EDUCATION

Depressive symptoms and associated factors among obstetrics and gynecology first-year residents

Elena Frank, PhD; Julia Rossi, MPH; Zhuo Zhao, MS; Karina Pereira-Lima, PhD, MSc; Srijan Sen, MD, PhD; Abigail Winkel, MD, MHPE; Helen K. Morgan, MD

BACKGROUND: Medical training is a high-stress period, and residents in surgical specialties face elevated risk for depression due to demanding schedules, intensity of training, persistent mistreatment, and long work hours. Obstetrics and gynecology is the only surgical specialty predominantly composed of women, who face higher rates of depression than men and greater increases in depressive symptoms during internship. These challenges, compounded by work-family conflict and policy pressures, contribute to workforce strain in a specialty already facing a projected shortage. Understanding the mental health burden is critical to supporting both trainee well-being and long-term workforce sustainability, yet depressive symptoms, risk factors, and treatment-seeking have not been systematically studied in obstetrics and gynecology trainees.

OBJECTIVE: To assess the prevalence of depressive symptoms and mental health treatment-seeking among obstetrics and gynecology first-year residents and identify associated demographic, psychological, and workplace factors.

STUDY DESIGN: First-year obstetrics and gynecology residents enrolled in the longitudinal Intern Health Study between 2009 and 2024 completed surveys at enrollment and quarterly throughout internship. Depressive symptoms were measured using the 9-item Patient Health Questionnaire. Pearson correlations, χ^2 tests, stepwise linear regression, and generalized estimating equation models were used to identify baseline and internship-related predictors of depressive symptoms. Sample weights were applied to address nonrepresentative sampling and attrition biases.

RESULTS: Of 1603 enrolled obstetrics and gynecology interns (86.5% women; median age 27 years), 1271 (79.3%) completed at least one internship assessment and were included in the analysis. During internship, 35.6% screened positive for depression on the 9-item Patient Health Questionnaire at one or more assessments, yet less than one-third of those affected sought mental health treatment. Baseline predictors of increased depressive symptoms included history of depression, higher baseline depressive symptoms, neuroticism, a difficult early family environment, and not being in a committed relationship. Internship-related factors associated with worsening symptoms included fewer sleep hours, longer work hours, and reported medical errors.

CONCLUSION: More than one-third of obstetrics and gynecology residents screened positive for depression on the 9-item Patient Health Questionnaire during their first year, yet treatment-seeking remained critically low. Both preexisting vulnerabilities and modifiable workplace stressors contributed to depressive symptoms. These findings are particularly notable given the female-predominant composition of the obstetrics and gynecology field and the growing representation of women in surgical specialties. The low treatment-seeking rate highlights the urgent need for improved mental health access, including opt-out service models. Structural interventions targeting workload and sleep are essential to support trainee well-being and long-term workforce sustainability in obstetrics and gynecology.

Key words: depressive symptoms, gender, mental health treatment seeking, obstetrics and gynecology, physician well-being, residency training, workforce sustainability

Introduction

Medical training is a high-stress period, with trainees in surgical specialties potentially at an elevated risk for depression due to more rigid schedules, intensity of training, and persistent mistreatment.[1,2] Even after the 2011 Accreditation Council for Graduate Medical Education duty hours reforms, surgical trainees continue to work more hours and sleep less than their nonsurgical counterparts.[3,4] Obstetrics and gynecology (OBGYN) is the only surgical specialty predominantly composed of women.[5] Women are at a heightened risk for depression compared to men[4] and experience greater increases in depressive symptoms over the course of internship.[6] Suffering depression is not only a source of distress in residents, increasing the risk of suicide, but has also been associated with increased reports of medical errors and reduced quality of patient care.[6]

Along with the demands of medical training, women face additional pressures (eg, higher work family conflict) reflective of broader gender inequities in professional and familial expectations that place them at a greater risk for depression, burnout, and emotional exhaustion compared to men.[6–8] Thus, contributing to a reduction of work hours or leaving full-time positions earlier than men,[9] which may have downstream effects on retention, pay, and advancement.[7,10] This is especially concerning in a workforce with a female majority, with a projected shortage of OBGYNs expected to progressively worsen over the next decade.[11] This trend is furthered by reproductive health policy challenges such as the 2022 *Dobbs* decision, which adds clinical, legal, and personal burden.[12] These pressures increase moral distress, anxiety, and depression among OBGYN physicians and early studies suggest this may contribute to more providers leaving the field.[13,14] Supporting long-term workforce sustainability in OBGYN starts with addressing trainee well-being.

Despite the importance of addressing mental health challenges in OBGYN

0002-9378

https://doi.org/10.1016/j.ajog.2026.07.023

AJOG at a Glance

Why was this study conducted?

Obstetrics and gynecology (OBGYN) is the only surgical specialty predominantly composed of women, who face higher rates of depression than men and greater increases in depressive symptoms during internship. Depressive symptoms, risk factors, and treatment-seeking have not been systematically studied in OBGYN trainees despite growing concerns about strains on the workforce.

Key findings

More than one-third of OBGYN interns screened positive for depression on the 9-item Patient Health Questionnaire during internship, yet less than one-third sought mental health treatment. Predictors included history of depression, higher baseline depressive symptoms, neuroticism, difficult early family environment, not being in a committed relationship, sleep deprivation, long work hours, and medical errors.

What does this add to what is known?

This is the first longitudinal study to systematically characterize depressive symptoms and treatment-seeking in OBGYN residents, underscoring the need for opt-out mental health models and structural interventions targeting workload and sleep.

physicians, depression prevalence, risk factors, and treatment-seeking among trainees remain poorly understood. Exploring both individual characteristics and the context of training is important to identify organizational interventions that can support the developing workforce. In this study, we assessed the prevalence of depressive symptoms and mental health treatment-seeking among OBGYN first-year residents (interns) and identified associated demographic, psychological, and workplace factors.

Materials and methods

The Intern Health Study is a prospective cohort study assessing stress and well-being during the first year of residency training (internship) in the United States.[1] Each year following the national residency match (mid-March), a convenience sample of US academic and community healthcare institutions across diverse practice settings and geographic locations provided names and e-mail addresses of graduating medical school students or incoming first-year residents. Additional information was also obtained from publicly available medical school match lists. Between 2009 and 2024, a total of 55,031 incoming interns at US residency institutions were invited via e-mail to participate in the 3 months prior to the start of internship (July 1). In total, 26,864 participants (48.8%) consented to enroll, representing 833 US programs across all specialties (Figure 1 provides additional recruitment details). The University of Michigan institutional review board approved the study. Participants provided electronic consent and were compensated $25 to $130 depending on survey completion and cohort year. The study followed the American Association for Public Opinion Research reporting guideline.

Participants completed a baseline survey at enrollment via the web assessing depressive symptoms using the 9-item Patient Health Questionnaire (PHQ-9).[15] Participants rated items on a 4-point Likert scale from 0 (not at all) to 3 (nearly every day). Total scores range from 0 to 27. This total score, indicating severity of depression symptoms, is consistent with the validated and recommended use of the measure. The diagnostic validity of the PHQ-9 is similar to that of clinician-administered assessments, with a PHQ-9 score equal to or greater than 10 having a sensitivity of 88% and specificity of 88% for a Diagnostic and Statistical Manual diagnosis of major depressive disorder.

The baseline survey also included basic demographic information (age, gender, race/ethnicity, and relationship status), personal history of depression, neuroticism, and early family environment information.

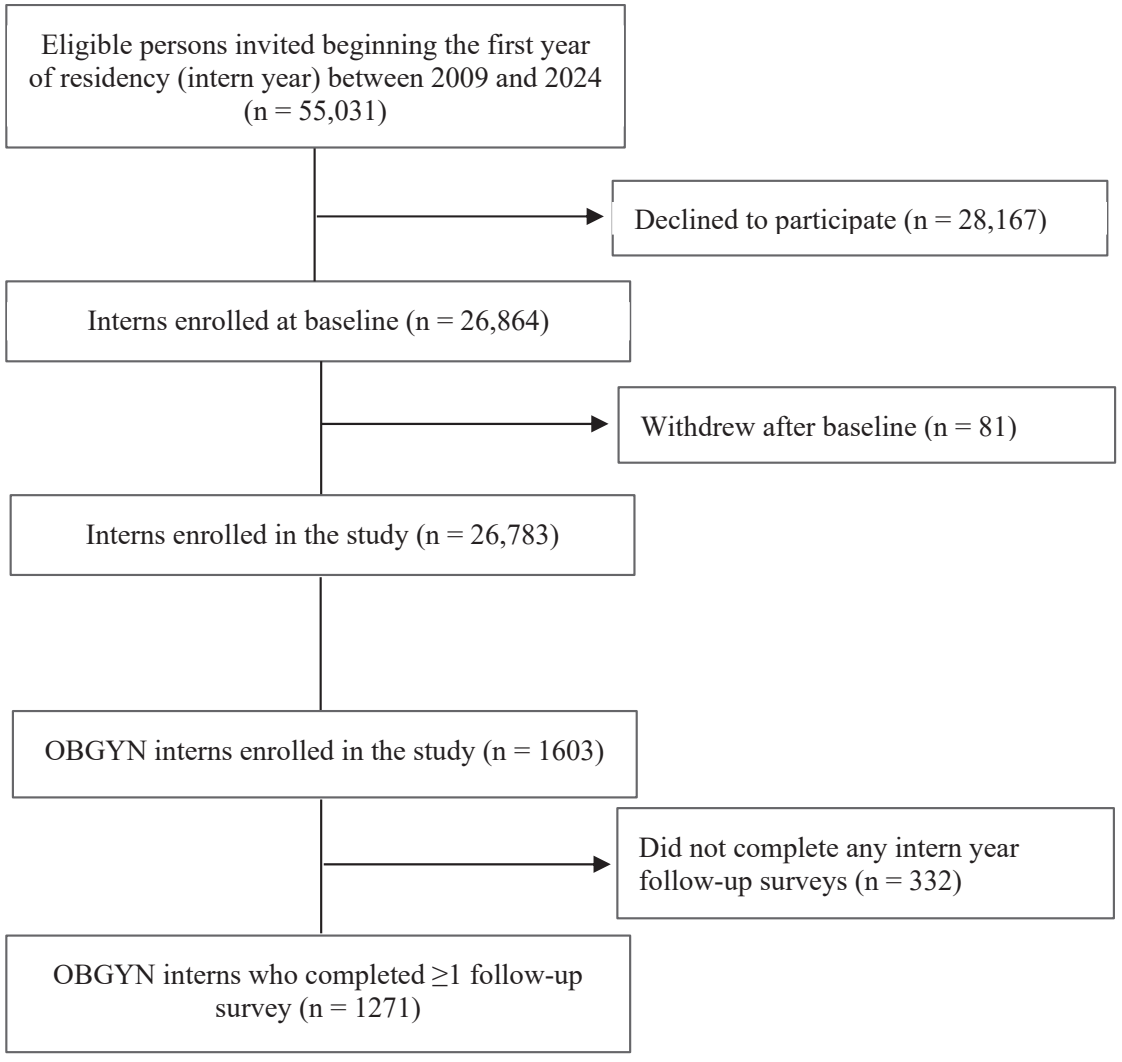
Personal history of depression was assessed using the following question with a "yes" or "no" response: "To the best of your recollection, have you EVER experienced an episode of depression (a 2-week period of your life when you felt down or lost interest or pleasure in your usual activities and also had difficulty concentrating or noticed changes in sleep, appetite, and energy or experienced thoughts of death or feelings of guilt)?".

Neuroticism was evaluated using the neuroticism subscale of the Neuroticism, Extraversion, and Openness to Experience Five-Factor Inventory, a 14-item scale assessing negative emotional constructs, such as anxiety, mood, worry, envy, and jealousy.[16] Participants rated items on a 5-point Likert scale from 0 (strongly disagree) to 4 (strongly agree), resulting in a total score of 0 to 56 (higher scores indicate increased likelihood of experiencing negative emotions, particularly in response to environmental stress).

Exposure to difficult early family environments during childhood and early adolescence (aged 5–15 years) was evaluated using the Risky Families Questionnaire.[17] Participants rated 13 childhood experiences (eg, insults, drinking, quarreling, loving, hugging) on a 6-point Likert scale from 1 (not at all) to 6 (very often), resulting in a total score of 0 to 65 (higher scores indicate a more chaotic, harsh, and abusive early family environment, whereas lower scores indicate a more organized, supportive, and nurturing family environment).

PHQ-9 depressive symptoms were also assessed via the web at 3, 6, 9, and 12 months during intern year.[1] Work hours, sleep hours, mental health treatment seeking, perceived medical errors, and stressful life events were also assessed

FIGURE 1
Enrollment flow diagram



quarterly. Stressful life events were assessed using an 11-item questionnaire adapted for the Intern Health Study where participants indicated if they experienced the death of family member, considerable other, or close friend; developed a disabling illness or injury lasting a month or more; a disabling physical illness or injury started or got worse in a family member, significant other, or close friend; a relationship with an intimate cohabiting partner ended; were involved in a physically violent relationship; suffered a significant financial loss or loss of property; had problems with debt; were physically assaulted or attacked; got married; and learned you or a partner were pregnant or had a child.[1]

Due to the substantial time lag between data collection and analysis, it was not possible to act in a timely manner on any personal information provided. Participants were informed of this in the consent document and received general information about depression as well as a list of resources for mental health counseling and treatment upon completing each survey.

In accordance with previously described methods,[4] we used a sample weighting strategy to reduce the potential bias introduced from differences between participants and nonparticipants and between participants who completed the follow-up surveys and those who did not.

To reduce potential bias owing to sampling in estimating the changes over time, we used data from the Association of American Medical Colleges on the overall characteristics of first-year OBGYN residents in the United States from 2009 to 2024 as the reference target. We performed a 2-step poststratification raking on the study data to construct weights such that the distribution of cohort year, sex, race, and ethnicity in the sample matched the

Association of American Medical Colleges data distribution. The first step was to generate between-cohort weights (w1b) with the raking variable to be cohort year; the second step was to generate weights within each cohort (w1w) with the raking variables to be sex, race, and ethnicity.

To account for differences between participants who completed the follow-up surveys and those who did not, we identified 3 baseline variables correlated with survey completion, cohort year, race, and baseline stressful life events, to calculate probability of participation using the PSMATCH SAS procedure (SAS, version 9.4 [SAS Institute]). The attrition weights (w2) for each person who completed the follow-up surveys were the inverse of the probability of participation.

The final weights were the product of poststratification weights and attrition weights (w1b×w1w×w2). We applied trimming and standardizing procedures to the final weights to address outliers and to ensure the weighted sample size matched the unweighted sample size. The final analytic weight, incorporating both poststratification weights and attrition weights, was applied to all primary analyses involving follow-up data, including the linear regression and generalized estimating equation (GEE) models.

We used Pearson correlations for continuous measures, and χ^2 tests for nominal measures, to examine associations between the mean change of PHQ-9 scores (difference between the mean of 4 quarterly internship PHQ-9 scores and the baseline PHQ-9 score) and baseline demographic and psychological factors previously associated with depressive symptoms among the broader intern physician population. Significant variables were subsequently entered into a stepwise linear regression model, to identify significant predictors while accounting for collinearity among variables. Variables were eligible to enter and remain in the model at $P<.15$. To explore internship factors associated with depressive symptoms, we included repeated quarterly follow-up measures of work hours, sleep hours, medical errors, and stressful life events in a GEE model adjusting for significant baseline predictors. For the GEE analysis, 4 repeated data were transformed to long format with one record for each quarterly internship assessment from 3 months through 12 months. The outcome was PHQ score at each quarterly assessment. Baseline PHQ score was included as an adjusting covariate to control for preinternship depressive symptoms. Analyses were conducted in SAS v9.4 (SAS Institute Inc). Two-sided $P<.05$ was considered significant.

Results

Of the 26,783 incoming first-year residents who enrolled in the overall study and completed the preinternship baseline survey, 1603 were entering OBGYN programs (median [interquartile range] age, 27 [26−28] years; 1369 women [86.5%]) (Table 1). Participants represented 191 residency programs across diverse regions of the United States (Northeast: 32.1%, Midwest: 22.5%, South: 29.8%, and West: 15.7%). After sample weighting, 879 (54.8%) participants reported a history of depression. Of 1603 residents, 1271 (79.3%) completed at least one internship assessment and were included in the analysis. Survey completion rates were 72.3%, 60.7%, 56.5%, and 51.3% at the 3, 6, 9, and 12 months of follow-up time points, respectively. The median PHQ-9 score increased from 2 (1−4) to 6 (3.3−8.8) from preinternship to internship. Prior to the start of internship, 54 (4.2%) of participants screened positive for depression on the PHQ-9 (≥10). During internship, 453 (35.6%) screened positive for depression at one or more quarterly assessments, with 20.1%, 20.3%, 22.3%, and 22.3% screening positive for depression at 3, 6, 9, and 12 months, respectively (Figure 2). Depressive symptoms were significantly increased at the 3-month assessment compared to preinternship (β=3.2, 95% confidence interval [CI], 3.0−3.4; $P<.001$). Depressive symptoms remained elevated throughout intern year with no significant difference between assessments at 3, 6, 9, and 12 months (Figure 3). Among interns who screened positive for depression on the PHQ-9 during intern year, 142 (31.4%) sought mental health treatment.

Baseline depression scores did not differ significantly between participants who completed only the baseline survey and were lost to all 4 quarterly follow-ups and those who completed at least one internship assessment (PHQ-9 score, mean=3.14, standard deviation=3.21 vs mean=2.79, standard deviation=2.96; Welch's t [446.04]=1.75, P=.081). Attrition weights were applied to account for differential follow-up based on observed baseline characteristics.

Among 11 baseline variables (Table 1) tested, 6 were significantly correlated with a change in depressive symptoms from the baseline assessment to the mean of the 4 internship assessments: history of depression (r=0.091, P=.001), lower baseline depressive symptoms (r=0.266, $P<.0001$), higher neuroticism (r=0.137, $P<.0001$), difficult early family environment (r=0.063, P=.025), not being in a committed relationship (r=0.074, P=.009), and lower work hours at baseline (r=0.065, P=.025). When these 6 variables were entered into the stepwise linear regression to account for collinearity between variables, 5 variables remained significant: history of depression (β=0.67; 95% CI, 0.28−1.06; $P<.001$), lower baseline depressive symptoms (β=0.59; 95% CI, 0.52−0.66; $P<.001$), higher neuroticism (β=0.16; 95% CI, 0.13−0.18; $P<.001$), difficult early family environment (β=0.02; 95% CI, 0.004−0.04; P=.018), and not being in a committed relationship (β=0.49; 95% CI, 0.13−0.85; P=.008). These 5 significant baseline variables, along with female sex, which was significant in prior analyses, were included as covariates in the GEE analysis with 4 repeated measures during internship. In these adjusted analyses, 3 within-internship variables, lower sleep hours, higher work hours, and medical errors, were associated with higher PHQ-9 scores during internship (Table 2).

Comment

Principal findings

In this 16-year national study, more than one-third of OBGYN residents screened positive for depression on the

TABLE 1
Baseline and internship characteristics of OBGYN first-year residents, 2009 to 2024

Characteristic				
Baseline	**Unweighted, N=1603**		**Weighted, N=1603**	
Median age (IQR), y	27.0	(26.0–28.0)	27.0	(26.0–28.0)
Gender, n (%)				
Men	213	13.5	194	12.4
Women	1369	86.5	1375	87.6
Race, n (%)				
Arab/Middle Eastern	14	0.9	18	1.2
Asian (eg, Indian, Chinese)	238	14.9	214	13.4
Black/African American	123	7.7	135	8.5
Latino/Hispanic	87	5.4	89	5.6
Native American	1	0.1	1	0.1
White	962	60.2	948	59.2
Multiracial	151	9.5	158	9.9
Other	22	1.4	37	2.3
Relationship status, n (%)				
Committed relationship	726	45.4	833	52.1
Not in committed relationship	872	54.6	766	47.9
Children, n (%)				
No	1516	94.9	1514	94.7
Yes	82	5.1	85	5.3
Depressive symptoms: median PHQ-9 score (IQR)[a]	2	(1–4)	2	(1–4)
Depression history, n (%)				
No	746	46.6	724	45.2
Yes	856	53.4	879	54.8
Neuroticism: median NEO Five-Factor Inventory score (IQR)[b]	24	(18–30)	24	(19–30)
Difficult early family environment: median Risky Families Questionnaire score (IQR)[c]	11	(6–17)	11	(6–17)
Median self-reported daily sleep hours (IQR)	8	(7–9)	8	(7–9)
Practice region, n (%)				
Midwest	357	22.5	350	22.4
Northeast	508	32.1	546	34.9
South	473	29.8	441	28.1
West	248	15.7	230	14.7
Internship	**Unweighted, N=1271**		**Weighted, N=1271**	
Average internship depressive symptoms: median PHQ-9 score (IQR)[a]	6	(3.3–8.5)	6	(3.3–8.8)
Physicians who met PHQ-9 criteria for depression during intern year (≥10), n (%)				
No	822	64.7	818	64.4
Yes	449	35.3	453	35.6

(continued)

TABLE 1
Baseline and internship characteristics of OBGYN first-year residents, 2009 to 2024 *(continued)*

Internship	Unweighted, N = 1271		Weighted, N = 1271	
Median self-reported weekly work hours (IQR)	70	(63.3–76)	70	(62.7–75.5)
Median self-reported daily sleep hours (IQR)	7	(6.1–7.5)	7	(6.1–7.5)
Self-reported medical error, n (%)				
No	781	61.5	788	62.0
Yes	490	38.6	483	38.0
Stressful life event, n (%)				
No	697	54.8	687	54.0
Yes	574	45.2	584	46.0
Sought mental health treatment when depressed, n (%)				
No	299	66.6	310	68.6
Yes	150	33.4	142	31.4

IQR, interquartile range; *OBGYN*, obstetrics and gynecology; *NEO*, Neuroticism, Extraversion, and Openness; *PHQ-9*, 9-item Patient Health Questionnaire.

[a] Range, 0–27. Scores of 0–4 indicate minimal depression, 5–9 mild depression, 10–14 moderate depression, 15–19 moderately severe depression, and 20–27 severe depression; [b] Range, 0–56; [c] Range, 0–65.

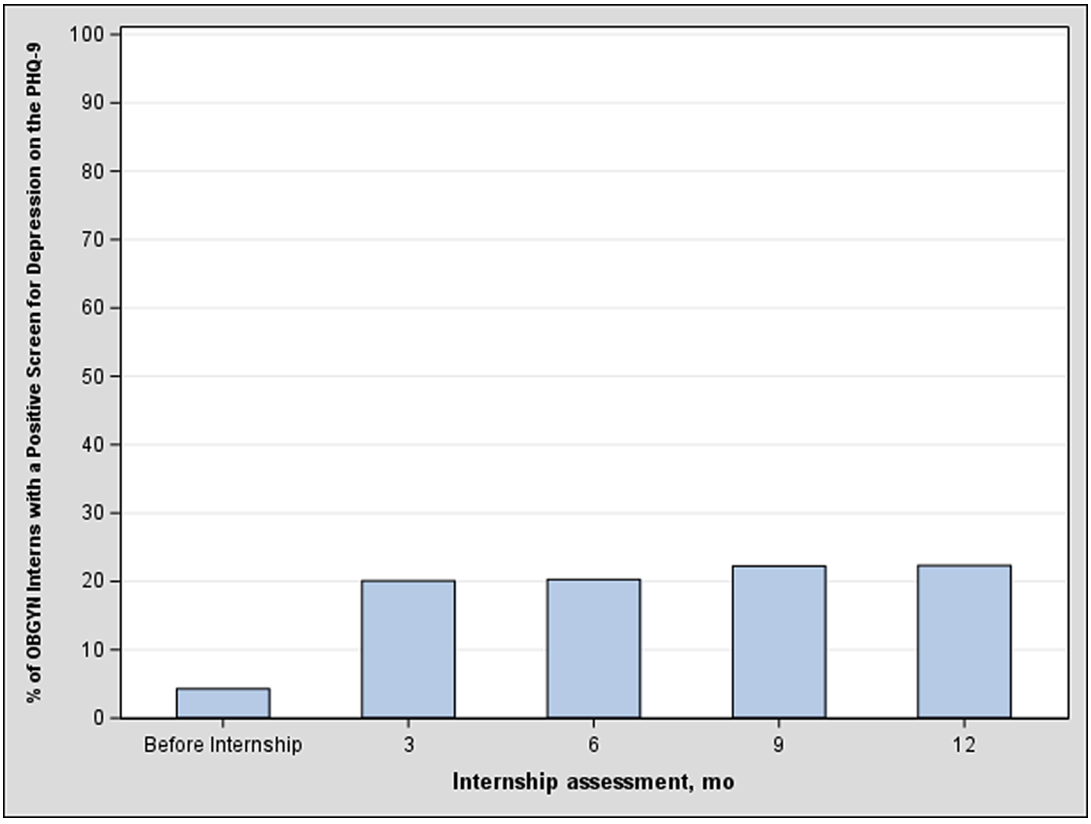
PHQ-9 during their first year, yet less than one-third of those affected sought mental health treatment. Both pre-internship factors (history of depression, baseline depression, and neuroticism) and internship-related factors (sleep hours, work hours, and medical errors) were associated with depressive symptoms.

Results in the context of what is known

The high prevalence of depressive symptoms and primary associated factors during OBGYN internship is consistent with prior data on intern year more broadly.[1–3] However, these mental health findings are particularly noteworthy given the OBGYN specialty's predominantly female composition, especially as more women enter surgical specialties. Trends in the rate of depression over time across all specialties have been reported previously,[4] with the average change in depressive symptoms with internship decreasing 24.4% from 2007 to 2019. This change was greater among women than among men and among nonsurgical interns. Prior literature also suggests that women physicians experience higher baseline rates of depressive symptoms and greater increases during internship and greater exposure to work-family conflict and structural stressors in training.[6–8] In this context, the high prevalence of those who met criteria for depression combined with low rates of treatment-seeking observed in our cohort may reflect, in part, the broader vulnerability of the female-majority specialty during early training. Thus, highlighting the need for early, proactive identification of at-risk trainees and improved access to confidential, low-barrier mental healthcare within residency programs.

The association between depressive symptoms and modifiable training factors such as sleep deprivation, work hours, and medical errors suggests that interventions targeting residency structure, not only individual-level resilience, may improve trainee well-being. A growing body of systematic reviews and meta-analyses has demonstrated that physician burnout and mental health strain are associated with lower quality of care, increased medical errors, and compromised patient safety.[18,19] Efforts to address physician burnout at the system level have included hiring scribes and using artificial intelligence to ease electronic health record burden, bringing on additional clerical staff, reassigning clinical tasks to other healthcare professionals, and redesigning health information technology workflows.[20–22] The use of artificial intelligence scribes was associated with a decrease in clinician burnout from 51.9% to 38.8% after 30 days,[23] while another study found that adding full-time clerical support for physician order entry and previsit planning reduced physicians reporting weekly burnout from 43% to 14% after 4 months.[24] While these approaches have shown promise in addressing burnout, reducing medical errors, and improving quality of care, their effectiveness may vary depending on institutional culture, staffing, and policy environments that influence implementation and sustainability. Importantly, they may also not fully address gendered dimensions of workload and mental health, including work-family conflict and the disproportionate domestic and caregiving responsibilities often experienced by women physicians during training.[6–8] As a result, the effectiveness of these interventions in mitigating sleep deprivation, long work hours, and medical errors, and ultimately reducing depressive symptoms, may vary across trainees depending on both institutional context and social and gendered conditions.

FIGURE 2
Frequency of OBGYN interns with a positive screen for depression on the Patient Health Questionnaire-9 across internship



This figure displays the frequency of OBGYN interns who screened positive for depression on the Patient Health Questionnaire-9 (≥10) at 4 quarterly assessment time points during intern year.

OBGYN, obstetrics and gynecology.

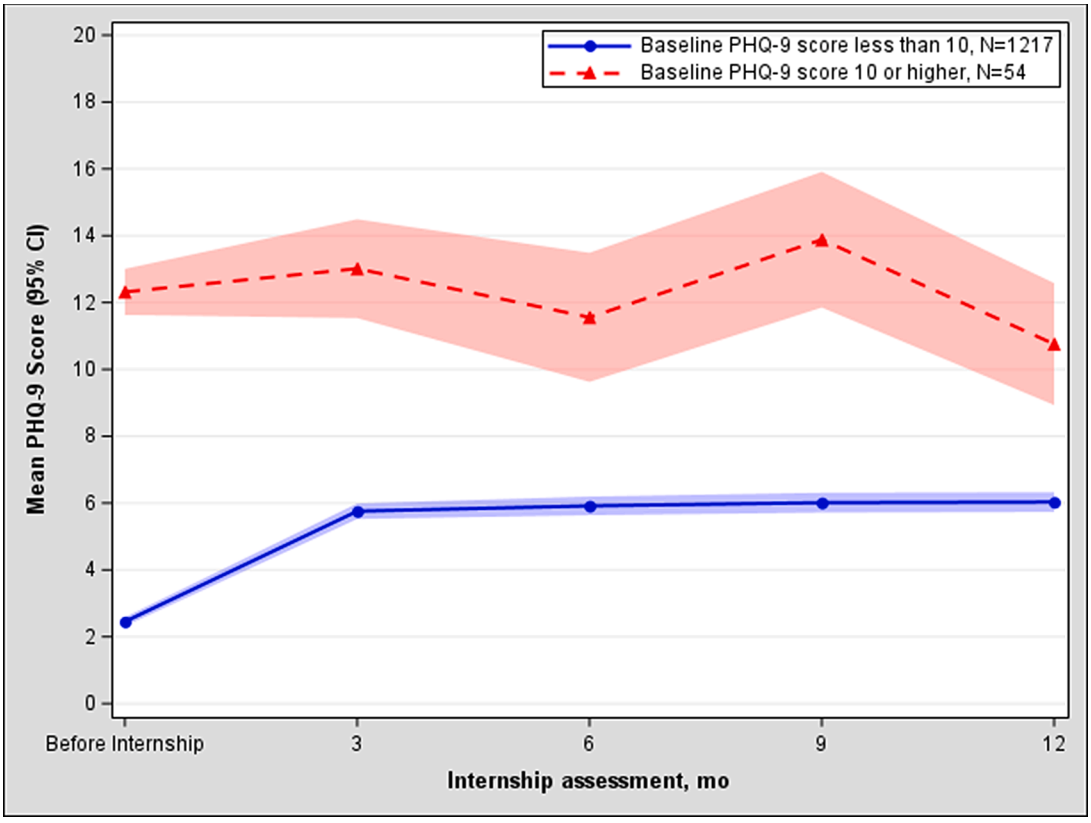
Recognizing the vulnerability of the OBGYN workforce to depression reflected in this study is essential to better attend to the mental health of trainees and support the long-term well-being of the workforce. Providing access to mental healthcare is essential in residency programs. Most residency programs use an opt-in model of mental health services. However, studies have demonstrated the success of opt-out mental health service models, showing considerable increases in residents' psychotherapy utilization, greater willingness to engage in mental healthcare, improved attitudes toward seeking help, and reduced mental illness stigma.[25,26] This opt-out model may be particularly impactful in residency settings and may warrant broader adoption compared to traditional opt-in approaches given the findings that less than one-third of those affected sought mental health treatment. Embedding mental health services as a default component of training infrastructure may therefore reduce barriers that are especially pronounced during internship.

Regional context also matters as workforce shortages may undermine these system-level interventions.[11] More recently, the *Dobbs* decision has contributed to increased moral distress, anxiety, and depression due to legal and ethical burdens among OBGYN physicians,[13,14] highlighting the need for both individualized and structural approaches to address its impact on OBGYN resident mental health.

Clinical implications

These findings indicate that supporting OBGYN intern mental health through both organizational structures and individual-level support is critical. Tailoring interventions to local resources, workforce capacity, and evolving policy environments will be essential to ensure meaningful impact. At the same time, efforts to normalize help-seeking, reduce stigma, and embed accessible mental healthcare into residency training can address persistent gaps in treatment utilization. Together, these strategies underscore that improving resident mental health is not solely an individual responsibility but a systems-level priority. In the context of a predominantly female specialty with established vulnerability to depressive symptoms during training,

FIGURE 3
Mean PHQ-9 scores across OBGYN internship



The line on this figure denotes the mean Patient Health Questionnaire (PHQ-9) for OBGYN interns from before the start of internship across the 4 quarterly internship assessment time points. The 95% confidence interval is represented in the shaded region. Weighted linear regression showed significant between-group differences in mean PHQ-9 scores at all follow-up assessments comparing baseline PHQ-9 scores less than 10 vs 10 or higher: internship 3, 6, 9, and 12 months, all $P<.0001$. Shaded bands indicate 95% confidence intervals around the weighted mean PHQ-9 scores.
OBGYN, obstetrics and gynecology; *PHQ-9*, 9-item Patient Health Questionnaire.

such structural support may be particularly important for sustaining workforce well-being and retention. Investing in the well-being of trainees early in their careers is essential not only for their personal health but also for maintaining a resilient physician workforce and ensuring the delivery of safe, high-quality patient care.

Research implications

Future research should assess additional psychological and interpersonal factors, such as social support, that may influence OBGYN trainee mental health. Further research is also needed to understand barriers to mental health treatment-seeking among OBGYN interns, including individual and structural factors that contribute to low utilization despite high rates of depressive symptoms. Furthermore, future work should further evaluate the uptake and effectiveness of opt-out mental health service models across diverse training environments to increase utilization of mental health services. In addition, studies should explore whether depressive symptoms cluster within training programs or are influenced by cohort-level or peer effects, including the potential for within-institution variation in symptom burden. Longitudinal designs are also needed to assess temporal patterns in depressive symptoms across internship, including whether rates change over time or vary by season. Finally, future research should examine how evolving reproductive health policy environments, including the post-*Dobbs* landscape, shape depressive symptoms and willingness to seek care and inform both individual and structural interventions to support trainee mental health.

Strengths and limitations

A major strength of our study is the longitudinal design and large national sample of OBGYN first-year residents. However, this study has several limitations. First, as an observational study with convenience sampling, we could not definitively determine the causes of changes in depressive symptoms. There may be additional unaccounted-for influential individual, institutional, and societal

TABLE 2
Factors associated with increased depressive symptoms

Factors	β	95% CI		*P* value
Baseline factors				
Baseline depressive symptoms	0.04	0.03	0.06	**<.001**
Neuroticism	0.03	0.02	0.03	**<.001**
Difficult early family environment	0.002	−0.002	0.01	**<.001**
Female gender	0.04	−0.11	0.19	.33
In a committed relationship	−0.02	−0.10	0.05	.61
Personal history of depression	0.16	0.07	0.25	**<.001**
Internship factors				
Sleep hours	−0.11	−0.15	−0.07	**<.001**
Work hours	0.005	0.002	0.01	**<.001**
Stressful life events	0.02	−0.04	0.08	.53
Medical errors	0.11	0.04	0.18	**.003**

GEE indicates generalized estimating equation; β coefficients are from a weighted GEE model. The outcome was PHQ-9 score at each completed quarterly internship assessment at 3, 6, 9, and 12 months. Baseline PHQ-9 score was included as an adjusting covariate. Other baseline covariates included neuroticism, difficult early family environment, female sex, committed relationship status, and history of depression. Internship predictors included average sleep hours, work hours, stressful life events, and reported medical errors. Statistically significant values are in **bold**.

CI, confidence interval.

factors. Second, missing data and participant attrition may have biased the findings, although this risk was lessened by the sample weighting. As invitee specialty was not available for all study cohorts, it was not possible to identify the specific proportion of invited OBGYN interns who opted into the study, only the overall enrollment rate across specialties. Third, depressive symptoms were assessed via self-report rather than a diagnostic interview. Fourth, the sample was restricted to the first year of residency; depressive symptom rates may vary among other groups of trainees and practicing physicians.

Conclusion

Given the current workforce shortage and composition of OBGYN practitioners,[11] addressing individual and system-level drivers of physician well-being is critical. The low treatment-seeking rate underscores the need for improved access to mental health treatment, potentially through opt-out models.[25,26] Furthermore, structural interventions targeting workload and sleep are essential to improve trainee well-being and support long-term workforce sustainability in OBGYN. ■

Acknowledgments

We would like to acknowledge and thank the first-year residents taking part in this research. Without their efforts, this study would not be possible.

References

1. Sen S, Kranzler HR, Krystal JH, et al. A prospective cohort study investigating factors associated with depression during medical internship. Arch Gen Psychiatry 2010;67:557–65.
2. Hu YY, Ellis RJ, Hewitt DB, et al. Discrimination, abuse, harassment, and burnout in surgical residency training. N Engl J Med 2019;381:1741–52.
3. Hughes TM, Waljee JF, Fang Y, et al. New-onset depression among surgical interns. JAMA Surg 2022;157:543–5.
4. Fang Y, Bohnert AS, Pereira-Lima K, et al. Trends in depressive symptoms and associated factors during residency, 2007 to 2019: a repeated annual cohort study. Ann Intern Med 2021;175:56–64.
5. Rayburn WF. Diversity in academic obstetrics and gynecology. Obstet Gynecol Clin North Am 2024;51:181–91.
6. Guille C, Frank E, Zhao Z, et al. Work-family conflict and the sex difference in depression among training physicians. JAMA Intern Med 2017;177:1766–72.
7. Frank E, Zhao Z, Sen S, et al. Gender disparities in work and parental status among early career physicians. JAMA Netw Open 2019;2:e198340.
8. Frank E, Zhao Z, Fang Y, et al. Experiences of work-family conflict and mental health symptoms by gender among physician parents during the COVID-19 pandemic. JAMA Netw Open 2021;4:e2134315.
9. Winkel AF, Morgan HK, Hammoud MM, et al. Burnout and well-being in trainees: findings from a national survey of US obstetrics and gynecology residents. J Grad Med Educ 2024;16:572–80.
10. Chen YW, Orlas C, Kim T, et al. Workforce attrition among male and female physicians working in US academic hospitals, 2014-2019. JAMA Netw Open 2023;6:e2323872.
11. Rayburn WF, Xierali IM. Projected shortages and distributional challenges of obstetrician-gynecologists in the United States, 2025-2035. Obstet Gynecol 2026;147:4–7.
12. Cates FE, Aiken CE, Aiken ARA. OB-GYN job satisfaction in a changing landscape of legal restrictions on their practice. Contraception 2026;161:111455.
13. Buchbinder M, Arora KS, McKetchnie SM, et al. Sources of moral distress among obstetrician-gynecologists after dobbs: a qualitative, multi-state study. SSM Qual Res Health 2024;6:100483.
14. Sabbath EL, McKetchnie SM, Arora KS, et al. US obstetrician-gynecologists' perceived impacts of post—*dobbs v jackson* state abortion bans. JAMA Netw Open 2024;7:e2352109.
15. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med 2001;16:606–13.
16. Murray G, Rawlings D, Allen NB, Trinder J. NEO five-factor inventory scores: psychometric properties in a community sample. Meas Eval Couns Dev 2003;36:140–9.

17. Taylor SE, Way BM, Welch WT, Hilmert CJ, Lehman BJ, Eisenberger NI. Early family environment, current adversity, the serotonin transporter promoter polymorphism, and depressive symptomatology. Biol Psychiatry 2006;60:671676.

18. Panagioti M, Panagopoulou E, Bower P, et al. Controlled interventions to reduce burnout in physicians: a systematic review and meta-analysis. JAMA Intern Med 2017;177:195–205.

19. Dewa CS, Loong D, Bonato S, et al. The relationship between resident burnout and safety-related and acceptability-related quality of healthcare: a systematic literature review. BMC Med Educ 2017;17:195.

20. Gottlieb M, Palter J, Westrick J, et al. Effect of medical scribes on throughput, revenue, and patient and provider satisfaction: a systematic review and meta-analysis. Ann Emerg Med 2021;77:180–9.

21. Tran BD, Rosenbaum K, Zheng K. An interview study with medical scribes on how their work may alleviate clinician burnout through delegated health IT tasks. J Am Med Inform Assoc 2021;28:907–14.

22. Hillmann W, Hayes BD, Marshall J, et al. Improving burnout through reducing administrative burden: a pilot of pharmacy-driven medication histories on a hospital medicine service. J Gen Intern Med 2021;36:2511–3.

23. Olson KD, Meeker D, Troup M, et al. Use of ambient AI scribes to reduce administrative burden and professional burnout. JAMA Netw Open 2025;8:e2534976.

24. Contratto E, Romp K, Estrada CA, Agne A, Willett LL. Physician order entry clerical support improves physician satisfaction and productivity. South Med J 2017;110:363–8.

25. Guldner G, Siegel JT, Broadbent C, et al. Use of an opt-out vs opt-in strategy increases use of residency mental health services. J Grad Med Educ 2024;16:195–201.

26. Kevern T, Davies DR, Stiel K, et al. Impact of opt-out therapy appointments on resident physicians' mental health, well-being, stigma, and willingness to engage. J Grad Med Educ 2023;15:734–7.

Author and article information

From the Michigan Neuroscience Institute, University of Michigan, Ann Arbor, MI (Frank, Zhao and Sen); Department of Obstetrics and Gynecology, Institute for Innovations in Medical Education, New York University Grossman School of Medicine, New York (Rossi and Winkel); Department of Anesthesiology, University of Michigan, Ann Arbor, MI (Pereira-Lima); Eisenberg Family Depression Center, University of Michigan, Ann Arbor, MI (Sen); Department of Psychiatry, University of Michigan, Ann Arbor, MI (Sen); and Departments of Obstetrics and Gynecology and Learning Health Sciences, University of Michigan, Ann Arbor, MI (Morgan).

Received April 27, 2026; revised July 15, 2026; accepted July 15, 2026.

The authors report no conflict of interest.

This work was supported by the Josiah Macy Jr. Foundation (grant ID P25-16).

Corresponding author: Elena Frank, PhD. emfrank@med.umich.edu