American Journal of Epidemiology

Vol. 190, No. 7
https://doi.org/10.1093/aje/kwab002
Advance Access publication:
January 11, 2021

## Systematic Review and Meta- and Pooled Analysis

# Is the US Gender Gap in Depression Changing Over Time? A Meta-Regression

**Jonathan M. Platt*, Lisa Bates, Justin Jager, Katie A. McLaughlin, and Katherine M. Keyes**

* Correspondence to Dr. Jonathan M. Platt, Department of Epidemiology, Mailman School of Public Health, Columbia University, 722 W. 168th Street, New York, NY 10032 (e-mail: jmp2198@cumc.columbia.edu).

*Initially submitted February 26, 2020; accepted for publication October 6, 2020.*

The depression gap refers to higher rates of depression among women than men. Change in the depression gap over time might elucidate social causes of this disparity—such as unequal college attendance or employment status. We conducted a meta-regression analysis to estimate variation in the depression gap over time by age, accounting for potential sources of variation between studies. Electronic databases and bibliographies were searched for English-language studies from January 1980 through October 2019; 144 independent estimates from US-representative samples met selection criteria ($n$ = 813,189). The depression gap was summarized as prevalence ratios among studies using diagnostic instruments and as standardized mean differences among symptom-based studies. Primary study measures were baseline study year (range, 1982–2017) and age (age groups ranging, in years, from 10–59 and 60 or older). Compared with respondents aged ≥60 years, depression prevalence was greater among respondents aged 10–19 (prevalence ratio = 1.26, 95% confidence interval: 1.02, 1.56). Over time, the depression gap did not change among adults, but it increased among adolescents (age-by-time interaction prevalence ratio = 1.05, 95% confidence interval: 1.01, 1.08). Results were similar for symptom-based studies. The present study finds no evidence of a change in the depression gender gap for US adults; however, the gap increased among adolescents. Greater attention to factors driving this widening disparity in adolescent depression is needed.

depression; depressive symptoms; gender; health disparities; time trends; United States

Abbreviations: CI, confidence interval; PR, prevalence ratio.

***Editor's note:*** *An invited commentary on this article appears on page 1207.*

Major depressive disorder is the leading cause of disability among Americans ages 15–44 years (1) and is more likely to affect women than men (2). This pattern, hereafter referred to as the depression gap, reflects meaningful differences in depression and is not solely an artifact of gender differences in reporting mental health symptoms or seeking treatment (3, 4). Also, even though the quantitative surveys providing evidence regarding the depression gap typically rely on binary categories that do not differentiate between sex assigned at birth and gender expression, the gap is typically described using the term gender. Given that caveat, we use gender throughout the present study.

The depression gap emerges in early adolescence, remains relatively stable throughout adulthood, and then decreases at later ages (5). Biological (6) and social stress (7) mechanisms have been explored to explain the gap, with the most robust evidence to date supporting social stress. As applied to gender, social stress theory suggests that gender might influence stress exposure and responses (8). In particular, women traditionally have had fewer opportunities in attaining higher education and full-time employment, which might act as social stressors (9). From an early age, women are typically socialized, through gender norms, to respond to stressors in depressogenic ways (10, 11). These factors might increase women's depression risk, and explain gender differences in depression (9). If so, changes in women's social positions, and therefore changes in these factors, should change the depression gap in turn.

Since the mid-20th century, education (12) and employment (13, 14) opportunities have become increasingly available to women. These changes in gendered social positions

likely reflect broader changes in norms and the process of gender socialization, which might influence both exposure and response to stressors, and might decrease the depression gap in adults in turn. Among adolescents, depressive symptoms have been increasing since 2012, to a greater extent among girls compared with boys (15). While the underlying factors contributing to the increase remain only speculative, the increase among young women might, unlike among adults, suggest an increasing gender gap.

Available evidence suggests that the depression gap might be changing (16, 17) but is inconclusive, in part due to 3 limitations. First, follow-up periods in single longitudinal studies are often too short to identify temporal trends in depression. Second, while existing studies suggest that the depression gap might vary over time, there could also be variation by age across time. Examining variation by both age and time is necessary to identify any temporal variation due to social change. Among reviews that have directly accounted for age in assessing temporal variation in the depression gap, most have focused on a single age group or developmental period (18, 19). A wide time span with age groups across the life course is necessary in order to fully characterize variation in the depression gap by both age and time. Third, less attention has been paid to gender differences in levels of depressive symptoms, which might be distinct from diagnostic depression categories (20).

We conducted a systematic review, meta-analysis, and meta-regression to characterize changes in the depression gap over time and across the life course. First, studies of gender differences in depression in recent decades were identified and summarized. Second, data from the systematic review were extracted to form the analytical sample of the meta-regression, which estimated the variation in the gap over time by age and accounted for other potential sources of variation between studies. Trends in the depression gap were considered separately based on diagnostic versus symptom-based depression tools, to explore whether variation has been different at a diagnostic threshold versus total depression symptoms.

## METHODS

### Identification of studies

The literature search focused on peer-reviewed research published in English-language journals between January 1980 and October 2019. The year 1980 was chosen as the lower limit because it coincided with changes to women's social positions in the United States that had been ongoing since the mid-20th century. The year 1980 also represented the introduction of the *Diagnostic and Statistical Manual of Mental Disorders*, version III, which was used to estimate the US population prevalence of psychiatric disorders in community-based psychiatric epidemiologic surveys (21, 22). Only studies of the US population were included, given the background of changing gendered social positions in the United States. Finally, the search focused on studies based on nationally representative sampling frames to avoid bias from gender-specific selection factors (e.g., clinical samples (23)). Also, US population-level samples typically have large sample sizes that maximize statistical precision of depression gap estimates.

The literature search and study-selection flow chart is detailed in Figure 1. The initial search included 5 electronic databases: PubMed, JSTOR, Embase, PsychInfo, and Scopus. The initial search yielded 1,007 potential abstracts. Bibliographies of related reviews and meta-analyses were also searched, which yielded 20 additional records.

After removing 218 duplicate studies, 809 abstracts were screened in more detail and additional studies were excluded based on the following criteria: No quantitative data were presented (e.g., qualitative study, narrative review); the sample included nonhuman subjects; gender-specific estimates were not presented; or depression measures were not based on a symptom-level interview (e.g., self-reported doctor diagnosed depression).

A second reviewer independently screened the 809 abstracts. Agreement between the 2 reviewers was very good ($\kappa = 0.827$, 95% confidence interval (CI): 0.788, 0.867) (24), and the reviewers further discussed any conflicting judgments to reach consensus. A resulting 452 studies met exclusion criteria and were removed.

The full text of the 357 remaining studies was reviewed in more detail. In the instance that the same data set was used for multiple studies, only the study with the most complete sample was included (i.e., the fewest stated restrictions to derive the analytical sample from the full study sample), to ensure the independence between depression gap estimates. The reasons for exclusion of the full-text reviewed studies were: study design (e.g., case control, sampling based on depression status) ($n = 158$), a non–nationally representative sampling strategy ($n = 123$), and duplicate data source ($n = 35$) (see Figure 1).

In total, 41 studies were included. Several of these studies included multiple estimates for different age groups, and each group was considered as an independent estimate (range: 1–17 estimates per study). Among longitudinal studies, only baseline data were included to avoid issues of within-study correlation of depression gap estimates and potential selection bias from attrition. Several studies did not measure depression at baseline but included it in later interviews. Estimates from the first follow-up interview where depression was measured were included, and the proportion lost to follow-up was extracted to consider the potential for selection bias. The full meta-analytical data set contained 144 independent estimates from nationally representative samples. The total sample size was 813,189 (52% women and girls).

### Data abstraction

For each estimate, the following information was collected: full citation, study name, baseline study year, sample size by gender, age range, depression effect measure, effect estimate and variance, depression instrument, and the period of symptom recall.

### Effect measures

The depression gap was summarized as a prevalence ratio among studies that reported depression based on a diagnostic

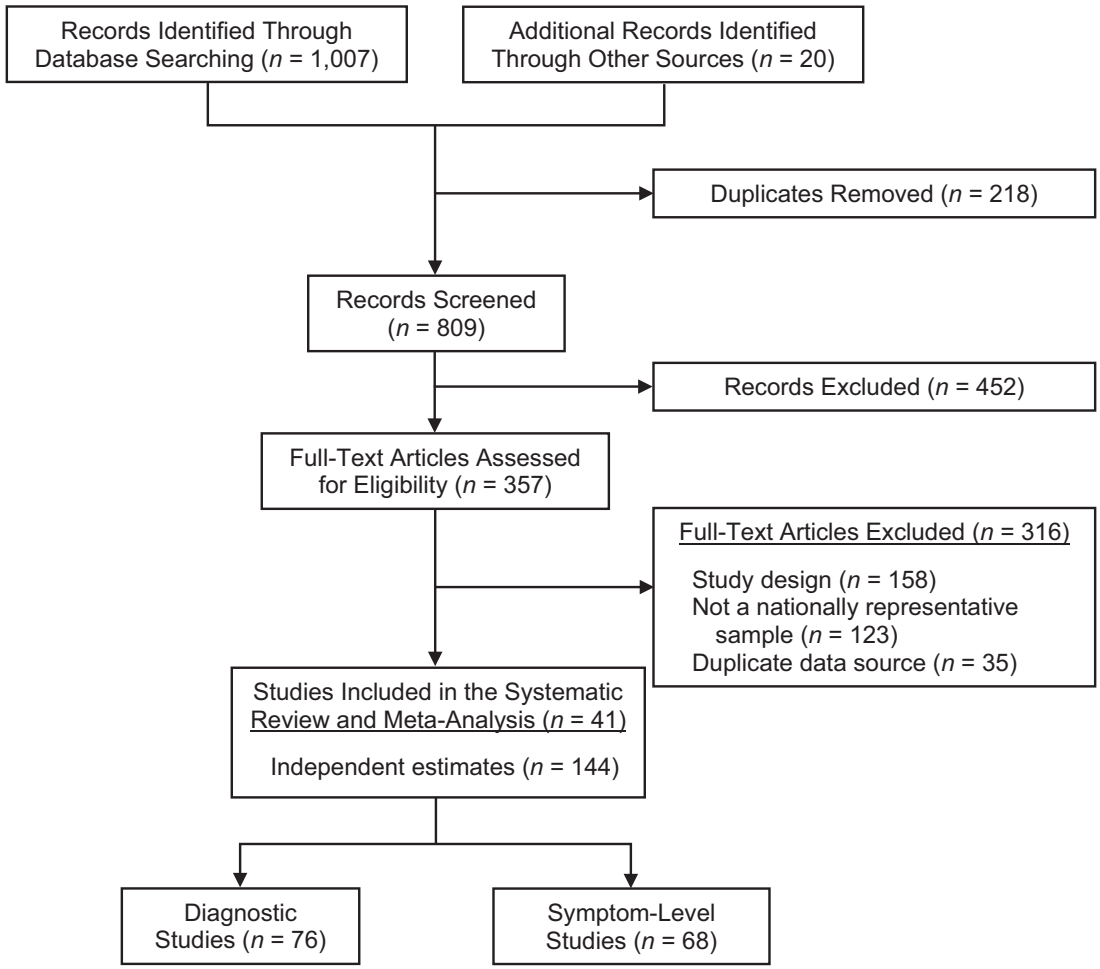


**Figure 1.** Literature search and study selection flowchart for an analysis of the gender gap in depression, United States, 1980–2019. Note: The general search strategy included: ("gender" OR "sex") AND ("male" AND "female") AND ("depress∗" OR "distress" OR "demoraliz∗" OR "internaliz∗"). The asterisk denotes a stem that might encompass various forms a word (e.g., depress∗ = depressive, depression, depressed, etc). Search terms were optimized using Medical Subject Headings terms and adapted for each database (PubMed, JSTOR, Embase, PsychInfo, and Scopus).

threshold (i.e., diagnostic studies). Among symptom-based studies, the depression gap was estimated as a standardized mean difference (calculated as the mean depression score among women minus the mean score among men, divided by the pooled standard deviation). Null, small, medium, and large effect sizes corresponded with standardized mean difference = 0 < 0.2, 0.2 < 0.5, 0.5 < 0.8, and 0.8 < 1.0, respectively (25). Studies were weighted by the inverse of the standard error of the log prevalence ratio and standardized mean difference, respectively (26).

### Independent variables

Continuous study year at baseline was the main independent variable (range 1982–2017) to estimate temporal variation in the depression gap.

Age was considered as an effect modifier of time, categorized in years: 10–19 (i.e., childhood/adolescence), 20–39 (i.e., early adulthood), 40–59 (i.e., middle adulthood), and 60 or older (i.e., older adults). Groupings were chosen in order to capture meaningful life periods, while also ensuring large enough samples within each group. Studies with wider age ranges (e.g., ages 18–65) were included in the descriptive analysis but not the meta-regression models (3 diagnostic studies (3.9%) and 6 symptom studies (8.8%)). Note that age across time also indexes birth cohort, thus age in the present review is a proxy for both age and birth cohort.

The depression instrument was considered as a confounding variable. Symptom-scale instruments included Children's Depression Inventory, 9-item Patient Health Questionnaire, or other versus the Center for Epidemiologic Studies–Depression scale. Diagnostic instruments included *Diagnostic and Statistical Manual of Mental Disorders* (DSM), version III/version III revised, or other diagnostic versus DSM version IV/version IV-revised. Only 1 study used DSM version V instruments, so it was grouped with DSM version IV instruments.

### Publication bias

We considered the potential for publication bias (27). However, given that the depression gap was often not the main focus of these articles, the magnitude of the depression gap would likely have little influence over whether a study was published. Nonetheless, to explore potential bias, a funnel plot was estimated for each set of studies. The degree of bias was tested using Egger's test (28). Additionally, the trim-and-fill procedure was used to estimate what the actual effect size would have been in the absence of any publication bias (29).

### Analysis

In order to characterize variation in the depression gap, the review was structured to estimate cross-study variation over time with meta-regression models. For each depression gap estimate, the baseline study year formed the main independent variable in the meta-regression model, representing change in the depression gap over time, accounting for differences in age and other potential sources of variation.

First, a descriptive analysis summarized the data sources, study designs, and analytical variables of all included studies. Additionally, a pooled depression gap was estimated to summarize the depression gap across all studies in the analytical sample, with random effects confidence intervals and prediction intervals (30). Second, meta-regression models estimated the average effects of time, age, the interaction between time and age, and instrument in depression gap estimates, using maximum likelihood estimation with robust standard errors. Analyses were implemented with "meta" (31) and "metafor" (32) packages in R (R Foundation for Statistical Computing, Vienna, Austria), version 3.5.1 (33).

### Multiple imputation

For 27 diagnostic study estimates (36%), data needed to compute the standard error of the effect estimate were not reported (i.e., only an unadjusted PR was reported). To minimize the amount of information lost due to missing data, models were estimated with imputed variance parameters from 100 imputed data sets using chained equations, combined with corrected standard errors, averaging coefficient vectors, variance-covariance matrices, and adding a nonnegative correction to variance-covariance matrices inversely proportional to the predictive ability of the imputation models, effectively widening confidence intervals where missing data values are poorly predicted by observed data (34). The large number of imputation models was chosen to achieve stability of imputed estimates and all abstracted study variables were used to imputed missing data. Recent simulations have reported that multiple imputations in meta-regression models are unbiased when missing values are weighting variables (i.e., within-study standard errors), rather than predictor variables (35). Imputed model estimates were compared with complete case-models to examine the degree of their robustness to missing data.

## RESULTS

### Descriptive summary

Tables 1 and 2 (and Web Tables 1 and 2, available at https://doi.org/10.1093/aje/kwab002) provide the descriptive details of the diagnostic and symptom-based studies that comprised the analytical sample. A descriptive summary of the study designs and sampling procedures can be found in Web Appendix 1.

Table 3 summarizes the distributions of all analytical variables. Of the 144 total estimates, 76 measured the depression gap with a diagnostic instrument and 68 measured the gap with symptom scores. Overall, the study year at baseline ranged from 1982 to 2017. The respondent ages ranged from 10 to 99 years old. Estimates from samples of ages 10–19 represented 35.5% of diagnostic and 48.6% of symptom-based estimates. Depression was assessed using criteria from *Diagnostic and Statistical Manual of Mental Disorders*, version IV/version IV revised, in 71 diagnostic studies (93.4%), and the Center for Epidemiologic Studies–Depression scale was used to measure depression in 42 symptom-based studies (61.7%). Among diagnostic studies, 97.4% of studies assessed past-year depression (2 studies assessed lifetime depression (36, 37)), so symptom period was not included as an independent variable; a sensitivity analysis included only studies of past-year depression to determine whether the meta-regression estimates were biased by the few studies with a longer recall period.

The effect sizes of all diagnostic and symptom-based depression gap estimates and a pooled summary depression gap are presented in Web Figures 1 (diagnostic studies) and 2 (symptom-based studies) and described in Web Appendix 1. Among diagnostic studies, prevalence ratios ranged from 1.26 (95% CI: 0.99, 1.59) to 4.23 (95% CI: 3.37, 5.31), and the pooled summary prevalence ratio was 2.01 (95% CI: 1.88, 2.14). Among symptom-based depression gap studies, standardized mean differences ranged from −0.12 (95% CI: −0.4, 0.16) to 0.59 (95% CI: 0.51, 0.67); the pooled summary standardized mean difference was 0.22 (95% CI: 0.19, 0.25), indicating a small effect size.

### Meta-regression

Meta-regression results are presented in Table 4. Main effects among diagnostic studies were estimated in model 1a. The depression gap with all model variables at their reference levels was 2.27 (95% CI: 1.48, 3.05). Overall, there was no evidence of change in the depression gap over time. The age effect was most pronounced among ages 10–19 years (prevalence ratio (PR) = 1.26, 95% CI: 1.02, 1.56), compared with the referent (i.e., respondents aged ≥60 years). Based on the exponentiated combined intercept and age coefficients, the depression gap was 2.86 among ages 10–19. The depression gap did not differ for any other age groups versus the referent.

Model 2a tested interaction between age group and study year. The interaction term for youngest age group was elevated (PR = 1.05, 95% CI: 1.01, 1.08), suggesting that, compared with ages ≥60, the depression gap had increased

**Table 1.** Studies of Gender Differences in Diagnostic Depression, Measured as Prevalence Ratios, United States

| First Author, Year (Reference No.) | Baseline Year | PR | SE | Age Minimum | Age Maximum[a] | No. of Men | No. of Women | Interview and Instrument | Data Source |
|---|---|---|---|---|---|---|---|---|---|
| Kessler, 1993 (54) | 1990 | 3.15 | 1.455 | 15 | 24 | 1,010 | 990 | CIDI DSM-III-R | NCS |
| | 1990 | 1.50 | 0.249 | 25 | 34 | 1,231 | 1,207 | CIDI DSM-III-R | NCS |
| | 1990 | 1.89 | 0.284 | 35 | 44 | 1,108 | 1,086 | CIDI DSM-III-R | NCS |
| | 1990 | 1.52 | 0.299 | 45 | 54 | 740 | 726 | CIDI DSM-III-R | NCS |
| Alaimo, 2002 (36) | 1991 | 2.28 | 0.968 | 15 | 16 | 365 | 389 | DIS | NHANES III |
| Grant, 1995 (70) | 1992 | 1.27 | 0.040 | 18 | 99 | 17,819 | 25,043 | AUDADIS DSM-IV[b] | NLAES |
| Dawson, 1997 (71) | 1992 | 1.60 | 0.065 | 18 | 99 | 17,819 | 25,043 | AUDADIS DSM-IV | NLAES |
| Kessler, 2010 (72) | 2001 | 1.58 | 0.196 | 18 | 34 | 1,375 | 1,658 | CIDI DSM-IV | NCS-R |
| | 2001 | 2.73 | 0.483 | 35 | 49 | 1,342 | 1,522 | CIDI DSM-IV | NCS-R |
| | 2001 | 1.48 | 0.247 | 50 | 64 | 854 | 1,068 | CIDI DSM-IV | NCS-R |
| | 2001 | 1.71 | 0.297 | 65 | 99 | 564 | 894 | CIDI DSM-IV | NCS-R |
| CBHSQ, 2004 (73) | 2004 | 2.62 | 0.131 | 12 | 17 | 11,363 | 10,938 | CIDI DSM-IV | NSDUH |
| CBHSQ, 2005 (74) | 2005 | 2.96 | 0.154 | 12 | 17 | 11,378 | 11,156 | CIDI DSM-IV | NSDUH |
| | 2005 | 2.09 | 0.104 | 18 | 25 | 10,697 | 10,444 | CIDI DSM-IV | NSDUH |
| | 2005 | 1.69 | 0.096 | 26 | 49 | 7,823 | 9,132 | CIDI DSM-IV | NSDUH |
| | 2005 | 2.00 | 0.260 | 50 | 99 | 3,142 | 3,420 | CIDI DSM-IV | NSDUH |
| CBHSQ, 2006 (75) | 2006 | 2.81 | 0.151 | 12 | 17 | 11,718 | 11,153 | CIDI DSM-IV | NSDUH |
| | 2006 | 1.81 | 0.096 | 18 | 25 | 9,158 | 11,526 | CIDI DSM-IV | NSDUH |
| | 2006 | 1.73 | 0.100 | 26 | 49 | 7,431 | 8,606 | CIDI DSM-IV | NSDUH |
| | 2006 | 1.67 | 0.217 | 50 | 99 | 2,888 | 3,804 | CIDI DSM-IV | NSDUH |
| CBHSQ, 2007 (76) | 2007 | 2.59 | 0.135 | 12 | 17 | 11,524 | 10,909 | CIDI DSM-IV | NSDUH |
| | 2007 | 1.97 | 0.100 | 18 | 25 | 10,645 | 11,542 | CIDI DSM-IV | NSDUH |
| | 2007 | 1.72 | 0.097 | 26 | 49 | 7,770 | 9,114 | CIDI DSM-IV | NSDUH |
| | 2007 | 1.67 | 0.214 | 50 | 99 | 2,857 | 3,509 | CIDI DSM-IV | NSDUH |
| CBHSQ, 2008 (77) | 2008 | 2.91 | 0.155 | 12 | 17 | 11,517 | 11,029 | CIDI DSM-IV | NSDUH |
| | 2008 | 2.11 | 0.103 | 18 | 25 | 11,166 | 12,039 | CIDI DSM-IV | NSDUH |
| | 2008 | 1.49 | 0.091 | 26 | 49 | 7,440 | 8,936 | CIDI DSM-IV | NSDUH |
| | 2008 | 2.14 | 0.306 | 50 | 99 | 2,996 | 3,613 | CIDI DSM-IV | NSDUH |
| CBHSQ, 2009 (78) | 2009 | 2.49 | 0.129 | 12 | 17 | 11,520 | 11,106 | CIDI DSM-IV | NSDUH |
| | 2009 | 1.93 | 0.096 | 18 | 25 | 11,104 | 11,900 | CIDI DSM-IV | NSDUH |
| | 2009 | 1.71 | 0.104 | 26 | 49 | 7,591 | 8,729 | CIDI DSM-IV | NSDUH |

**Table continues**

**Table 1.** Continued

| First Author, Year (Reference No.) | Baseline Year | PR | SE | Age Minimum | Age Maximum[a] | No. of Men | No. of Women | Interview and Instrument | Data Source |
|---|---|---|---|---|---|---|---|---|---|
| CBHSQ, 2009 (78) | 2009 | 1.67 | 0.213 | 50 | 99 | 3,060 | 3,690 | CIDI DSM-IV | NSDUH |
| CBHSQ, 2010 (79) | 2010 | 2.71 | 0.146 | 12 | 17 | 11,140 | 10,820 | CIDI DSM-IV | NSDUH |
| | 2010 | 2.27 | 0.091 | 18 | 25 | 17,283 | 16,788 | CIDI DSM-IV | NSDUH |
| CBHSQ, 2011 (80) | 2011 | 2.69 | 0.138 | 12 | 17 | 12,028 | 11,482 | CIDI DSM-IV | NSDUH |
| | 2011 | 1.95 | 0.077 | 18 | 25 | 17,178 | 17,123 | CIDI DSM-IV | NSDUH |
| Verplaetse, 2016 (81) | 2012 | 2.01 | 0.218 | 18 | 99 | 15,715 | 20,386 | DSM-V | NESARC 3 |
| CBHSQ, 2013 (82) | 2013 | 2.07 | 0.405 | 12 | 12 | 1,824 | 1,713 | CIDI DSM-IV | NSDUH |
| | 2013 | 3.42 | 0.523 | 13 | 13 | 1,963 | 1,849 | CIDI DSM-IV | NSDUH |
| | 2013 | 4.23 | 0.550 | 14 | 14 | 2,026 | 1,865 | CIDI DSM-IV | NSDUH |
| | 2013 | 3.34 | 0.366 | 15 | 15 | 1,882 | 1,868 | CIDI DSM-IV | NSDUH |
| | 2013 | 2.54 | 0.260 | 16 | 16 | 1,940 | 1,890 | CIDI DSM-IV | NSDUH |
| | 2013 | 2.70 | 0.283 | 17 | 17 | 1,914 | 1,760 | CIDI DSM-IV | NSDUH |
| | 2013 | 1.96 | 0.095 | 18 | 25 | 10,671 | 11,543 | CIDI DSM-IV | NSDUH |
| | 2013 | 1.71 | 0.251 | 26 | 29 | 1,376 | 1,603 | CIDI DSM-IV | NSDUH |
| | 2013 | 1.26 | 0.171 | 30 | 34 | 1,529 | 1,802 | CIDI DSM-IV | NSDUH |
| | 2013 | 1.64 | 0.269 | 35 | 39 | 1,317 | 1,562 | CIDI DSM-IV | NSDUH |
| | 2013 | 1.57 | 0.237 | 40 | 44 | 1,437 | 1,671 | CIDI DSM-IV | NSDUH |
| | 2013 | 1.31 | 0.194 | 45 | 49 | 1,440 | 1,613 | CIDI DSM-IV | NSDUH |
| | 2013 | 1.43 | 0.282 | 50 | 54 | 837 | 951 | CIDI DSM-IV | NSDUH |
| | 2013 | 1.54 | 0.374 | 55 | 59 | 711 | 909 | CIDI DSM-IV | NSDUH |
| | 2013 | 1.41 | 0.376 | 60 | 64 | 674 | 719 | CIDI DSM-IV | NSDUH |
| | 2013 | 3.53 | 1.399 | 65 | 99 | 1,302 | 1,659 | CIDI DSM-IV | NSDUH |
| CBHSQ, 2017 (83) | 2017 | 2.59 | 0.610 | 12 | 12 | 1,329 | 1,269 | CIDI DSM-IV | NSDUH |
| | 2017 | 4.04 | 0.696 | 13 | 13 | 1,507 | 1,423 | CIDI DSM-IV | NSDUH |
| | 2017 | 3.63 | 0.507 | 14 | 14 | 1,492 | 1,385 | CIDI DSM-IV | NSDUH |
| | 2017 | 3.68 | 0.416 | 15 | 15 | 1,460 | 1,427 | CIDI DSM-IV | NSDUH |
| | 2017 | 2.46 | 0.246 | 16 | 16 | 1,508 | 1,389 | CIDI DSM-IV | NSDUH |
| | 2017 | 2.20 | 0.206 | 17 | 17 | 1,419 | 1,418 | CIDI DSM-IV | NSDUH |
| | 2017 | 2.10 | 0.296 | 18 | 18 | 1,070 | 1,036 | CIDI DSM-IV | NSDUH |
| | 2017 | 1.90 | 0.263 | 19 | 19 | 976 | 1,002 | CIDI DSM-IV | NSDUH |
| | 2017 | 1.51 | 0.180 | 20 | 20 | 973 | 954 | CIDI DSM-IV | NSDUH |
| | 2017 | 1.59 | 0.208 | 21 | 21 | 922 | 984 | CIDI DSM-IV | NSDUH |

**Table continues**

**Table 1.** Continued

| First Author, Year (Reference No.) | Baseline Year | PR | SE | Age Minimum | Age Maximum[a] | No. of Men | No. of Women | Interview and Instrument | Data Source |
|---|---|---|---|---|---|---|---|---|---|
| CBHSQ, 2017 (83) | 2017 | 2.21 | 0.319 | 22 | 22 | 1,000 | 1,033 | CIDI DSM-IV | NSDUH |
| | 2017 | 1.85 | 0.253 | 23 | 23 | 1,006 | 1,155 | CIDI DSM-IV | NSDUH |
| | 2017 | 1.55 | 0.205 | 24 | 24 | 975 | 1,139 | CIDI DSM-IV | NSDUH |
| | 2017 | 1.48 | 0.206 | 25 | 25 | 1,061 | 1,183 | CIDI DSM-IV | NSDUH |
| | 2017 | 1.83 | 0.196 | 26 | 29 | 2,117 | 2,580 | CIDI DSM-IV | NSDUH |
| | 2017 | 1.36 | 0.139 | 30 | 34 | 2,631 | 3,088 | CIDI DSM-IV | NSDUH |
| | 2017 | 2.06 | 0.267 | 35 | 39 | 2,231 | 2,551 | CIDI DSM-IV | NSDUH |
| | 2017 | 1.91 | 0.249 | 40 | 44 | 1,945 | 2,387 | CIDI DSM-IV | NSDUH |
| | 2017 | 1.67 | 0.214 | 45 | 49 | 2,075 | 2,450 | CIDI DSM-IV | NSDUH |
| | 2017 | 1.30 | 0.274 | 50 | 54 | 901 | 1,093 | CIDI DSM-IV | NSDUH |
| | 2017 | 1.64 | 0.342 | 55 | 59 | 931 | 1,138 | CIDI DSM-IV | NSDUH |
| | 2017 | 1.87 | 0.448 | 60 | 64 | 948 | 1,013 | CIDI DSM-IV | NSDUH |
| | 2017 | 1.57 | 0.343 | 65 | 99 | 2,077 | 2,381 | CIDI DSM-IV | NSDUH |

Abbreviations: AUDADIS, Alcohol Use Disorder and Associated Disabilities Interview Schedule; CBHSQ, Center for Behavioral Health Statistics and Quality; CIDI, Composite International Diagnostic Interview; DIS, Diagnostic Interview Scale; DSM, *Diagnostic and Statistical Manual of Mental Disorders*; NCS, National Comorbidity Survey; NCS-R, National Comorbidity Survey Replication; NESARC, National Epidemiologic Survey of Alcoholism and Related Conditions; NHANES, National Health and Nutrition Epidemiologic Survey; NLAES, National Longitudinal Alcohol Epidemiologic Survey; NSDUH, National Survey of Drug Use and Health; PR, prevalence ratio; SE, standard error.

[a] When the sample age range was described as all ages (e.g., 18 years and up), upper bound was coded as 99. All studies used cross-sectional study design.

[b] Lifetime recall period. All other studies used prior-year recall.

**Table 2.** Studies of Gender Differences in Symptom-Based Depression, Measured as Standardized Mean Differences, United States

| Author, Year (Reference No.) | Baseline Year | SMD | SE | Age Minimum | Age Maximum[a] | No. of Men | No. of Women | Instrument | Study Design[b] | Data Source | Retention[b] | Recall Period[b] |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Ferketich, 2000 (84) | 1982 | 0.26 | 0.02 | 30 | 99 | 2,886 | 5,007 | CESD | C | NHANES I | | 1 |
| Everson-Rose, 2004 (85) | 1986 | 0.31 | 0.02 | 24 | 34 | 333 | 407 | CESD | BL | ACLS | | 1 |
| | 1986 | 0.23 | 0.05 | 35 | 44 | 228 | 363 | CESD | BL | ACLS | | 1 |
| | 1986 | −0.01 | 0.02 | 45 | 54 | 168 | 222 | CESD | BL | ACLS | | 1 |
| | 1986 | 0.08 | 0.05 | 55 | 64 | 251 | 434 | CESD | BL | ACLS | | 1 |
| | 1986 | 0.12 | 0.05 | 65 | 74 | 239 | 526 | CESD | BL | ACLS | | 1 |
| | 1986 | 0.23 | 0.05 | 75 | 99 | 139 | 307 | CESD | BL | ACLS | | 1 |
| Inaba, 2005 (86) | 1994 | 0.27 | 0.05 | 28 | 39 | 1,372 | 1,413 | CESD | BL | NSFH-2 | | 1 |
| | 1994 | 0.22 | 0.05 | 40 | 49 | 1,013 | 987 | CESD | BL | NSFH-2 | | 1 |
| | 1994 | 0.38 | 0.05 | 50 | 59 | 594 | 716 | CESD | BL | NSFH-2 | | 1 |
| | 1994 | 0.29 | 0.04 | 60 | 78 | 856 | 1,220 | CESD | BL | NSFH-2 | | 1 |
| Marmorstein, 2009 (87) | 1995 | 0.12 | 0.02 | 12 | 12 | 262 | 329 | CESD | BL | Add Health | | 1 |
| | 1995 | 0.22 | 0.02 | 13 | 13 | 1,039 | 1,218 | CESD | BL | Add Health | | 1 |
| | 1995 | 0.30 | 0.02 | 14 | 14 | 1,319 | 1,472 | CESD | BL | Add Health | | 1 |
| | 1995 | 0.34 | 0.02 | 15 | 15 | 1,778 | 1,883 | CESD | BL | Add Health | | 1 |
| | 1995 | 0.31 | 0.02 | 16 | 16 | 2,061 | 1,991 | CESD | BL | Add Health | | 1 |
| | 1995 | 0.19 | 0.02 | 17 | 17 | 1,981 | 1,940 | CESD | BL | Add Health | | 1 |
| | 1995 | 0.21 | 0.05 | 18 | 18 | 1,512 | 1,427 | CESD | BL | Add Health | | 1 |
| | 1995 | 0.34 | 0.05 | 19 | 19 | 237 | 159 | CESD | BL | Add Health | | 1 |
| Neumark-Sztainer, 2000 (88) | 1997 | 0.26 | 0.04 | 10 | 10 | 239 | 267 | CDI | C | CFS | | 1 |
| | 1997 | 0.02 | 0.02 | 11 | 11 | 254 | 305 | CDI | C | CFS | | 1 |
| | 1997 | 0.04 | 0.05 | 12 | 12 | 386 | 461 | CDI | C | CFS | | 1 |
| | 1997 | 0.29 | 0.05 | 13 | 13 | 420 | 484 | CDI | C | CFS | | 1 |
| | 1997 | 0.22 | 0.05 | 14 | 14 | 370 | 462 | CDI | C | CFS | | 1 |
| | 1997 | 0.31 | 0.02 | 15 | 15 | 361 | 503 | CDI | C | CFS | | 1 |
| | 1997 | 0.32 | 0.04 | 16 | 16 | 399 | 497 | CDI | C | CFS | | 1 |
| | 1997 | 0.25 | 0.02 | 17 | 17 | 314 | 372 | CDI | C | CFS | | 1 |
| Mumford, 2013 (89) | 2000 | 0.48 | 0.02 | 15 | 15 | 815 | 765 | MHI-D | L (BL = 1997) | NLSY97 | 88 | 1 |
| | 2000 | 0.33 | 0.02 | 16 | 16 | 819 | 774 | MHI-D | L (BL = 1997) | NLSY97 | 88 | 1 |

Table continues

**Table 2.** Continued

| Author, Year (Reference No.) | Baseline Year | SMD | SE | Age Minimum | Age Maximum[a] | No. of Men | No. of Women | Instrument | Study Design[b] | Data Source | Retention[b] | Recall Period[b] |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Mumford, 2013 (89) | 2000 | 0.29 | 0.05 | 17 | 17 | 811 | 773 | MHI-D | L (BL = 1997) | NLSY97 | 88 | 1 |
| | 2000 | 0.22 | 0.02 | 18 | 18 | 766 | 767 | MHI-D | L (BL = 1997) | NLSY97 | 88 | 1 |
| | 2000 | 0.23 | 0.05 | 19 | 19 | 657 | 681 | MHI-D | L (BL = 1997) | NLSY97 | 88 | 1 |
| Song, 2011 (90) | 2005 | 0.30 | 0.04 | 21 | 64 | 167 | 188 | CESD | C | Study-specific | | 2 |
| | 2005 | 0.21 | 0.02 | 21 | 64 | 187 | 225 | CESD | C | Study-specific | | 2 |
| | 2005 | 0.11 | 0.04 | 21 | 64 | 939 | 1,124 | CESD | C | Study-specific | | 2 |
| Shiovitz-Ezra, 2009 (91) | 2005 | 0.23 | 0.05 | 57 | 64 | 521 | 484 | CESD | C | NSHAP | | 2 |
| | 2005 | 0.16 | 0.04 | 65 | 74 | 543 | 537 | CESD | C | NSHAP | | 2 |
| | 2005 | 0.09 | 0.05 | 75 | 85 | 373 | 499 | CESD | C | NSHAP | | 2 |
| Haroz, 2014 (92) | 2006 | 0.16 | 0.14 | 11 | 12 | 95 | 99 | CESD-10R | C | Growing up With Media | | 1 |
| | 2006 | 0.49 | 0.10 | 13 | 14 | 201 | 191 | CESD-10R | C | Growing up With Media | | 1 |
| | 2006 | 0.11 | 0.11 | 15 | 17 | 192 | 172 | CESD-10R | C | Growing up With Media | | 1 |
| | 2009 | 0.47 | 0.06 | 13 | 14 | 585 | 785 | CESD-10R | C | Growing up With Media | | 1 |
| | 2009 | 0.27 | 0.05 | 15 | 17 | 856 | 1,096 | CESD-10R | C | Teen Health and Tech | | 1 |
| | 2009 | 0.22 | 0.04 | 18 | 18 | 954 | 1,404 | CESD-10R | C | Teen Health and Tech | | 1 |
| | 2006 | 0.25 | 0.14 | 18 | 18 | 94 | 106 | CESD-10R | C | Teen Health and Tech | | 1 |
| Wang, 2010 (93) | 2006 | 0.29 | 0.05 | 11 | 11 | 1,164 | 1,186 | DFB | C | HBSC | | 1 |
| | 2006 | 0.44 | 0.02 | 12 | 12 | 892 | 951 | DFB | C | HBSC | | 1 |
| | 2006 | 0.49 | 0.02 | 13 | 13 | 789 | 997 | DFB | C | HBSC | | 1 |
| | 2006 | 0.51 | 0.02 | 14 | 14 | 721 | 742 | DFB | C | HBSC | | 1 |
| | 2006 | 0.59 | 0.02 | 15 | 15 | 793 | 804 | DFB | C | HBSC | | 1 |
| Oksuzyan, 2010 (94) | 2006 | 0.13 | 0.02 | 50 | 54 | 640 | 1,013 | CESD | L (BL = 1992) | HRS | 85 | 1 |
| | 2006 | 0.05 | 0.05 | 55 | 59 | 1,051 | 1,472 | CESD | L (BL = 1992) | HRS | 85 | 1 |
| | 2006 | 0.10 | 0.05 | 60 | 64 | 936 | 1,463 | CESD | L (BL = 1992) | HRS | 85 | 1 |

**Table continues**

**Table 2.** Continued

| Author, Year (Reference No.) | Baseline Year | SMD | SE | Age Minimum | Age Maximum[a] | No. of Men | No. of Women | Instrument | Study Design[b] | Data Source | Retention[b] | Recall Period[b] |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Oksuzyan, 2010 (94) | 2006 | 0.11 | 0.05 | 65 | 69 | 1,537 | 1879 | CESD | L (BL = 1992) | HRS | 85 | 1 |
| | 2006 | 0.12 | 0.05 | 70 | 74 | 1,267 | 1,560 | CESD | L (BL = 1992) | HRS | 85 | 1 |
| | 2006 | 0.16 | 0.02 | 75 | 79 | 906 | 1,128 | CESD | L (BL = 1992) | HRS | 85 | 1 |
| | 2006 | 0.11 | 0.02 | 80 | 84 | 647 | 917 | CESD | L (BL = 1992) | HRS | 85 | 1 |
| | 2006 | 0.04 | 0.04 | 85 | 89 | 344 | 649 | CESD | L (BL = 1992) | HRS | 85 | 1 |
| | 2006 | 0.04 | 0.02 | 90 | 99 | 142 | 379 | CESD | L (BL = 1992) | HRS | 85 | 1 |
| Thibodeau, 2014 (95) | 2008 | 0.29 | 0.02 | 18 | 29 | 550 | 500 | PHQ-9 | C | NHANES 2008 | | 1 |
| | 2008 | 0.34 | 0.05 | 30 | 39 | 431 | 447 | PHQ-9 | C | NHANES 2008 | | 1 |
| | 2008 | 0.30 | 0.05 | 40 | 49 | 391 | 452 | PHQ-9 | C | NHANES 2008 | | 1 |
| | 2008 | 0.23 | 0.02 | 50 | 59 | 418 | 400 | PHQ-9 | C | NHANES 2008 | | 1 |
| | 2008 | 0.29 | 0.05 | 60 | 69 | 434 | 459 | PHQ-9 | C | NHANES 2008 | | 1 |
| | 2008 | 0.25 | 0.05 | 70 | 99 | 483 | 482 | PHQ-9 | C | NHANES 2008 | | 1 |
| Bushman, 2012 (96) | 2011 | 0.14 | 0.05 | 18 | 90 | 251 | 549 | CESD | C | Study-specific | | 1 |
| Gettler, 2016 (97) | 2011 | 0.14 | 0.04 | 20 | 60 | 1,505 | 933 | PHQ-9 | C | NHANES 2011–2012 | | 1 |
| Margraf, 2016 (98) | 2013 | −0.12 | 0.14 | 18 | 99 | 1,252 | 1,786 | DASS-D | BL | Bochum Optimism and Mental Health | | 1 |

Abbreviations: ACLS, American's Changing Lives Survey; Add Health, The National Longitudinal Study of Adolescent to Adult Health; CDI, Children's Depression Inventory; CESD, Center for Epidemiologic Studies–Depression scale; CFS, Commonwealth Fund Survey of Adolescent Girls and Boys; DASS-D, Depression Anxiety Stress Scales–Depression subscale; DFB, Depressive Feelings and Behaviors; HBSC, Health Behavior in School-Aged Children; HRS, Health and Retirement Study; MHI-D, Mental Health Inventory–Depression scale; NHANES, National Health and Nutrition Epidemiologic Survey; NLSY97, National Longitudinal Survey of Youth 1997; NSFH-2, National Survey of Families and Households 2; NSHAP, National Social Life, Health, and Aging Project; PHQ-9, 9-item Patient Health Questionnaire; SE, standard error; SMD, standardized mean difference.

[a] When the sample age range was described as all ages (e.g., 18 years and up), upper bound was coded as 99.

[b] Study design: C, cross-sectional, BL, baseline interview of a longitudinal study; L, other wave of longitudinal study (BL year). Retention: refers to follow-up rate in samples that are not cross-sectional or baseline longitudinal. Recall period: 1, past week to 6 months; 2, past year.

**Table 3.** Distributions of All Variables Used in Meta-Regression Models Analyzing Studies of Gender Differences in Depression, United States, 1982–2017

| Variable | Diagnostic Studies (*n* = 76) | | Symptom-Based Studies (*n* = 68) | |
|---|---|---|---|---|
| | No. | % | No. | % |
| Year[a,b] | 2010 (6.9) | | 2001 (7.8) | |
| Age group, year | | | | |
| All[c] | 3 | 3.9 | 6 | 8.8 |
| 10–19 | 27 | 35.5 | 33 | 48.5 |
| 20–39 | 24 | 31.6 | 4 | 5.9 |
| 40–59 | 11 | 14.5 | 9 | 13.2 |
| ≥60 | 11 | 14.5 | 16 | 23.5 |
| Symptom period | | | | |
| Prior year | 74 | 97.4 | | |
| Lifetime | 2 | 2.6 | | |
| Instrument | | | | |
| DSM-III/III-R | 4 | 5.3 | | |
| DSM-IV/IV-R | 71 | 93.4 | | |
| DSM-5 | 1 | 1.3 | | |
| CESD | | | 42 | 61.7 |
| CDI | | | 8 | 11.8 |
| PHQ-9 | | | 7 | 10.3 |
| Other | | | 11 | 16.2 |

Abbreviations: CDI, Children's Depression Inventory; CESD, Center for Epidemiologic Studies–Depression scale; DSM, *Diagnostic and Statistical Manual of Mental Disorders*; PHQ-9, 9-item Patient Health Questionnaire.

[a] Values are expressed as mean (standard deviation).

[b] Year range: 1990–2017 (diagnostic studies); 1982–2013 (symptom-based studies).

[c] Studies were not included to estimate age effects.

among the youngest ages over the study period. There was no evidence of time changes among any other age groups. These results are also presented graphically in Web Figure 3.

Main effects among symptom-based studies were estimated in model 1b (Table 4). In these studies, the depression gap with all variables at their reference levels was 0.3 (0.09, 0.51). There was no evidence of change over time overall. Compared to samples of persons aged ≥60 years, the depression gap was greater only among the youngest ages (10–19 years; standardized mean difference = 0.40, based on combined intercept (0.30) and age 10–19 (0.10) model coefficients). In model 2b, the interaction term for youngest age group was elevated (standardized mean difference = 0.03, 95% CI: 0.01, 0.05), suggesting that, compared with the oldest ages, the symptom-based depression gap increased over the study periods among the youngest ages. These results are also presented graphically in Web Figure 4. Compared with studies that measured depression with the Center for Epidemiologic Studies–Depression scale, the depression gap was higher in the 7 studies that used the Patient Health Questionnaire (standardized mean difference = 0.14, 95% CI: 0, 0.28) and other instruments (standardized mean difference = 0.13, 95% CI: 0.04, 0.22).

### Multiple imputation

In a sensitivity analysis, missing variance information was multiply imputed for 27 diagnostic studies. The depression gap with all variables at the reference level was slightly larger than in the unimputed model (PR = 2.49, 95% CI: 1.28, 4.88), and the age-by-time interaction tests were similar to the unimputed estimates (ages 10–19 years, PR = 1.20, 95%: 1.01, 1.39; no other age differences vs. the referent). The imputed random-effects model pooled PR was not appreciably different from the complete-case analysis (PR = 1.97, 95% CI: 1.82, 2.14) (imputed data not shown). Overall, results suggested that the complete-case analysis was not appreciably biased by missing data.

### Publication bias

Funnel plots are shown in Web Figures 5 (diagnostic studies) and 6 (symptom-based studies). In the symptom-based model, Egger's test indicated no evidence of publication bias (intercept = −1.19, 95% CI: –3.5, 1.1), although the trim-and-fill procedure imputed 23 additional studies to achieve funnel plot symmetry. Imputing these studies

**Table 4.** Meta-Regression Model Estimates for Analysis of Studies of Gender Differences in Depression, United States, 1982–2017

| Variable | Diagnostic Depression Gap | | | | Symptom-Based Depression Gap | | | |
|---|---|---|---|---|---|---|---|---|
| | Model 1a | | Model 2a[a] | | Model 1b | | Model 2b[a] | |
| | PR | 95% CI | PR | 95% CI | SMD | 95% CI | SMD | 95% CI |
| Intercept | 2.27 | 1.48, 3.05 | 2.30 | 1.58, 3.36 | 0.30 | 0.09, 0.51 | 0.54 | 0.34, 0.75 |
| Study year | 1.00 | 0.99, 1.01 | 1.00 | 0.98, 1.01 | −0.01 | −0.02, 0.01 | −0.01 | −0.02, 0.00 |
| Age group, years[b] | | | | | | | | |
| 10–19 | 1.26 | 1.02, 1.56 | 0.43 | 0.21, 0.88 | 0.10 | 0.03, 0.16 | 0.08 | −0.13, 0.29 |
| 20–39 | 0.89 | 0.73, 1.08 | 0.41 | 0.06, 0.77 | 0.03 | −0.10, 0.15 | −0.02 | −0.20, 0.16 |
| 40–59 | 0.94 | 0.59, 1.50 | 0.86 | 0.30, 1.43 | −0.02 | −0.14, 0.11 | −0.03 | −0.09, 0.03 |
| Ages 10–19 × study year | | | 1.05 | 1.01, 1.08 | | | 0.03 | 0.01, 0.05 |
| Ages 20–39 × study year | | | 1.02 | 0.99, 1.04 | | | 0.00 | −0.01, 0.01 |
| Ages 40–59 × study year | | | 1.00 | 0.98, 1.03 | | | 0.01 | −0.01, 0.02 |
| Diagnostic depression instrument[b] | | | | | | | | |
| DSM-III/IIIR | 0.94 | 0.59, 1.50 | 0.80 | 0.49, 1.32 | | | | |
| Other | 1.08 | 0.80, 1.48 | 1.03 | 0.74, 1.43 | | | | |
| Symptom-based depression instrument[b] | | | | | | | | |
| CDI | | | | | −0.06 | −0.22, 0.09 | −0.05 | −0.21, 0.11 |
| PHQ-9 | | | | | 0.13 | 0.03, 0.23 | 0.14 | 0.00, 0.28 |
| Other | | | | | 0.16 | 0.02, 0.30 | 0.13 | 0.04, 0.22 |

Abbreviations: CDI, Children's Depression Inventory; CESD, Center for Epidemiologic Studies–Depression scale; CI, confidence interval; DSM, *Diagnostic and Statistical Manual of Mental Disorders*; PHQ-9, Patient Health Questionnaire; PR, prevalence ratio; SMD, standardized mean difference.

[a] Adjusted for all model 1 variables.

[b] Age group referent: ≥60 years. Diagnostic referent: DSM-IV. Symptom-based referent: CESD.

increased the pooled effect size from 0.22 to standardized mean difference = 0.27 (prediction interval: −0.026, 0.57). In the diagnostic-based model, Egger's test indicated no evidence of publication bias (intercept = −0.266, 95% CI:−1.78, 1.24). The trim-and-fill procedure imputed no additional studies.

## DISCUSSION

The purpose of this systematic review and meta-regression was to review studies of the depression gap and characterize changes in the gap over time. To our knowledge, this is the largest study to examine changes in the depression gap over time by age in the United States. There were 4 central findings. First, women's depression risk was twice that of men overall, and the effect size was moderate among symptom-based studies. Second, there was no variation over time among adults ages 20 years or older, which does not support the hypothesis that changing gendered social positions are narrowing the depression gap. Third, the depression gap increased over time among respondents ages 10–19. Fourth, variation in the magnitude of the symptom-based depression gap was related to differences in depression instrument.

Concordant with nearly all of the depression gap literature, the present meta-analysis identified an appreciable depression gap between men and women. Findings were generally consistent between diagnostic depression and symptom-based depression measures, although more variation in the depression gap was found in studies of depression symptom scales. This variation was likely due in part to differences in the depression instrument across these studies. Symptom scales like the Center for Epidemiologic Studies–Depression scale, the most commonly used instrument in these studies, correlate with diagnostic depression but likely measure more general psychological distress and demoralization constructs (38). This variation should be considered when measuring and interpreting the depression gap using symptom scales in future individual studies.

While the meta-regression results suggested no change in the depression gap over time on average, there was heterogeneity in the time effects by age group, which is potentially indicative of cohort effects. Among adults ages 20 or older, there was no variation over time in the depression gap. Evidence of changes in the adult depression gap to date has been mixed. Some have reported a narrowing gender depression gap among younger adults over time, among individuals born from 1905–1965 (21, 39) and young adults ages 18–25 from 2005–2014 (16). In contrast, other studies have reported no effects or an increasing depression gap over time, among individuals born from 1936–1975 (40) and individuals born from 1915–1955 (41). To some extent, these differences reflected the period of recall, the age and birth year of respondents, and the depression instrument. The present meta-regression sought to account for these sources of heterogeneity across studies and estimate a summary of overall variation in a wider age range to the present day.

The time period covered by the present study coincides with broad changes to women's social positions in the United States. It was hypothesized that these changes would narrow the depression gap, but the results do not support a clear effect on the depression gap among adults. While it could be that the depression gap is not influenced by social position, the lack of an effect could also reflect both positive and negative consequences of changing position on the depression gap. On one hand, changing social position is indicated by greater opportunities in the workplace and access to personal socioeconomic (13, 42) and psychosocial resources among women (43, 44). Greater resources might reduce exposure to stress (45) and mitigate the effects of stressors (46, 47) in ways that influence the risk of depression (9, 48, 49). On the other hand, these changes might increase exposure to role conflict– and overload–related stressors that could increase women's depression risk (50, 51). Alternatively, regardless of the impact of changing social positions on women's exposure to stress, stress responses might still remain gendered (52, 53) in ways that are deleterious to women, thus sustaining the depression gap magnitude. Future research to elucidate these complex and potentially countervailing effects of gendered social and economic changes would add to the current understanding of depression gap trends.

Among the youngest respondents, however, the depression gap was appreciably larger than among respondents age ≥60. This pattern has been reported by individual (54) and meta-analytical (55) studies of age effects in the depression gap, which suggest that the depression gap peaks around ages 13–15. This peaking corresponds with the onset of puberty, which marks significant neurobiological changes (56) but also substantial changes in adolescents' social context, marked by increases in psychosocial stressors and interpersonal conflict among peers (57). The development of secondary sex characteristics and other physical changes, such as acne or increased adipose tissue, serve as additional sources of potential negative social interactions (58). These changes have been shown to increase the risk of depression and anxiety, especially in adolescent girls (59), whose experiences might be exacerbated by depressogenic coping strategies such as rumination (10).

In addition to identifying age effects overall, the interaction between age and study year indicated that the depression gap has increased among adolescents since 1982. These results align with previous studies showing that the adolescent depression gap has been increasing and emerging at earlier ages for several generations (16, 40). Causes of these trends are not clear, but again, changes in the adolescent social environment have been hypothesized. The prevalence of online harassment and bullying has increased over the past 20 years and is more frequently experienced by girls (60). While social media use entails a diverse set of exposures with potentially positive effects on adolescent self-esteem (61), problematic use is more common among girls (62). To date, social media use is inconsistently linked to depressed mood (63, 64), and more detailed research is needed before any particular mechanism is implicated. Fortunately, recent cohort studies have included more detailed measures of social media use (65). These studies can address the role social media use plays in depression across cohorts (66). Regardless of the causes of these emergent trends, clinicians should pay particular attention to adolescent girls as a high-risk group for depression. Additionally, it is unknown

whether these trends will be limited to adolescence or the gap will continue to widen through adulthood. To date, there is little evidence of gender differences in depression recurrence (67, 68); however, future studies that follow adolescent cohorts longitudinally into adulthood are necessary to fully answer this important question.

The findings of this meta-regression should be interpreted in light of several limitations. First, the majority of national samples were cross-sectional in design and were only able to assess prevalent depression status. Second, there was significant heterogeneity in age ranges across studies. A consequence of making the age group categories comparable across studies (i.e., observed ages) involved truncating the age ranges within each sample (i.e., true ages), potentially introducing measurement error because observed age range was sometimes different from the true age range. This measurement error was likely nondifferential as it was not related to the depression gap outcome, so any bias would attenuate age and age-by-time interaction estimates. In this study, the reported age effects were robust to an alternative set of age groupings (i.e., 10–17, 18–25, 26–35, 36–45, 46–55, ≥56), suggesting that the age trends were not artifactual. Finally, among included studies of diagnostic depression, there were missing data and evidence of potential publication bias, which might have distorted the summary estimates of the depression gap. However, evidence from multiple imputation models, and trim-and-fill sensitivity analyses, suggested that this bias was minimal. Finally, this study complements recent international research showing associations between country-level gender equality and the depression gap (55, 69); however, it was not structured to directly test the putative social mechanisms that might explain the depression gap itself. While mental health is the result of multiple interacting exposures, identifying changes over time highlights the important role of the social environment, which is more dependent on historical and social context than relatively immutable biological determinants. Future research to directly test the role of changing gender social positions over time in mediating any changes in the depression gap would be an important contribution to understanding determinants of the gender depression gap.

In conclusion, with a sample of 813,189 respondents spanning 8 decades of age and a time period of 35 years, the present study finds evidence of a persistent gender gap in depression that appears to be increasing in adolescents over time. Future research is needed to understand the causes of these changes in the gender gap in adolescents, in order to inform depression prevention and treatment efforts, and reverse potentially growing depression disparities among young people.

## ACKNOWLEDGMENTS

Author affiliations: Department of Epidemiology, Mailman School of Public Health, Columbia University, New York, New York, United States (Jonathan M. Platt, Lisa Bates, Katherine M. Keyes); T. Denny School of Social and Family Dynamics, Arizona State University, Tempe, Arizona , United States (Justin Jager); and Department of Psychology, Harvard University, Cambridge, Massachusetts, United States (Katie A. McLaughlin).

This work was funded by the National Institute of Mental Health (grant T32-MH1304343 to J.M.P. and grants R01-MH103291, R01-MH106482, and R56-119194 to K.A.M.) and National Institute on Alcohol Abuse and Alcoholism (grant R01-AA026861 to K.M.K.).

The authors thank Dr. Sharon Schwartz for her constructive feedback in earlier drafts of this manuscript.

Conflict of interest: none declared.

## REFERENCES

1. Center for Behavioral Health Statistics and Quality, Substance Abuse and Mental Health Services Administration. *Key Substance Use and Mental Health Indicators in the United States: Results From the 2016 National Survey on Drug Use and Health*. Rockville, MD: Substance Abuse and Mental Health Services Administration; 2017. https://www.samhsa.gov/data/report/key-substance-use-and-mental-health-indicators-united-states-results-2016-national-survey. Accessed November 13, 2018
2. Karg RS, Bose J, Batts KR, et al. Past year mental disorders among adults in the United States: Results from the 2008–2012 Mental Health Surveillance Study. In: *CBHSQ Data Review*. Rockville, MD: Substance Abuse and Mental Health Services Administration; 2014. https://www.samhsa.gov/data/sites/default/files/NSDUH-DR-N2MentalDis-2014-1/Web/NSDUH-DR-N2MentalDis-2014.htm. Accessed December 21, 2018.
3. Young MA, Fogg LF, Scheftner WA, et al. Sex differences in the lifetime prevalence of depression: does varying the diagnostic criteria reduce the female/male ratio? *J Affect Disord*. 1990;18(3):187–192.
4. Gove WR, Geerken MR. Response bias in surveys of mental health: an empirical investigation. *Am J Sociol*. 1977;82(6):1289–1317.
5. Kessler RC. Epidemiology of women and depression. *J Affect Disord*. 2003;74(1):5–13.
6. Angold A, Worthman CW. Puberty onset of gender differences in rates of depression: a developmental, epidemiologic and neuroendocrine perspective. *J Affect Disord*. 1993;29(2–3):145–158.
7. Horwitz AV. The sociological study of mental illness: A critique and synthesis of four perspectives. In: *Handbook of the Sociology of Mental Health*. New York, NY: Springer; 2013:95–112.
8. Aneshensel CS. Social stress: theory and research. *Annu Rev Sociol*. 1992;18(1):15–38.
9. Aneshensel CS, Phelan JC, Bierman A. The sociology of mental health: surveying the field. In: *Handbook of the Sociology of Mental Health*. New York, NY: Springer; 2013:1–19.
10. Johnson DP, Whisman MA. Gender differences in rumination: a meta-analysis. *Personal Individ Differ*. 2013;55(4):367–374.
11. Risman BJ, Davis G. From sex roles to gender structure. *Curr Sociol*. 2013;61(5–6):733–755.

12. England P, Li S. Desegregation stalled: the changing gender composition of college majors, 1971-2002. *Gend Soc*. 2006; 20(5):657–677.
13. England P. *Households, Employment, and Gender: A Social, Economic, and Demographic View*. Abingdon, UK: Routledge; 2017.
14. Lee M, Mather M. *US Labor Force Trends*. 2008. https://www.prb.org/wp-content/uploads/2008/07/63.2uslabor.pdf. Accessed March 7, 2019.
15. Keyes KM, Gary D, O'Malley PM, et al. Recent increases in depressive symptoms among US adolescents: trends from 1991 to 2018. *Soc Psychiatry Psychiatr Epidemiol*. 2019; 54(8):987–996.
16. Mojtabai R, Olfson M, Han B. National trends in the prevalence and treatment of depression in adolescents and young adults. *Pediatrics*. 2016;138(6): e20161878–e20161878.
17. Kessler RC, McRae JA. Trends in the relationship between sex and psychological distress: 1957–1976. *Am Sociol Rev*. 1981;46(4):443–452.
18. Twenge JM, Nolen-Hoeksema S. Age, gender, race, socioeconomic status, and birth cohort difference on the children's depression inventory: a meta-analysis. *J Abnorm Psychol*. 2002;111(4):578–588.
19. Jane Costello E, Erkanli A, Angold A. Is there an epidemic of child or adolescent depression? *J Child Psychol Psychiatry*. 2006;47(12):1263–1271.
20. Lewinsohn PM, Solomon A, Seeley JR, et al. Clinical implications of "subthreshold" depressive symptoms. *J Abnorm Psychol*. 2000;109(2):345–351.
21. Robins, Regier LN, DA. *Psychiatric Disorders in America: The Epidemiologic Catchment Area Study*. New York, NY: Free Press; 1991.
22. Weissman MM, Livingston B, Leaf PJ, et al. Affective disorders. In: *Psychiatric Disorders in America*. New York, NY: Free Press; 1991:53–80.
23. Tousignant M, Brosseau R, Tremblay L. Sex biases in mental health scales: do women tend to report less serious symptoms and confide more than men? *Psychol Med*. 1987;17(01): 203–215.
24. Fleiss JL, Levin B, Paik MC. *Statistical Methods for Rates and Proportions*. Hoboken, NJ: John Wiley & Sons; 2013.
25. Cohen J. *Statistical Power Analysis for the Social Sciences*. New York, NY: Erlbaum Associates; 1988.
26. Lipsey MW, Wilson DB. *Practical Meta-Analysis*. Newbury Park, CA: SAGE Publications, Inc; 2001.
27. Sutton AJ. Publication bias. In: *The Handbook of Research Synthesis and Meta-Analysis*. New York, NY, US: Russell Sage Foundation; 2009:435–452.
28. Egger M, Smith GD, Schneider M, et al. Bias in meta-analysis detected by a simple, graphical test. *BMJ*. 1997;315(7109):629–634.
29. Duval S, Tweedie R. Trim and fill: a simple funnel-plot–based method of testing and adjusting for publication bias in meta-analysis. *Biometrics*. 2000;56(2):455–463.
30. IntHout J, Ioannidis JP, Rovers MM, et al. Plea for routinely presenting prediction intervals in meta-analysis. *BMJ Open*. 2016;6(7):e010247.
31. Schwarzer G. Meta: an R package for meta-analysis. *R news*. 2007;7(3):40–45.
32. Viechtbauer W. Conducting meta-analyses in R with the metafor package. *J Stat Softw*. 2010;36(3):1–48.
33. Team RC. R: A language and environment for statistical computing. 2013. http://www.R-project.org/. Accessed March 25, 2025.
34. Pigott TD. Handling missing data. In: *The Handbook of Research Synthesis and Meta-Analysis*. New York, NY: Russell Sage Foundation; 2009:399–416.
35. Ellington EH, Bastille-Rousseau G, Austin C, et al. Using multiple imputation to estimate missing data in meta-regression. *Methods Ecol Evol*. 2015;6(2):153–163.
36. Alaimo K, Olson CM, Frongillo EA. Family food insufficiency, but not low family income, is positively associated with dysthymia and suicide symptoms in adolescents. *J Nutr*. 2002;132(4):719–725.
37. Hasin DS, Goodwin RD, Stinson FS, et al. Epidemiology of major depressive disorder: results from the National Epidemiologic Survey on Alcoholism and Related Conditions. *Arch Gen Psychiatry*. 2005;62(10):1097–1106.
38. Dohrenwend BP, Dohrenwend BS. Perspectives on the past and future of psychiatric epidemiology. The 1981 Rema Lapouse Lecture. *Am J Public Health*. 1982;72(11): 1271–1279.
39. Wickramaratne PJ, Weissman MM, Leaf PJ, et al. Age, period and cohort effects on the risk of major depression: results from five United States communities. *J Clin Epidemiol*. 1989;42(4):333–343.
40. Kessler RC, McGonagle KA, Nelson CB, et al. Sex and depression in the National Comorbidity Survey. II: cohort effects. *J Affect Disord*. 1994;30(1):15–26.
41. Yang Y, Lee LC. Sex and race disparities in health: cohort variations in life course patterns. *Soc Forces*. 2009;87(4): 2093–2124.
42. Lorant V, Deliège D, Eaton W, et al. Socioeconomic inequalities in depression: a meta-analysis. *Am J Epidemiol*. 2003;157(2):98–112.
43. Rafaeli-Mor E, Steinberg J. Self-complexity and well-being: a review and research synthesis. *Pers Soc Psychol Rev*. 2002; 6(1):31–58.
44. Gove WR. The relationship between sex roles, marital status, and mental illness. *Soc Forces*. 1972;51(1):34–44.
45. Persson R, Hansen Å-M, Ohlsson K, et al. Physiological and psychological reactions to work in men and women with identical job tasks. *Eur J Appl Physiol*. 2009;105(4):595–606.
46. Kunz-Ebrecht SR, Kirschbaum C, Steptoe A. Work stress, socioeconomic status and neuroendocrine activation over the working day. *Soc Sci Med*. 2004;58(8):1523–1530.
47. Brydon L, Edwards S, Mohamed-Ali V, et al. Socioeconomic status and stress-induced increases in interleukin-6. *Brain Behav Immun*. 2004;18(3):281–290.
48. Aneshensel CS. Marital and employment role-strain, social support, and depression among adult women. In Hobfoll SE, ed. *The Series in Clinical and Community Psychology. Stress, Social Support, and Women*. Washington, DC: Hemisphere Publishing Corp;. 1986;99–114.
49. Gove WR, Tudor JF. Adult sex roles and mental illness. *Am J Sociol*. 1973;78(4):812–835.
50. Hecht LM. Role conflict and role overload: different concepts, different consequences. *Sociol Inq*. 2001;71(1): 111–121.
51. Pearlin LI. Sex roles and depression. In: *Life-Span Developmental Psychology*. New York, NY: Elsevier; 1975: 191–207.
52. Horwitz AV. *Creating Mental Illness*. Chicago, IL: University of Chicago Press; 2002.
53. Rogers A, Pilgrim D. *A Sociology of Mental Health and Illness*. London, UK: McGraw-Hill Education; 2014.
54. Kessler RC, McGonagle KA, Swartz M, et al. Sex and depression in the National Comorbidity Survey I: lifetime prevalence, chronicity and recurrence. *J Affect Disord*. 1993; 29(2–3):85–96.

55. Salk RH, Hyde JS, Abramson LY. Gender differences in depression in representative national samples: meta-analyses of diagnoses and symptoms. *Psychol Bull*. 2017;143(8):783–822.
56. Blakemore S-J, Burnett S, Dahl RE. The role of puberty in the developing adolescent brain. *Hum Brain Mapp*. 2010;31(6):926–933.
57. Hankin BL, Mermelstein R, Roesch L. Sex differences in adolescent depression: stress exposure and reactivity models. *Child Dev*. 2007;78(1):279–295.
58. Vicary JR, Klingaman LR, Harkness WL. Risk factors associated with date rape and sexual assault of adolescent girls. *J Adolesc*. 1995;18(3):289.
59. Copeland W, Shanahan L, Miller S, et al. Outcomes of early pubertal timing in young women: a prospective population-based study. *Am J Psychiatry*. 2010;167(10):1218–1225.
60. Kessel Schneider S, O'Donnell L, Smith E. Trends in cyberbullying and school bullying victimization in a regional census of high school students, 2006–2012. *J Sch Health*. 2015;85(9):611–620.
61. Steinfield C, Ellison NB, Lampe C. Social capital, self-esteem, and use of online social network sites: a longitudinal analysis. *J Appl Dev Psychol*. 2008;29(6):434–445.
62. Lenhart A, Smith A, Anderson M, et al. *Teens, Technology and Friendships*. Washington DC: Pew Research Center; 2015.
63. Augner C, Hacker GW. Associations between problematic mobile phone use and psychological parameters in young adults. *Int J Public Health*. 2012;57(2):437–441.
64. Orben A, Przybylski AK. The association between adolescent well-being and digital technology use. *Nat Hum Behav*. 2019;3(2):173–182.
65. Bagot K, Matthews SA, Mason M, et al. Current, future and potential use of mobile and wearable technologies and social media data in the ABCD study to increase understanding of contributors to child health. *Dev Cogn Neurosci*. 2018;32:121–129.
66. Paulus MP, Squeglia LM, Bagot K, et al. Screen media activity and brain structure in youth: evidence for diverse structural correlation networks from the ABCD study. *Neuroimage*. 2019;185:140–153.
67. Kuehner C. Gender differences in unipolar depression: an update of epidemiological findings and possible explanations. *Acta Psychiatr Scand*. 2003;108(3):163–174.
68. Burcusa SL, Iacono WG. Risk for recurrence in depression. *Clin Psychol Rev*. 2007;27(8):959–985.
69. Yu S. Uncovering the hidden impacts of inequality on mental health: a global study. *Transl Psychiatry*. 2018;8(1):1–10.
70. Grant BF. Comorbidity between DSM-IV drug use disorders and major depression: results of a national survey of adults. *J Subst Abuse*. 1995;7(4):482–497.
71. Dawson DA, Grant BF. Family history of alcoholism and gender: their combined effects on DSM-IV alcohol dependence and major depression. *J Stud Alcohol*. 1997;59(1):97–106.
72. Kessler RC, Birnbaum HG, Shahly V, et al. Age differences in the prevalence and co-morbidity of DSM-IV major depressive episodes: results from the WHO World Mental Health Survey Initiative. *Depress Anxiety*. 2010;27(4):351–364.
73. Center for Behavioral Health Statistics and Quality. *National Survey on Drug Use and Health: Mental Health Detailed Tables*. Rockville, MD: Substance Abuse and Mental Health Services Administration; 2004. https://www.samhsa.gov/data/report/2004-nsduh-detailed-tables. Accessed January 7, 2019.
74. Center for Behavioral Health Statistics and Quality. *National Survey on Drug Use and Health: Mental Health Detailed Tables*. Rockville, MD: Substance Abuse and Mental Health Services Administration; 2005. https://www.samhsa.gov/data/report/2005-nsduh-detailed-tables. Accessed January 7, 2019.
75. Center for Behavioral Health Statistics and Quality. *National Survey on Drug Use and Health: Mental Health Detailed Tables*. Rockville, MD: Substance Abuse and Mental Health Services Administration; 2006. https://www.samhsa.gov/data/report/2006-nsduh-detailed-tables. Accessed January 7, 2019.
76. Center for Behavioral Health Statistics and Quality. *National Survey on Drug Use and Health: Mental Health Detailed Tables*. Rockville, MD: Substance Abuse and Mental Health Services Administration; 2007. https://www.samhsa.gov/data/report/2007-nsduh-detailed-tables. Accessed January 7, 2019.
77. Center for Behavioral Health Statistics and Quality. *National Survey on Drug Use and Health: Mental Health Detailed Tables*. Rockville, MD: Substance Abuse and Mental Health Services Administration; 2008. https://www.samhsa.gov/data/report/2008-nsduh-detailed-tables. Accessed January 7, 2019.
78. Center for Behavioral Health Statistics and Quality. *National Survey on Drug Use and Health: Mental Health Detailed Tables*. Rockville, MD: Substance Abuse and Mental Health Services Administration; 2009. https://www.samhsa.gov/data/report/2009-nsduh-detailed-tables. Accessed January 7, 2019.
79. Center for Behavioral Health Statistics and Quality. *National Survey on Drug Use and Health: Mental Health Detailed Tables*. Rockville, MD: Substance Abuse and Mental Health Services Administration; 2010. https://www.samhsa.gov/data/report/results-2010-national-survey-drug-use-and-health-detailed-tables. Accessed January 7, 2019.
80. Center for Behavioral Health Statistics and Quality. *National Survey on Drug Use and Health: Mental Health Detailed Tables*. Rockville, MD: Substance Abuse and Mental Health Services Administration; 2011. https://www.samhsa.gov/data/report/results-2011-national-survey-drug-use-and-health-detailed-tables. Accessed January 7, 2019.
81. Verplaetse TL, Smith PH, Pittman BP, et al. Associations of gender, smoking, and stress with transitions in major depression diagnoses. *Yale J Biol Med*. 2016;89(2):123–129.
82. Center for Behavioral Health Statistics and Quality. *National Survey on Drug Use and Health: Mental Health Detailed Tables*. Rockville, MD: Substance Abuse and Mental Health Services Administration; 2013. https://www.samhsa.gov/data/sites/default/files/NSDUH-DetTabs2013.pdf. Accessed January 7, 2019.
83. Center for Behavioral Health Statistics and Quality. *National Survey on Drug Use and Health: Mental Health Detailed Tables*. Rockville, MD: Substance Abuse and Mental Health Services Administration; 2017. https://www.samhsa.gov/data/report/2017-nsduh-detailed-tables. Accessed November 13, 2018.
84. Ferketich AK, Schwartzbaum JA, Frid DJ, et al. Depression as an antecedent to heart disease among women and men in the NHANES I study. *Arch Intern Med*. 2000;160(9):1261–1268.
85. Everson-Rose SA, House JS, Mero RP. Depressive symptoms and mortality risk in a national sample: confounding effects of health status. *Psychosom Med*. 2004;66(6):823–830.

86. Inaba A, Thoits PA, Ueno K, et al. Depression in the United States and Japan: gender, marital status, and SES patterns. *Soc Sci Med*. 2005;61(11):2280–2292.
87. Marmorstein NR. Longitudinal associations between alcohol problems and depressive symptoms: early adolescence through early adulthood. *Alcohol Clin Exp Res*. 2009;33(1): 49–59.
88. Neumark-Sztainer D, Hannan PJ. Weight-related behaviors among adolescent girls and boys: results from a national survey. *Arch Pediatr Adolesc Med*. 2000;154(6):569–577.
89. Mumford EA, Liu W, Hair EC, et al. Concurrent trajectories of BMI and mental health patterns in emerging adulthood. *Soc Sci Med*. 2013;98:1–7.
90. Song L. Social capital and psychological distress. *J Health Soc Behav*. 2011;52(4):478–492.
91. Shiovitz-Ezra S, Leitsch S, Graber J, et al. Quality of life and psychological health indicators in the national social life, health, and aging project. *J Gerontol B Psychol Sci Soc Sci*. 2009;64(suppl 1):i30–i37.
92. Haroz EE, Ybarra ML, Eaton WW. Psychometric evaluation of a self-report scale to measure adolescent depression: the CESDR-10 in two national adolescent samples in the United States. *J Affect Disord*. 2014;158:154–160.
93. Wang Z, Goswami D, Fitzgibbon H, et al. Gender-specific gene expression of a novel zinc finger transcription factor, KLF7, in the prefrontal cortex of subjects with major depression. *Biol Psychiatry*. 2010;67(9):135S.
94. Oksuzyan A, Crimmins E, Saito Y, et al. Cross-national comparison of sex differences in health and mortality in Denmark, Japan and the US. *Eur J Epidemiol*. 2010;25(7): 471–480.
95. Thibodeau MA, Asmundson GJG. The PHQ-9 assesses depression similarly in men and women from the general population. *Personal Individ Differ*. 2014;56: 149–153.
96. Bushman BJ, Moeller SJ, Konrath S, et al. Investigating the link between liking versus wanting self-esteem and depression in a nationally representative sample of American adults. *J Pers*. 2012;80(5):1453–1469.
97. Gettler LT, Oka RC. Are testosterone levels and depression risk linked based on partnering and parenting? Evidence from a large population-representative study of US men and women. *Soc Sci Med*. 2016;163:157–167.
98. Margraf J, Lavallee K, Zhang X, et al. Social rhythm and mental health: a cross-cultural comparison. *PLoS One*. 2016; 11(3):e0150312.