

Journal Pre-proof



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

PII: S0002-9378(26)00372-8

DOI: <https://doi.org/10.1016/j.ajog.2026.07.023>

Reference: YMOB 16844

To appear in: *American Journal of Obstetrics and Gynecology*

Received Date: 27 April 2026

Revised Date: 15 July 2026

Accepted Date: 15 July 2026

Please cite this article as: Frank E, Rossi J, Zhao Z, Pereira-Lima K, Sen S, Winkel A, Morgan HK, Depressive Symptoms and Associated Factors Among Obstetrics and Gynecology First-Year Residents, *American Journal of Obstetrics and Gynecology* (2026), doi: <https://doi.org/10.1016/j.ajog.2026.07.023>.

This is a PDF of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability. This version will undergo additional copyediting, typesetting and review before it is published in its final form. As such, this version is no longer the Accepted Manuscript, but it is not yet the definitive Version of Record; we are providing this early version to give early visibility of the article. Please note that Elsevier's sharing policy for the Published Journal Article applies to this version, see: <https://www.elsevier.com/about/policies-and-standards/sharing#4-published-journal-article>. Please also note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

© 2026 The Author(s). Published by Elsevier Inc.

Title: Depressive Symptoms and Associated Factors Among Obstetrics and Gynecology First-Year Residents

Authors:

Elena Frank, PhD¹ (ORCID: [0000-0001-7862-9042](#))

Julia Rossi, MPH² (ORCID: [0009-0002-0940-4004](#))

Zhuo Zhao, MS¹ (ORCID: [0000-0001-9339-1896](#))

Karina Pereira-Lima, PhD, MSc³ (ORCID: [0000-0003-4313-743X](#))

Srijan Sen, MD, PhD^{1,4,5} (ORCID: [0000-0003-4495-495X](#))

Abigail Winkel, MD, MHPE² (ORCID: [0000-0001-6149-5842](#))

Helen K Morgan, MD⁶ (ORCID: [0000-0002-6674-5377](#))

¹Michigan Neuroscience Institute, University of Michigan, Ann Arbor

²Department of Obstetrics and Gynecology, Institute for Innovations in Medical Education, New York University Grossman School of Medicine, New York

³Department of Anesthesiology, University of Michigan, Ann Arbor

⁴Eisenberg Family Depression Center, University of Michigan, Ann Arbor

⁵Department of Psychiatry, University of Michigan, Ann Arbor

⁶Departments of Obstetrics and Gynecology and Learning Health Sciences, University of Michigan, Ann Arbor

Address for Correspondence:

Elena Frank, PhD

Michigan Neuroscience Institute

University of Michigan

Arbor Lakes Bldg 1, Fl 2, Ste 2400

Ann Arbor MI 48105-3640

Phone: (734)764-5192

Email: emfrank@med.umich.edu

Funding Sources: This work was supported by the Josiah Macy Jr. Foundation.

Conflict of Interest Statement: The authors report no conflicts of interest.

1 **Word Count (main text): 3245**

2 **Word Count (abstract): 392**

3 **Tweetable Statement:** In a 16-year national cohort study, we found that over one-third of
4 OBGYN interns screened positive for depression on the PHQ-9, but less than one-third of those
5 affected sought mental health treatment.

6 **Short Title:** Depressive Symptoms Among Obstetrics and Gynecology First-Year Residents

7

AJOG at a Glance:

A. Why was this study conducted?

- 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.

B. What are the key findings?

- Over one-third of OBGYN interns screened positive for depression on the PHQ-9 during internship, yet fewer 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.

C. What does this study add to what is already 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.

1 Abstract

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

11 **Objective(s):** To assess the prevalence of depressive symptoms and mental health treatment-
12 seeking among OBGYN first-year residents and identify associated demographic, psychological,
13 and workplace factors.

14 **Study Design:** First-year OBGYN residents enrolled in the longitudinal Intern Health Study
15 between 2009 and 2024 completed surveys at enrollment and quarterly throughout internship.
16 Depressive symptoms were measured using the 9-item Patient Health Questionnaire (PHQ-9).
17 Pearson correlations, χ^2 tests, stepwise linear regression, and generalized estimating equation
18 (GEE) models were used to identify baseline and internship-related predictors of depressive
19 symptoms. Sample weights were applied to address nonrepresentative sampling and attrition
20 biases.

21 **Results:** Of 1,603 enrolled OBGYN interns (86.5% women; median age 27 years), 1,271
22 (79.3%) completed at least one internship assessment and were included in the analysis. During

internship, 35.6% screened positive for depression on the PHQ-9 at one or more assessments, yet fewer 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(s): More than one-third of OBGYN residents screened positive for depression on the PHQ-9 during their first year, yet treatment-seeking remained critically low. Both pre-existing vulnerabilities and modifiable workplace stressors contributed to depressive symptoms. These findings are particularly notable given the female-predominant composition of the OBGYN 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 OBGYN.

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

1 Introduction

2 Medical training is a high-stress period, with trainees in surgical specialties potentially at
3 elevated risk for depression due to more rigid schedules, intensity of training, and persistent
4 mistreatment.^{1,2} Even after the 2011 ACGME duty hours reforms, surgical trainees continue to
5 work more hours and sleep less than their non-surgical counterparts.^{3,4} Obstetrics and
6 gynecology (OBGYN) is the only surgical specialty predominantly composed of women.⁵
7 Women are at heightened risk for depression compared to men⁴ and experience greater
8 increases in depressive symptoms over the course of internship.⁶ Suffering depression is not only
9 a source of distress in residents, increasing the risk of suicide, but has also been associated with
10 increased reports of medical errors and reduced quality of patient care.⁶

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

Despite the importance of addressing mental health challenges in OBGYN 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 email 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 email 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 (see Figure 1 for 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 AAPOR 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 (DSM) diagnosis of major depressive disorder.

The baseline survey also included basic demographic information (age, gender, race/ethnicity, 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 two 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, 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 (NEO-FFI), 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 (age 5 to 15 years) was evaluated using the Risky Families Questionnaire (RFQ).¹⁷ Participants rated 13 childhood experiences (e.g. 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 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, significant 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; 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.

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

4 To reduce potential bias owing to sampling in estimating the changes over time, we used data
5 from the Association of American Medical Colleges (AAMC) on the overall characteristics of
6 first-year OBGYN residents in the US from 2009 to 2024 as the reference target. We performed
7 a 2-step post-stratification raking on the study data to construct weights such that the distribution
8 of cohort year, sex, race and ethnicity in the sample matched the AAMC data distribution. The
9 first step was to generate between-cohort weights ($w1b$) with the raking variable to be cohort
10 year; the second step was to generate weights within each cohort ($w1w$) with the raking variables
11 to be sex, race and ethnicity.

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

18 The final weights were the product of post-stratification weights and attrition weights ($w1b \times$
19 $w1w \times w2$). We applied trimming and standardizing procedures to the final weights to address
20 outliers and to ensure the weighted sample size matched the unweighted sample size. The final
21 analytic weight, incorporating both poststratification weights and attrition weights, was applied

to all primary analyses involving follow-up data, including the linear regression and GEE models.

We used Pearson correlations for continuous measures, and X^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 $P < 0.15$. 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 generalized estimating equation (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 pre-internship baseline survey, 1603 were entering OBGYN programs (median [IQR] age, 27 [26-28] years; 1369 women [86.5%]) (Table 1). Participants represented 191 residency programs across diverse regions of the US (Northeast: 32.1%, Midwest: 22.5%, South: 29.8%, 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-month follow-up time points, respectively. The median PHQ-9 score increased from 2 [1-4] to 6 [3.3-8.8] from pre-internship 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 ($\beta = 3.2$, 95% CI, 3.0-3.4; $p < 0.001$). 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 four quarterly follow-ups and those who completed at least one internship assessment (PHQ-9 score, Mean = 3.14, SD = 3.21 vs. Mean = 2.79, SD = 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 colinearity between variables, 5 variables remained significant: history of depression ($\beta=0.67$, 95% CI, 0.28-1.06; $p < 0.001$), lower baseline depressive symptoms ($\beta=0.59$, 95% CI, 0.52-0.66; $p < 0.001$), higher neuroticism ($\beta=-0.16$, 95% CI, 0.13-0.18; $p < 0.001$), difficult early family environment ($\beta=0.02$, 95% CI, 0.004-0.04; $p=0.018$), and not being in a committed relationship ($\beta=0.49$, 95% CI, 0.13-0.85; $p=0.008$). 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 PHQ-9 during their first year, yet fewer than one-third of those affected sought mental health treatment. Both pre-internship factors (history of depression, baseline depression, neuroticism) and internship-related factors (sleep hours, work hours, 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.^{1,2,3} 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

1 change in depressive symptoms with internship decreasing 24.4% from 2007 to 2019. This
2 change was greater among women than men and among nonsurgical interns. Prior literature also
3 suggests that women physicians experience higher baseline rates of depressive symptoms and
4 greater increases during internship, as well as greater exposure to work-family conflict and
5 structural stressors in training.^{6,7,8} In this context, the high prevalence of those who met criteria
6 for depression combined with low rates of treatment-seeking observed in our cohort may reflect,
7 in part, the broader vulnerability of the female-majority specialty during early training. Thus,
8 highlighting the need for early, proactive identification of at-risk trainees and improved access to
9 confidential, low-barrier mental health care within residency programs.

10 The association between depressive symptoms and modifiable training factors such as sleep
11 deprivation, work hours, and medical errors suggests that interventions targeting residency
12 structure, not only individual-level resilience, may improve trainee well-being. A growing body
13 of systematic reviews and meta-analyses has demonstrated that physician burnout and mental
14 health strain are associated with lower quality of care, increased medical errors, and
15 compromised patient safety.^{18,19} Efforts to address physician burnout at the system level have
16 included hiring scribes and using AI to ease EHR burden, bringing on additional clerical staff,
17 reassigning clinical tasks to other healthcare professionals, and redesigning health IT
18 workflows.^{20,21,22} The use of AI scribes was associated with a decrease in clinician burnout from
19 51.9% to 38.8% after 30 days,²³ while another study found that adding full-time clerical support
20 for physician order entry and previsit planning reduced physicians reporting weekly burnout
21 from 43% to 14% after four months.²⁴ While these approaches have shown promise in addressing
22 burnout, reducing medical errors, and improving quality of care, their effectiveness may vary
23 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.^{6,7,8} 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.

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 health care 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 significant increases in residents' psychotherapy utilization, greater willingness to engage in mental health care, 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 fewer 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 health care 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, 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. Further, future work should further evaluate the uptake and effectiveness of opt-out mental health service models across diverse training environments to increase utilization of

1 mental health services. In addition, studies should explore whether depressive symptoms cluster
2 within training programs or are influenced by cohort-level or peer effects, including the potential
3 for within-institution variation in symptom burden. Longitudinal designs are also needed to
4 assess temporal patterns in depressive symptoms across internship, including whether rates
5 change over time or vary by season. Finally, future research should examine how evolving
6 reproductive health policy environments, including the post-*Dobbs* landscape, shape depressive
7 symptoms and willingness to seek care, as well as inform both individual and structural
8 interventions to support trainee mental health.

9 *Strengths and Limitations*

10 A major strength of our study is the longitudinal design and large national sample of OBGYN
11 first-year residents. However, this study has several limitations. First, as an observational study
12 with convenience sampling, we could not definitively determine the causes of changes in
13 depressive symptoms. There may be additional unaccounted for influential individual,
14 institutional, and societal factors. Second, missing data and participant attrition may have biased
15 the findings, although this risk was lessened by the sample weighting. As invitee specialty was
16 not available for all study cohorts, it was not possible to identify the specific proportion of
17 invited OBGYN interns who opted into the study, only the overall enrollment rate across
18 specialties. Third, depressive symptoms were assessed via self-report rather than a diagnostic
19 interview. Fourth, the sample was restricted to the first year of residency; depressive symptom
20 rates may vary among other groups of trainees and practicing physicians.

21 *Conclusions*

22 Given the current workforce shortage and composition of OBGYN practitioners,¹¹ addressing
23 individual and system-level drivers of physician well-being is critical. The low treatment-seeking

- 1 rate underscores the need for improved access to mental health treatment, potentially through
- 2 opt-out models.^{25,26} Further, structural interventions targeting workload and sleep are essential to
- 3 improve trainee well-being and support long-term workforce sustainability in OBGYN.

- 1 **Additional Contributions:** We would like to acknowledge and thank the first-year residents
- 2 taking part in this research. Without their efforts this study would not be possible.

3

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(6):557-565. doi:[10.1001/archgenpsychiatry.2010.41](https://doi.org/10.1001/archgenpsychiatry.2010.41)
2. Hu YY, Ellis RJ, Hewitt DB, et al. Discrimination, abuse, harassment, and burnout in surgical residency training. *N Engl J Med*. 2019;381(18):1741-1752. doi:[10.1056/NEJMsa1903759](https://doi.org/10.1056/NEJMsa1903759)
3. Hughes TM, Waljee JF, Fang Y, et al. New-onset depression among surgical interns. *JAMA Surg*. 2022;157(6):543-545. doi:[10.1001/jamasurg.2022.0618](https://doi.org/10.1001/jamasurg.2022.0618)
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(1):56-64. doi:[10.7326/M21-1594](https://doi.org/10.7326/M21-1594)
5. Rayburn WF. Diversity in Academic Obstetrics and Gynecology. *Obstet Gynecol Clin North Am*. 2024;51(1):181-191. doi:[10.1016/j.ogc.2023.11.003](https://doi.org/10.1016/j.ogc.2023.11.003)
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(12):1766-1772. doi:[10.1001/jamainternmed.2017.5138](https://doi.org/10.1001/jamainternmed.2017.5138)
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(8):e198340. doi:[10.1001/jamanetworkopen.2019.8340](https://doi.org/10.1001/jamanetworkopen.2019.8340)

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(11):e2134315. doi:[10.1001/jamanetworkopen.2021.34315](https://doi.org/10.1001/jamanetworkopen.2021.34315)
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(5):572-580. doi:[10.4300/JGME-D-23-00554.1](https://doi.org/10.4300/JGME-D-23-00554.1)
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(7):e2323872. doi:[10.1001/jamanetworkopen.2023.23872](https://doi.org/10.1001/jamanetworkopen.2023.23872)
11. Rayburn WF, Xierali IM. Projected shortages and distributional challenges of obstetrician-gynecologists in the united states, 2025-2035. *Obstet Gynecol*. 2026;147(1):4-7. doi:[10.1097/AOG.0000000000006079](https://doi.org/10.1097/AOG.0000000000006079)
12. Cates FE, Aiken CE, Aiken ARA. OB-GYN job satisfaction in a changing landscape of legal restrictions on their practice. *Contraception*. 2026;111455. doi:[10.1016/j.contraception.2026.111455](https://doi.org/10.1016/j.contraception.2026.111455)
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. doi:[10.1016/j.ssmqr.2024.100483](https://doi.org/10.1016/j.ssmqr.2024.100483)
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(1):e2352109. doi:[10.1001/jamanetworkopen.2023.52109](https://doi.org/10.1001/jamanetworkopen.2023.52109)

- 1 15. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity
2 measure. *J Gen Intern Med*. 2001;16(9):606-613. doi:[10.1046/j.1525-](https://doi.org/10.1046/j.1525-1497.2001.016009606.x)
3 [1497.2001.016009606.x](https://doi.org/10.1046/j.1525-1497.2001.016009606.x)
- 4 16. Murray G, Rawlings D, Allen NB, Trinder J. NEO five-factor inventory scores:
5 psychometric properties in a community sample. *Meas Eval Couns Dev*. 2003;36(3):140-
6 149. doi:[10.1080/07481756.2003.11909738](https://doi.org/10.1080/07481756.2003.11909738)
- 7 17. Taylor SE, Way BM, Welch WT, Hilmert CJ, Lehman BJ, Eisenberger NI. Early family
8 environment, current adversity, the serotonin transporter promoter polymorphism, and
9 depressive symptomatology. *Biol Psychiatry*. 2006;60(7):671-676.
10 doi:[10.1016/j.biopsych.2006.04.019](https://doi.org/10.1016/j.biopsych.2006.04.019)
- 11 18. Panagioti M, Panagopoulou E, Bower P, et al. Controlled interventions to reduce burnout
12 in physicians: a systematic review and meta-analysis. *JAMA internal medicine*. 2017
13 Feb;177(2):195-205. doi:[10.1001/jamainternmed.2016.7674](https://doi.org/10.1001/jamainternmed.2016.7674)
- 14 19. Dewa CS, Loong D, Bonato S, et al. The relationship between resident burnout and
15 safety-related and acceptability-related quality of healthcare: a systematic literature
16 review. *BMC medical education*. 2017 Nov 9;17(1):195. doi:[10.1186/s12909-017-1040-y](https://doi.org/10.1186/s12909-017-1040-y)
- 17 20. Gottlieb M, Palter J, Westrick J, et al. Effect of medical scribes on throughput, revenue,
18 and patient and provider satisfaction: A systematic review and meta-analysis. *Ann Emerg*
19 *Med*. 2021;77(2):180-189. doi:[10.1016/j.annemergmed.2020.07.031](https://doi.org/10.1016/j.annemergmed.2020.07.031)

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(5):907-914. doi:[10.1093/jamia/ocaa345](https://doi.org/10.1093/jamia/ocaa345)
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(8):2511-2513. doi:[10.1007/s11606-020-06066-9](https://doi.org/10.1007/s11606-020-06066-9)
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(10):e2534976. doi:[10.1001/jamanetworkopen.2025.34976](https://doi.org/10.1001/jamanetworkopen.2025.34976)
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(5):363-368. doi:[10.14423/SMJ.0000000000000645](https://doi.org/10.14423/SMJ.0000000000000645)
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(2):195-201. doi:[10.4300/JGME-D-23-00460.1](https://doi.org/10.4300/JGME-D-23-00460.1)
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(6):734-737. doi:[10.4300/JGME-D-22-00898.1](https://doi.org/10.4300/JGME-D-22-00898.1)

1 **Tables**2 **Table 1. Baseline and Internship Characteristics of OBGYN First-Year Residents, 2009-**
3 **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 (e.g. 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
Multi-racial		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)*		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)†		24	(18-30)	24 (19-30)
Difficult early family environment: median Risky Families Questionnaire score (IQR)‡		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)*		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
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

1

2

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	0.33
In a committed relationship	-0.02	-0.10	0.05	0.61
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	0.53
Medical errors	0.11	0.04	0.18	0.003

GEE indicates generalized estimating equation; PHQ-9, 9-item Patient Health Questionnaire; CI, confidence interval. 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.

1 **Figures**

2 **Figure 1. Enrollment Flow Diagram**

3 **Figure 2. Frequency of OBGYN Interns with a Positive Screen for Depression on the** 4 **Patient Health Questionnaire-9 Across Internship**

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

7 **Figure 3. Mean PHQ-9 Scores Across OBGYN Internship**

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

Figure 1. Study Flow Diagram



