics	101 (50.0) ^a	Training physicians	<30 (61.8)	69.9	Longitudinal	HANDS (≥9)	Baseline	Active surveillance	Daily review during 1 mo
Havashino et al, ³⁹ 2012	Japan	2009	Several	836 (69.8)	Any career level	<39 (22.9)	7.9	Longitudinal	WHO-5 (<13)	Baseline	Self-reported	1 y after baseline
Kalmbach et al, ⁴⁰ 2017	United States	2012	Several	1215 (58.0)	Training physicians	Mean (SD): 27.5 (2.7)	48.9	Longitudinal	PHQ-9 (≥10)	Third and sixth mo of internship ^b	Self-reported	Third and sixth mo of internship
Kang et al, ⁴¹ 2013	South Korea	2010	Several	86 (58.5)	Training physicians	<31 (77.9)	25.6	Cross-sectional	PRIME-MD-2 (yes to either item)	Single measure	Self-reported	Single measure
de Oliveira et al, ¹⁶ 2013	United States	2011	Anesthesiology	1345 (54.4) ^c	Training physicians	<31 (53.8)	43.0	Cross-sectional	HANDS (≥9)	Single measure	Self-reported	Single measure
Sen et al, ¹⁴ 2010	United States	2009	Several	740 (58.2)	Training physicians	<31 (85.7)	54.5	Longitudinal	PHQ-9 (≥10)	Preinternship + third, sixth, ninth, and twelfth mo of internship	Self-reported	Third, sixth, ninth, and twelfth mo of internship
Sen et al, ¹⁵ 2013	United States	2007	Several	2323 (58.0)	Training physicians	Mean (SD): 27.5 (3.0)	50.9	Longitudinal	PHQ-9 (≥10)	Preinternship + third, sixth, ninth, and twelfth mo of internship	Self-reported	Third, sixth, ninth, and twelfth mo of internship
Shanafelt et al, ⁴² 2010	United States	2008	Surgical	7905 (32.0)	Any career level	Median (IQR): 51 (43-59)	13.3	Cross-sectional	PRIME-MD-2 (yes to either item)	Single measure	Self-reported	Single measure
Tawfik et al, ⁴³ 2018	United States	2014	Several	6586 (19.2)	Any career level	Median (IQR): 56 (45-63)	32.9	Cross-sectional	PRIME-MD-2 (yes to either item)	Single measure	Self-reported	Single measure
West et al, ¹³ 2009	United States	2003	Internal medicine	380 (88.4)	Training physicians	<31 (63.2)	37.9	Longitudinal	PRIME-MD-2 (yes to either item)	Every 6 mo from residency onset to completion	Self-reported	Every 3 mo from residency onset to completion
West et al, ⁴⁴ 2006 ^d	United States	2003	Internal medicine	184 (84.0) ^d	Training physicians	<31 (70.1)	35.9	Longitudinal	PRIME-MD-2 (yes to either item)	Every 6 mo from residency onset to completion	Self-reported	Every 3 mo from residency onset to completion

Abbreviations: HANDS, Harvard National Depression Screening Day Scale; IQR, interquartile range; PHQ-9, 9-item Patient Health Questionnaire; PRIME-MD-2, 2-item Primary Care Evaluation of Mental Disorders; WHO-5, 5-item World Health Organization Well-being Index.

^a Number of participants included in the active surveillance of medical errors.
^b A baseline assessment of depressive symptoms was performed to exclude physicians with a positive screening for depression before internship onset.

^c Responses to the question, "I make mistakes with negative consequences to patients."

^d Included only in the meta-analysis of medical errors associated with depressive symptoms. A more recent publication with a more comprehensive population (West et al¹³) reported on depressive symptoms associated with medical errors.

wards and to actively review all medical records and medication orders using structured data forms.¹² When evaluated by the established quality assessment criteria, 6 studies (55%) were considered as methodologically strong on the basis of design^{12-15,39,40,44}; 8 (73%), on the basis of sample size^{13-16,39,40,42,43}; 5 (46%), on the basis of ascertainment of depressive symptoms measure^{12,14-16,40}; 8 (73%), on the basis of representativeness of the sample^{12,14-16,39,40,42,43}; and all, on the basis of descriptive characteristics of participants.^{12-16,39-44} Detailed quality indicators for each study are available in eTable 3 in the Supplement.

Of the 11 included studies, 1 (9%) was used only in the meta-analysis of medical errors associated with subsequent depressive symptoms.⁴⁴ The reason for excluding this study from the other analyses is that a more recent article reported data on depressive symptoms associated with subsequent medical errors in a more comprehensive sample of physicians.¹³ Because the more recent study did not report data on medical errors associated with subsequent depressive symptoms, the previous study was included in this directionality meta-analysis and excluded from all other analyses to avoid overlapping data. The approach recommended by Zhang and Yu²⁷ was used for computing RR estimates in 2 studies that reported associations of depressive symptoms and medical errors in the format of an odds ratio.^{13,44}

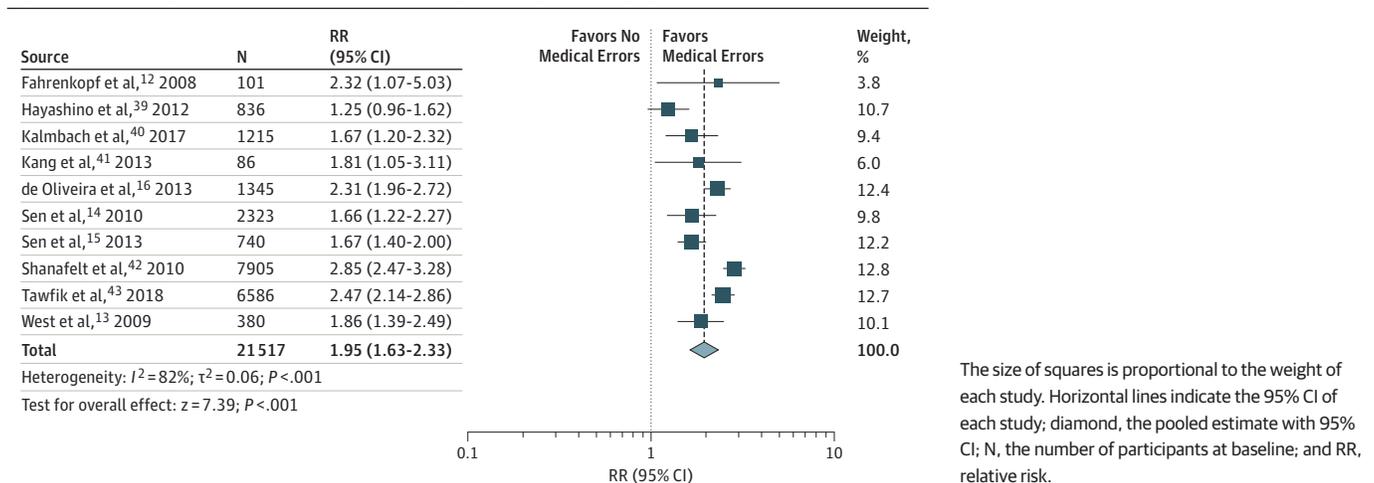
Associations Between Depressive Symptoms and Medical Errors

Meta-analytic pooling of the associations between depressive symptoms and medical errors yielded a summary RR of 1.95 (95% CI, 1.63-2.33), with high heterogeneity across the studies ($\chi^2 = 49.91$; $P < .001$; $I^2 = 82\%$; $\tau^2 = 0.06$) (Figure 2). The sensitivity analysis, in which the meta-analysis was serially repeated after exclusion of each study, demonstrated that no individual study had an implication for the overall RR estimate of more than 0.12 points (these estimates varied from 1.85 [95% CI, 1.56-2.19] to 2.07 [95% CI, 1.77-2.43]) (eFigure 1 in the Supplement).

Direction of the Associations

All of the 7 longitudinal studies included in the present review investigated the association of physician depressive symptoms in the next 1,¹² 3,^{13-15,40,44} or 12 months.³⁹ One study⁴⁴ was removed from the first directionality analysis because a later publication, which included a more comprehensive sample, also reported on data regarding depressive symptoms associated with subsequent medical errors.¹³ Meta-analytic pooling of physician depression associated with medical errors resulted in a pooled RR of 1.62 (95% CI, 1.43-1.84), with low heterogeneity across studies ($\chi^2 = 5.77$; $P = .33$; $I^2 = 13\%$; $\tau^2 < 0.01$) (Figure 3).

Figure 2. Meta-analysis of the Association Between Physician Depressive Symptoms and Medical Errors



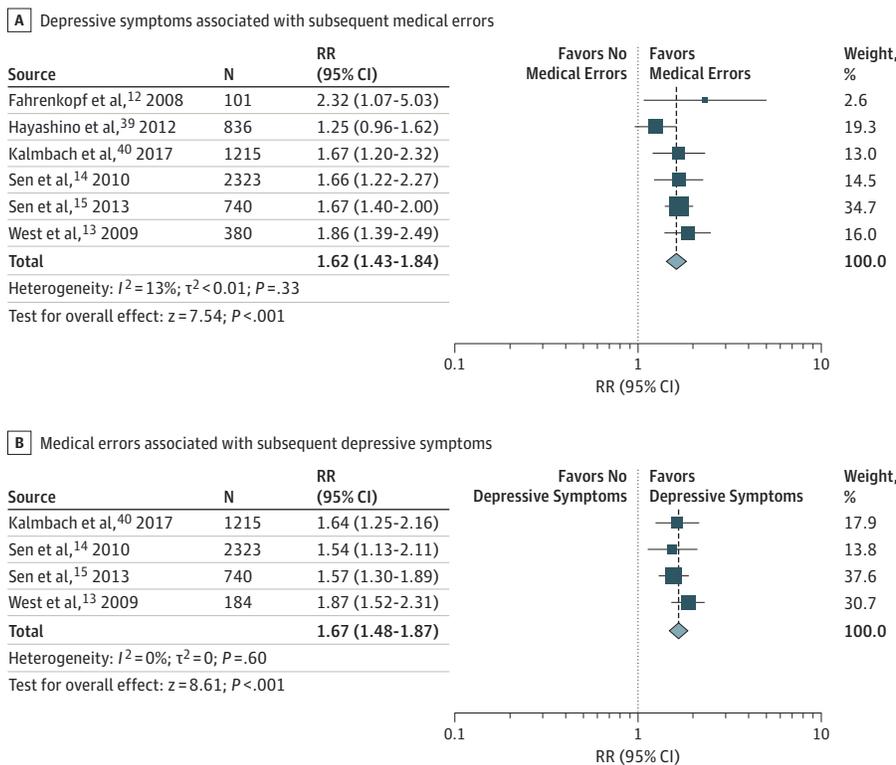
Similarly, 4 of the 7 longitudinal studies provided data on medical errors associated with depressive symptoms in the next 3 months.^{14,15,40,44} Meta-analytic pooling of these 4 studies (involving 4462 physicians) resulted in a summary RR of 1.67 (95% CI, 1.48-1.87), with low heterogeneity across studies ($\chi^2 = 1.85$; $P = .60$; $I^2 = 0\%$; $\tau^2 = 0$), suggesting that the association between physician depression and medical errors is bidirectional (Figure 3).

Associations Stratified by Study-Level Characteristics

To identify potential sources of heterogeneity, we performed subgroup meta-analysis of studies stratified by different study-level characteristics when at least 2 studies were available in each comparator subgroup. Studies with exclusively surgical specialties yielded a summary RR estimate that was significantly higher than the summary RR estimate in studies that also included nonsurgical specialties (2.59 [95% CI, 2.10-3.16] vs 1.79 [95% CI, 1.46-3.16]). Furthermore, US studies yielded higher estimates of the association between depression and medical errors compared with non-US studies (2.10 [95% CI, 1.77-2.46] vs 1.39 [95% CI, 1.00-1.93]). Summary RR estimates for studies assessing depressive symptoms through the HANDS or the PRIME-MD-2 were significantly higher compared with the ones identified through the PHQ-9 (HANDS: 2.32 [95% CI, 1.97-2.72]; PRIME-MD-2: 2.39 [95% CI, 1.97-2.86]; PHQ-9: 1.67 [95% CI, 1.45-1.92]) (eFigure 2 in the Supplement). No statistically significant differences in RR estimates were found between subgroups of studies stratified by physician career level or studies inquiring physicians about medical errors in the past 3 or 12 months.

A single study assessed depressive symptoms associated with medication errors actively surveyed in the next month.¹² The sensitivity analysis that excluded this study did not show a significant reduction in heterogeneity statistics (from 1.95; 95% CI, 1.63-2.33; $\chi^2 = 49.91$; $P < .001$; $I^2 = 82\%$; $\tau^2 = 0.06$ to 1.94; 95% CI, 1.61-2.33; $\chi^2 = 49.88$; $P < .001$; $I^2 = 84\%$; $\tau^2 = 0.06$). In contrast,

Figure 3. Meta-analyses of Long-term Studies of the Association Between Physician Depressive Symptoms and Medical Errors



the sensitivity analysis that excluded the only study³⁹ that used the WHO-5 to assess physician depressive symptoms resulted in a reduction in all heterogeneity statistics (from 1.95; 95% CI, 1.63-2.33; $\chi^2 = 49.91$; $P < .001$; $I^2 = 82\%$; $\tau^2 = 0.06$ to 2.07; 95% CI, 1.77-2.43; $\chi^2 = 31.91$; $P < .001$; $I^2 = 75\%$; $\tau^2 = 0.04$) (eFigure 1 in the Supplement). Metaregression results revealed that RR estimates did not significantly vary with baseline survey year (estimate = 0.01; 95% CI, -0.05 to 0.07; QM [statistic for the test of moderators] = 0.14; $P = .71$) or percentage of female physicians (estimate = -0.06; 95% CI, -1.13 to 1.00; QM = 0.01; $P = .91$) (eFigure 3 in the Supplement).

When evaluated by the quality assessment indicators, longitudinal studies yielded summary RR estimates that were significantly lower compared with those from the cross-sectional studies (1.62; 95% CI, 1.43-1.84; $\chi^2 = 5.77$; $P = .33$; $I^2 = 13\%$; $\tau^2 < 0.01$ vs 2.51; 95% CI, 2.20-2.83; $\chi^2 = 5.44$; $P = .14$; $I^2 = 45\%$; $\tau^2 < 0.01$). No statistically significant differences in RR estimates were found between subgroups of studies stratified by sample size, ascertainment of the depression measure, representativeness of the sample, or descriptive characteristics of the participants (eFigure 4 in the Supplement).

Assessment of Publication Bias

A funnel plot of studies that reported on physician depressive symptoms associated with medical errors is presented in eFigure 5 in the Supplement). The Egger test indicated the absence of significant publication bias (intercept = -2.79; $P = .12$).

Discussion

This systematic review and meta-analysis of 11 studies involving 21 517 physicians demonstrated an association between physician depressive symptoms and an increased risk for perceived medical errors (RR, 1.95; 95% CI, 1.63-2.33). We also found that the magnitude of the associations of physician depressive symptoms and perceived medical errors were relatively consistent across studies that assessed training and practicing physicians, providing additional evidence that physician depression has implications for the quality of care delivered by physicians at different career stages.

Subgroup meta-analysis of studies stratified by different study-level characteristics identified study design, specialty type, geographic region, and depressive symptoms measure as possible sources of heterogeneity in this meta-analysis. The 6 longitudinal studies that assessed physician depressive symptoms associated with subsequent medical errors yielded a significantly lower summary RR estimate compared with the 4 cross-sectional studies included in this meta-analysis (1.62 [95% CI, 1.43-1.84] vs 2.51 [95% CI, 2.20-2.83]), but a significant increased risk for medical errors among physicians with depressive symptoms was identified in both study designs.

Similarly, although the summary RR estimates for studies that included nonsurgical specialties, that were from non-US countries, and that used the PHQ-9 as a measure of depressive symptoms were significantly lower than the summary RR estimates identified for their reference subgroups, the estimates were still statistically significant for all analyzed subgroups. These results support the main finding that depressive symptoms are associated with an increased risk for medical errors among physicians.

In line with these results, sensitivity analysis demonstrated that no individual study was associated with the overall RR estimate by more than 0.12 points (overall RR estimates in sensitivity analysis varied from 1.85 [95% CI, 1.56-2.19] to 2.07 [95% CI, 1.77-2.43]). The study that accounted for the largest variation in the magnitude of RR estimates (from 1.95 [95% CI, 1.63-2.33] to 2.07 [95% CI, 1.77-2.43]) used the WHO-5 for the ascertainment of depression in Japanese physicians.³⁹ The WHO-5 was originally designed as a measure of subjective well-being and has been validated as a depression screening instrument.⁴⁵ Studies conducted in primary care settings have suggested that the WHO-5's broad statements tend to favor sensitivity at the cost of specificity when screening for depression in the general population,⁴⁶⁻⁴⁹ which might have been a source of heterogeneity in the present study.

A previous meta-analysis has associated physician burnout and emotional distress with patient safety outcomes.⁵⁰ The present meta-analysis advances the findings of this past work in different ways. First, the issue of quantifying heterogeneous constructs of emotional distress in the same meta-analysis was overcome by focusing on depressive symptoms, which have well-established clinical criteria and methods of assessment.^{20,51} Similarly, by working with RR instead of odds ratio estimates, we were able to more accurately estimate the magnitude of the association between depressive symptoms and perceived medical errors.^{52,53} Furthermore, the analysis of 7 longitudinal studies^{12-15,39,40,44} allowed us to demonstrate that physician depressive symptoms are associated with future medical errors (RR, 1.62; 95% CI, 1.43-1.84; $n = 5595$ physicians from 6 studies^{12-15,39,40}) and that medical errors are associated with future depressive symptoms in physicians (RR, 1.67; 95% CI, 1.48-1.87; $n = 4462$ physicians from 4 studies^{14,15,40,44}). Taken together, these data suggest that the association between physician depression and medical errors is bidirectional. To our knowledge, this study is the first to systematically review the direction of the associations between physician depressive symptoms and medical errors.

Studies have recommended the addition of physician well-being to the Triple Aim of enhancing the patient experience of care, improving the health of populations, and reducing the per capita cost of health care.⁵⁴⁻⁵⁷ Results of the present study endorse the Quadruple Aim movement by demonstrating not only that medical errors are associated with physician health but also that physician depressive symptoms are associated with subsequent errors. Given that few physicians with depression seek treatment^{58,59} and that recent evidence has pointed to the lack of organizational interventions aimed at reducing physician depressive symptoms,²⁵ our findings underscore the need for institutional policies to remove barriers to the delivery of evidence-based treatment to physicians with depression. Investments in patient safety have been associated with significant reductions in health care costs,⁶⁰ and the bidirectional associations between physician depressive symptoms and perceived medical errors verified by this meta-analysis suggest that physician well-being is critical to patient safety. Further studies are needed to explore these associations. Such research should investigate whether systematic interventions for reducing depressive symptoms could be factors in decreased medical errors.

Limitations

This systematic review and meta-analysis has some limitations. First, 10 of 11 studies included relied on self-report measures of medical errors.^{13-16,39-44} Although substantial differences in RR estimates and heterogeneity statistics were not identified by sensitivity analysis that removed the only study that assessed medical errors through active surveillance,¹² the small sample size of the referred study limited its weight in the overall meta-analysis. Furthermore, although self-reported errors have been found to be highly correlated with recorded events,⁶¹ the self-report nature of the included studies may have introduced bias to the present results. For instance, physicians with depression may be more likely to perceive medical errors, which may drive the association between depressive symptoms and medical errors. However, the secondary meta-analyses of longitudinal studies that assessed depressive symptoms associated with subsequent medical errors and medical errors associated with future depressive symptoms demonstrated significantly increased risk estimates, which suggests the existence of bidirectional temporal associations between physician depressive symptoms and perceived medical errors. Similarly, all included studies examined and ascertained depressive symptoms from self-report inventories that varied in sensitivity and specificity. Therefore, the results demonstrated the presence of associations between depressive symptoms and perceived medical errors rather than the association between a clinical diagnosis of depression and medical errors.

Second, the 10 studies that evaluated self-reported medical errors included general questions about either major,^{13-15,39,40,42-44} harmful,¹⁶ or any⁴¹ medical errors. By doing so, these studies might have underestimated particular acts and omissions with potential to harm that physicians might not have considered to be a major, harmful, or any medical error. In the only study that assessed errors

through active surveillance, more than 60% of the observed medical errors were considered to be potentially harmful,¹² which suggests that a large portion of medical errors committed by physicians could have negative consequences for patients.

Third, the small number of studies included in some of the subgroups may have biased some of the subgroup analysis results.⁶² Fourth, despite the significant overall effect of the meta-analytic model of medical errors associated with subsequent depressive symptoms, few studies (4 studies with 4462 physicians)^{14,15,40,44} were included in this directional analysis, which might also have introduced bias to the results. Fifth, most studies (9 of 11) assessed US physicians.^{12-16,40,42-44} Therefore, the results may not be generalizable to physicians in other countries.

Sixth, although the 3 studies that evaluated both practicing and training physicians included the largest number of physicians in this meta-analysis (15 327 of 21 517),^{39,42,43} most of the included studies (8 of 11) exclusively assessed populations of training physicians.^{12-16,40,41,44} Although the subgroup meta-analysis that stratified studies by physician career level did not identify significant differences between the 2 subgroups, generalizations of the present study results to populations of practicing physicians should be done with caution. Seventh, all references included were from full-text articles published in peer-reviewed journals. Although no evidence of publication bias was verified by Egger test, the exclusion of unpublished data and gray literature might have introduced selection bias to this analysis.

Conclusions

By combining data from multiple studies, this systematic review and meta-analysis found that physician depressive symptoms were associated with increased risk for perceived medical errors and that the association between depressive symptoms and perceived errors was bidirectional. Future research is needed to evaluate the associations of physician depressive symptoms with objective measures of medical errors, such as active surveillance. Studies that include physicians from different countries could answer whether cultural and socioeconomic aspects play a role in the associations between depressive symptoms and errors. Future research is also needed into the degree to which interventions for reducing physician depressive symptoms could mitigate medical errors and improve physician well-being and patient care.

ARTICLE INFORMATION

Accepted for Publication: October 4, 2019.

Published: November 27, 2019. doi:[10.1001/jamanetworkopen.2019.16097](https://doi.org/10.1001/jamanetworkopen.2019.16097)

Open Access: This is an open access article distributed under the terms of the [CC-BY License](https://creativecommons.org/licenses/by/4.0/). © 2019 Pereira-Lima K et al. *JAMA Network Open*.

Corresponding Author: Karina Pereira-Lima, PhD, Department of Psychiatry, University of Michigan Medical School, 205 Zina Pitcher Pl, Ann Arbor, MI 48109-5720 (pereiralima.k@gmail.com).

Author Affiliations: Department of Psychiatry, University of Michigan Medical School, Ann Arbor (Pereira-Lima, Sen); Department of Psychiatry, Federal University of São Paulo, São Paulo, Brazil (Pereira-Lima); Department of Pathology, Memorial Sloan Kettering Cancer Center, New York, New York (Mata); Department of Neuroscience and Behavior, University of São Paulo, Ribeirão Preto, São Paulo, Brazil (Loureiro, Crippa, Bolsoni).

Author Contributions: Dr Pereira-Lima 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: Pereira-Lima, Mata, Loureiro, Crippa, Sen.

Acquisition, analysis, or interpretation of data: Pereira-Lima, Loureiro, Crippa, Bolsoni.

Drafting of the manuscript: Pereira-Lima.

Critical revision of the manuscript for important intellectual content: All authors.

Statistical analysis: Pereira-Lima.

Obtained funding: Pereira-Lima, Loureiro, Sen.

Administrative, technical, or material support: Crippa, Bolsoni.

Supervision: Mata, Loureiro, Crippa, Sen.

Conflict of Interest Disclosures: Dr Crippa reported receiving grants from CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico) and CNPq/FAPESP (Fundação de Amparo à Pesquisa do Estado de São Paulo), other (cannabidiol supply) from BSPG Pharm, and grants from Prati-Donaduzzi Pharmaceutical and University Global Partnership Network outside the submitted work, and reported holding a patent (Mechoulam R, Crippa JA, Guimaraes FS, Zuardi A, Hallak, JE, and Breuer A, inventors. Fluorinated CBD compounds, compositions and uses thereof. Pub. No: WO/2014/108899; International Application No: PCT/IL2014/050023; Def. US no. Reg. 62193296; July 29, 2015; INPI on August 19, 2015 [BR1120150164927] issued; University of São Paulo has an agreement with Prati-Donaduzzi [Toledo, Brazil] to "develop a pharmaceutical product containing synthetic CBD and prove its safety and therapeutic efficacy in the treatment of epilepsy, schizophrenia, Parkinson's disease, and anxiety disorders."). No other disclosures were reported.

Funding/Support: This study was funded in part by the Arnold P. Gold Foundation Research Institute (Drs Pereira-Lima, Mata, Loureiro, and Sen). Dr Pereira-Lima was supported by research fellowship grants 2016/13410-O and 2018/21480-4 from FAPESP. Drs Loureiro and Crippa were supported by a productivity fellowship grant from the CNPq. Ms Bolsoni was supported by research fellowship award grant 2016/01801-5 from FAPESP. Dr Sen was supported by grant RO1MH101459 from the National Institutes of Health.

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; and decision to submit the manuscript for publication.

Disclaimer: The opinions, results, and conclusions herein are those of the authors and do not reflect the official policy or position of the funding sources.

REFERENCES

1. Institute of Medicine Committee on Quality of Health Care in America; Kohn LT, Corrigan JM, Donaldson MS, eds. *To Err Is Human: Building a Safer Health System*. Vol 6. Washington, DC: National Academies Press; 2000.
2. James JT. A new, evidence-based estimate of patient harms associated with hospital care. *J Patient Saf*. 2013;9(3):122-128. doi:10.1097/PTS.0b013e3182948a69
3. Makary MA, Daniel M. Medical error—the third leading cause of death in the US. *BMJ*. 2016;353:i2139. doi:10.1136/bmj.i2139
4. Weingart SN, Wilson RM, Gibberd RW, Harrison B. Epidemiology of medical error. *BMJ*. 2000;320(7237):774-777. doi:10.1136/bmj.320.7237.774
5. Jha AK, Larizgoitia I, Audera-Lopez C, Prasopa-Plaizier N, Waters H, Bates DW. The global burden of unsafe medical care: analytic modelling of observational studies. *BMJ Qual Saf*. 2013;22(10):809-815. doi:10.1136/bmjqs-2012-001748
6. Johnson WG, Brennan TA, Newhouse JP, et al. The economic consequences of medical injuries: implications for a no-fault insurance plan. *JAMA*. 1992;267(18):2487-2492. doi:10.1001/jama.1992.03480180073032
7. Bates DW, Spell N, Cullen DJ, et al; Adverse Drug Events Prevention Study Group. The costs of adverse drug events in hospitalized patients. *JAMA*. 1997;277(4):307-311. doi:10.1001/jama.1997.03540280045032
8. Thomas EJ, Studdert DM, Newhouse JP, et al. Costs of medical injuries in Utah and Colorado. *Inquiry*. 1999;36(3):255-264.
9. Van Den Bos J, Rustagi K, Gray T, Halford M, Ziemkiewicz E, Shreve J. The \$17.1 billion problem: the annual cost of measurable medical errors. *Health Aff (Millwood)*. 2011;30(4):596-603. doi:10.1377/hlthaff.2011.0084
10. Mata DA, Ramos MA, Bansal N, et al. Prevalence of depression and depressive symptoms among resident physicians: a systematic review and meta-analysis. *JAMA*. 2015;314(22):2373-2383. doi:10.1001/jama.2015.15845
11. Rotenstein LS, Ramos MA, Torre M, et al. Prevalence of depression, depressive symptoms, and suicidal ideation among medical students: a systematic review and meta-analysis. *JAMA*. 2016;316(21):2214-2236. doi:10.1001/jama.2016.17324
12. Fahrenkopf AM, Sectish TC, Barger LK, et al. Rates of medication errors among depressed and burnt out residents: prospective cohort study. *BMJ*. 2008;336(7642):488-491. doi:10.1136/bmj.39469.763218.BE
13. West CP, Tan AD, Habermann TM, Sloan JA, Shanafelt TD. Association of resident fatigue and distress with perceived medical errors. *JAMA*. 2009;302(12):1294-1300. doi:10.1001/jama.2009.1389

14. Sen S, Kranzler HR, Krystal JH, et al. A prospective cohort study investigating factors associated with depression during medical internship. *Arch Gen Psychiatry*. 2010;67(6):557-565. doi:10.1001/archgenpsychiatry.2010.41
15. Sen S, Kranzler HR, Didwania AK, et al. Effects of the 2011 duty hour reforms on interns and their patients: a prospective longitudinal cohort study. *JAMA Intern Med*. 2013;173(8):657-662. doi:10.1001/jamainternmed.2013.351
16. de Oliveira GS Jr, Chang R, Fitzgerald PC, et al. The prevalence of burnout and depression and their association with adherence to safety and practice standards: a survey of United States anesthesiology trainees. *Anesth Analg*. 2013;117(1):182-193. doi:10.1213/ANE.0b013e3182917da9
17. Heyhoe J, Birks Y, Harrison R, O'Hara JK, Cracknell A, Lawton R. The role of emotion in patient safety: are we brave enough to scratch beneath the surface? *J R Soc Med*. 2016;109(2):52-58. doi:10.1177/0141076815620614
18. Frank E, Prien RF, Jarrett RB, et al. Conceptualization and rationale for consensus definitions of terms in major depressive disorder: remission, recovery, relapse, and recurrence. *Arch Gen Psychiatry*. 1991;48(9):851-855. doi:10.1001/archpsyc.1991.01810330075011
19. DeRubeis RJ, Siegle GJ, Hollon SD. Cognitive therapy versus medication for depression: treatment outcomes and neural mechanisms. *Nat Rev Neurosci*. 2008;9(10):788-796. doi:10.1038/nrn2345
20. Oquendo MA, Bernstein CA, Mayer LES. A key differential diagnosis for physicians-major depression or burnout [published online July 17, 2019]?. *JAMA Psychiatry*. doi:10.1001/jamapsychiatry.2019.1332
21. Pereira-Lima K, Gupta RR, Guille C, Sen S. Residency program factors associated with depressive symptoms in internal medicine interns: a prospective cohort study. *Acad Med*. 2019;94(6):869-875. doi:10.1097/ACM.0000000000002567
22. Guille C, Frank E, Zhao Z, et al. Work-family conflict and the sex difference in depression among training physicians. *JAMA Intern Med*. 2017;177(12):1766-1772. doi:10.1001/jamainternmed.2017.5138
23. Tsutsumi A, Kawanami S, Horie S. Effort-reward imbalance and depression among private practice physicians. *Int Arch Occup Environ Health*. 2012;85(2):153-161. doi:10.1007/s00420-011-0656-1
24. Center C, Davis M, Detre T, et al. Confronting depression and suicide in physicians: a consensus statement. *JAMA*. 2003;289(23):3161-3166. doi:10.1001/jama.289.23.3161
25. Petrie K, Crawford J, Baker STE, et al. Interventions to reduce symptoms of common mental disorders and suicidal ideation in physicians: a systematic review and meta-analysis. *Lancet Psychiatry*. 2019;6(3):225-234. doi:10.1016/S2215-0366(18)30509-1
26. Moher D, Liberati A, Tetzlaff J, Altman DG; PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Open Med*. 2009;3(3):e123-e130.
27. Zhang J, Yu KF. What's the relative risk? a method of correcting the odds ratio in cohort studies of common outcomes. *JAMA*. 1998;280(19):1690-1691. doi:10.1001/jama.280.19.1690
28. Schünemann H. *The GRADE Handbook*. London, UK: Cochrane Collaboration; 2013.
29. Löwe B, Spitzer RL, Gräfe K, et al. Comparative validity of three screening questionnaires for DSM-IV depressive disorders and physicians' diagnoses. *J Affect Disord*. 2004;78(2):131-140. doi:10.1016/S0165-0327(02)00237-9
30. Higgins JPT, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ*. 2003;327(7414):557-560. doi:10.1136/bmj.327.7414.557
31. Higgins JPT, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Stat Med*. 2002;21(11):1539-1558. doi:10.1002/sim.1186
32. Sterne JAC, Jüni P, Schulz KF, Altman DG, Bartlett C, Egger M. Statistical methods for assessing the influence of study characteristics on treatment effects in 'meta-epidemiological' research. *Stat Med*. 2002;21(11):1513-1524. doi:10.1002/sim.1184
33. van Houwelingen HC, Arends LR, Stijnen T. Advanced methods in meta-analysis: multivariate approach and meta-regression. *Stat Med*. 2002;21(4):589-624. doi:10.1002/sim.1040
34. Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *BMJ*. 1997;315(7109):629-634. doi:10.1136/bmj.315.7109.629
35. Sterne JAC, Egger M. Funnel plots for detecting bias in meta-analysis: guidelines on choice of axis. *J Clin Epidemiol*. 2001;54(10):1046-1055. doi:10.1016/S0895-4356(01)00377-8.
36. R Development Core Team. The R Project For Statistical Computing. <http://www.R-project.org>. Accessed August 1, 2019.
37. Schwarzer G. meta: an R package for meta-analysis. *R News*. 2007;7(3):40-45.

38. Viechtbauer W. Conducting meta-analysis in R with the metafor package. *J Stat Softw*. 2010;36(3):1-48. doi:10.18637/jss.v036.i03
39. Hayashino Y, Utsugi-Ozaki M, Feldman MD, Fukuhara S. Hope modified the association between distress and incidence of self-perceived medical errors among practicing physicians: prospective cohort study. *PLoS One*. 2012;7(4):e35585. doi:10.1371/journal.pone.0035585
40. Kalmbach DA, Arnedt JT, Song PX, Guille C, Sen S. Sleep disturbance and short sleep as risk factors for depression and perceived medical errors in first-year residents. *Sleep*. 2017;40(3):zsw073. doi:10.1093/sleep/zsw073
41. Kang E-K, Lihm H-S, Kong E-H. Association of intern and resident burnout with self-reported medical errors. *Korean J Fam Med*. 2013;34(1):36-42. doi:10.4082/kjfm.2013.34.1.36
42. Shanafelt TD, Balch CM, Bechamps G, et al. Burnout and medical errors among American surgeons. *Ann Surg*. 2010;251(6):995-1000. doi:10.1097/SLA.0b013e3181bfdab3
43. Tawfik DS, Profit J, Morgenthaler TI, et al. Physician burnout, well-being, and work unit safety grades in relationship to reported medical errors. *Mayo Clin Proc*. 2018;93(11):1571-1580. doi:10.1016/j.mayocp.2018.05.014
44. West CP, Huschka MM, Novotny PJ, et al. Association of perceived medical errors with resident distress and empathy: a prospective longitudinal study. *JAMA*. 2006;296(9):1071-1078. doi:10.1001/jama.296.9.1071
45. Topp CW, Østergaard SD, Søndergaard S, Bech P. The WHO-5 well-being index: a systematic review of the literature. *Psychother Psychosom*. 2015;84(3):167-176. doi:10.1159/000376585
46. Henkel V, Mergl R, Kohnen R, Maier W, Möller H, Hegerl U. The WHO-5 wellbeing index performed the best in screening for depression in primary care. *BMJ*. 2003;8(5):155. doi:10.1159/000376585
47. Primack BA. The WHO-5 Wellbeing Index performed the best in screening for depression in primary care. *ACP J Club*. 2003;139(2):48. doi:10.1136/ebm.8.5.155
48. Mergl R, Seidscheck I, Allgaier AK, Möller HJ, Hegerl U, Henkel V. Depressive, anxiety, and somatoform disorders in primary care: prevalence and recognition. *Depress Anxiety*. 2007;24(3):185-195. doi:10.1002/da.20192
49. Saipanish R, Lotrakul M, Sumrithe S. Reliability and validity of the Thai version of the WHO-Five Well-Being Index in primary care patients. *Psychiatry Clin Neurosci*. 2009;63(2):141-146. doi:10.1111/j.1440-1819.2009.01933.x
50. Panagioti M, Geraghty K, Johnson J, et al. Association between physician burnout and patient safety, professionalism, and patient satisfaction: a systematic review and meta-analysis. *JAMA Intern Med*. 2018;178(10):1317-1330. doi:10.1001/jamainternmed.2018.3713
51. Schwenk TL, Gold KJ. Physician burnout—a serious symptom, but of what? *JAMA*. 2018;320(11):1109-1110. doi:10.1001/jama.2018.11703
52. Norton EC, Dowd BE. Log odds and the interpretation of logit models. *Health Serv Res*. 2018;53(2):859-878. doi:10.1111/1475-6773.12712
53. Norton EC, Dowd BE, Maciejewski ML. Odds ratios—current best practice and use. *JAMA*. 2018;320(1):84-85. doi:10.1001/jama.2018.6971
54. Bodenheimer T, Sinsky C. From triple to quadruple aim: care of the patient requires care of the provider. *Ann Fam Med*. 2014;12(6):573-576. doi:10.1370/afm.1713
55. Sikka R, Morath JM, Leape L. The Quadruple Aim: care, health, cost and meaning in work. *BMJ Qual Saf*. 2015;24(10):608-610. doi:10.1136/bmjqs-2015-004160
56. West CP. Physician well-being: expanding the triple aim. *J Gen Intern Med*. 2016;31(5):458-459. doi:10.1007/s11606-016-3641-2
57. Wallace JE, Lemaire JB, Ghali WA. Physician wellness: a missing quality indicator. *Lancet*. 2009;374(9702):1714-1721. doi:10.1016/S0140-6736(09)61424-0
58. Kay M, Mitchell G, Clavarino A, Doust J. Doctors as patients: a systematic review of doctors' health access and the barriers they experience. *Br J Gen Pract*. 2008;58(552):501-508. doi:10.3399/bjgp08X319486
59. Guille C, Speller H, Laff R, Epperson CN, Sen S. Utilization and barriers to mental health services among depressed medical interns: a prospective multisite study. *J Grad Med Educ*. 2010;2(2):210-214. doi:10.4300/JGME-D-09-00086.1
60. Agency for Healthcare Research Quality. National scorecard on rates of hospital-acquired conditions 2010 to 2015: interim data from National Efforts to Make Health Care Safer. <https://www.ahrq.gov/hai/pfp/2015-interim.html>. Published December 2016. Accessed September 20, 2019.
61. Weingart SN, Callanan LD, Ship AN, Aronson MD. A physician-based voluntary reporting system for adverse events and medical errors. *J Gen Intern Med*. 2001;16(12):809-814. doi:10.1046/j.1525-1497.2001.10231.x

62. Kontopantelis E, Springate DA, Reeves D. A re-analysis of the Cochrane Library data: the dangers of unobserved heterogeneity in meta-analyses. *PLoS One*. 2013;8(7):e69930. doi:10.1371/journal.pone.0069930

SUPPLEMENT.

eMethods. Syntaxes Used in Database Searches

eTable 1. Sensitivities and Specificities of Commonly Used Instruments With Cutoff Scores Adopted by Individual Studies Included in This Meta-analysis

eTable 2. Medical Errors Measurements Adopted by Individual Studies

eTable 3. Detailed Quality Assessment

eFigure 1. Sensitivity Analysis

eFigure 2. Subgroup Meta-analyses Stratified by Study-Level Characteristics

eFigure 3. Bubble Plots Displaying Meta-Regression Results

eFigure 4. Meta-analyses of the Associations Between Physician Depressive Symptoms and Medical Errors Stratified by Study Quality Indicators

eFigure 5. Funnel Plot

eReferences.