Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Equal treatment but fewer labels: a paired comparison of anxiety and depression diagnosis and treatment in US adults

Yueqi Cui

yueqicui@foxmail.com

Independent Researcher https://orcid.org/0009-0007-2513-4063

Research Article

Keywords: anxiety, depression, diagnosis, labeling, treatment gap, comorbidity, within-person

Posted Date: November 4th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11600530/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: The authors declare no competing interests.

Abstract

Background.

Anxiety disorders are reported to be less likely to be diagnosed than depression, but most evidence compares rates across different people and different instruments. Whether the gap reflects differential labeling of the same patient is unknown.

Methods.

We used the US National Health Interview Survey (NHIS, 2019–2025) to identify 17,901 adults reporting weekly-or-more symptoms of both anxiety and depression, comparing within-person diagnosis and treatment rates. Analyses used survey-weighted marginal contrasts, McNemar's test, conditional logistic regression, and TOST equivalence testing. Sensitivity analyses included E-value computation. Severity-matched comparisons used GAD-7 and PHQ-8 waves (2019, 2022, 2025; n = 67,233 concordant).

Results.

Within persons with both symptom types, anxiety was diagnosed 62.2% vs 73.8% for depression (− 11.7 pp, 95% CI − 12.5 to − 10.9; McNemar χ^2=1434.9, $p < 0.001$); the conditional logit OR was 0.225 (0.207–0.245). Treatment rates were formally equivalent (difference 0.25 pp, 90% CI − 0.12 to 0.62; TOST $p < 0.001$ at ± 1 pp). The gap was severity-dependent (+ 0.3, − 1.3, − 7.3, − 6.1 pp across bands) and present in all three years. It was not explained by comorbidity, insurance, access, treatment measurement, or survey year; the E-value was 3.6 (typical confounders RR 1.2–2.0).

Conclusions.

Among US adults with equivalent symptom burdens, anxiety is labeled less often than depression while being treated at the same rate. This labeling asymmetry is severity-dependent, stable across survey years, and contributes to the population-level identification gap for anxiety.

Highlights

- Within the same adult, anxiety is diagnosed 11.7 pp less often than depression.

- Treatment rates are equivalent, showing a labeling rather than treatment gap.

- The diagnosis gap is severity-dependent, largest at moderate symptom levels.

- Anxiety medication users are 3 times as likely to lack a matching diagnosis.

- The gap narrows from − 14.1 pp (2019) to − 9.8 pp (2025) and is robust.

1. Introduction

Anxiety disorders form the most common group of mental disorders and are underdiagnosed in all care settings (Penninx et al., 2021); in the US, 8.1% of adults screened positive for high anxiety symptoms in the 2019 NHIS (Daly and Robinson, 2021) — yet in the World Mental Health Surveys of 21 countries, most adults with a 12-month anxiety disorder received no treatment (Alonso et al., 2018; Evans-Lacko et al., 2018; Wang et al., 2007). Large treatment gaps for mood disorders are documented in population surveys worldwide, including recent nationwide data from South Asia (Sarkar et al., 2026). Increasing treatment use in the US has not, however, been accompanied by comparable gains in population mental health (Mojtabai and Jorm, 2015). Several studies further suggest that anxiety is less often labeled than depression (Lecrubier, 2007; Cornelius et al., 2014). Most of this evidence compares the fraction of people with each condition who receive a diagnosis, so the comparison rests on two non-overlapping groups defined by different symptom measures. Because symptom frequency, severity, comorbidity, and help-seeking all differ between people with anxiety and people with depression, it is unclear how much of the apparent “recognition gap” reflects differential labeling of the same patient versus compositional differences between the two populations.

We exploit a design feature of the National Health Interview Survey (NHIS): the same symptom frequency items are asked for anxiety and depression, and separate lifetime diagnosis items are asked for each. This permits a **within-person** comparison — for the same adult, with the same symptom burden, is an anxiety diagnosis more or less likely than a depression diagnosis? And conditional on treatment, is anxiety treated as often as depression?

We combine this paired design with a severity-matched validation using the GAD-7 and PHQ-8, which were administered in three NHIS waves (2019, 2022, and 2025). These scales have been used in NHIS to describe symptom burden and to compare measurement properties (Zablotsky et al., 2022) but not, to our knowledge, to match anxiety and depression severity within the same individuals. We report results that are robust to comorbidity, insurance, usual source of care, treatment definition, and multiplicity correction.

2. Methods

2.1 Data source

We used the NHIS public-use files for 2019–2025, restricted to Sample Adults aged ≥ 18 years with valid final weights. (National Center for Health Statistics, 2020–2026) We identified adults reporting weekly-or-more symptoms of both anxiety and depression (ANXFREQ_A and DEPFREQ_A ≤ 2, where 1 = daily, 2 = weekly, 3 = monthly, 4 = a few times a year, 5 = never). Adults with missing symptom responses were excluded. The analytic sample comprised 17,901 adults with both symptom types (8.6% of 207,064

adults). For population-level proportions we report the complete-case estimate; treating missing symptom responses as never yields a slightly lower figure, and the alternative is noted where used. For severity matching we used the 2019, 2022, and 2025 waves, in which the GAD-7 and PHQ-8 were administered (83,863 records, of which 67,233 had concordant GAD-7 and PHQ-8 severity bands and both diagnosis items observed).

2.2 Measures

Diagnosis. Lifetime anxiety diagnosis (ANXEV_A: 1 = yes, 2 = no) and lifetime depression diagnosis (DEPEV_A) were coded 1/0, with 7/8/9 coded missing.

Labeling. We use "labeling" to denote the presence of a recorded diagnosis (NHIS: "ever told by a doctor"). This measure conflates clinician diagnosis and patient recall of that diagnosis. The medication-user analysis (§ 3.4) bounds the recall component: among adults certainly in contact with the health system, the missing-label rate is 3.0 times higher for anxiety. These data cannot distinguish among (a) clinician diagnosis, (b) patient recall, and (c) documentation in records; we therefore interpret the observed asymmetry strictly as a deficit in recorded labeling. This limitation is not unique to NHIS: self-reported and claims-based diagnoses have been shown to identify distinct risk profiles in population studies (Wu et al., 2027).

Treatment. Anxiety medication (ANXMED_A), depression medication (DEPMED_A), and psychotherapy in the past 12 months (MHTHRPY_A) were each coded 1/0. Any treatment = medication or psychotherapy. Medication items have no recall window; the psychotherapy item refers to the past 12 months.

Severity. In the 2019, 2022, and 2025 waves, GAD-7 (0–21) and PHQ-8 (0–24) totals were computed from their items, with standard cut-points applied (none 0–4, mild 5–9, moderate 10–14, severe ≥ 15). Our computed scores agreed with the official GADCAT_A and PHQCAT_A in 100.0% of 83,863 records. (Spitzer et al., 2006; Kroenke et al., 2009; Zablotsky et al., 2022)

2.3 Statistical analysis

We report six analyses. (i) Diagnosis and treatment rates by condition. (ii) The within-person contrast in diagnosis, estimated as a survey-weighted marginal difference with the covariance between the two estimates accounted for in the standard error. (iii) McNemar's test on the discordant pairs (continuity-corrected). (iv) Conditional logistic regression stratified by person, which conditions out all time-invariant person-level confounders; time-varying confounders, such as changes in health status, could not be addressed (Hosmer and Lemeshow, 2000). (v) Two one-sided tests (TOST) of the treatment contrast at margins of ± 1 and ± 2 pp, set to reflect the study's measurement precision and between-wave variation rather than the observed difference (Schuirmann, 1987; Lakens, 2017). (vi) A four-way decomposition by diagnosis and treatment.

The TOST margins were chosen on empirical grounds: ±1 pp is smaller than the study precision (± 0.8 pp on the difference), so equivalence at this margin demonstrates that any true difference is smaller than

we can resolve; ±2 pp is smaller than the observed between-wave variation in treatment rates (range 0.2–0.6 pp). Both margins are well below the diagnosis gap (− 11.7 pp), confirming that the two contrasts are qualitatively different. All estimates use the final annual weight with pseudo-strata and pseudo-PSU re-coded within survey year; design-based inference uses Taylor linearization.

Sensitivity analyses covered the daily symptom threshold, a medication-only treatment definition, exclusion of 2020, pooling of severity bands, and sex- and ethnicity-stratified analyses with Benjamini-Hochberg correction. Analyses used R 4.6.1 with the $survey$ package. To quantify the robustness of the diagnosis gap to unmeasured confounding, we computed the E-value using the odds-ratio approximation (VanderWeele and Ding, 2017)—the minimum strength of association that an unmeasured confounder would need with both symptom type and diagnosis to explain away the observed association. Validation studies report that self-reported psychiatric diagnosis has low-to-moderate sensitivity (0.37–0.74) and specificity of 0.72–0.96 (Sanchez-Villegas et al., 2008; Leikauf and Federman, 2009). Because these estimates derive from populations with lower disorder prevalence than our sample, a quantitative correction is not applicable and the magnitude of any bias cannot be quantified from these data.

Missing data. The analytic restriction (both symptom types reported weekly or more) excludes adults with missing symptom responses. Within the analytic sample (n = 17,901), 42 adults (0.2%) had a missing anxiety diagnosis, 35 (0.2%) a missing depression diagnosis, and 63 (0.4%) either; these records are excluded from the four-way diagnostic decomposition, which is therefore based on 17,838 adults with both diagnoses observed. Item-level missingness within the analytic sample was under 0.3% for every other variable used. Population-level estimates are complete-case within each quantity unless otherwise noted. No imputation was performed.

Sample size and precision. The study uses complete NHIS files rather than a sample drawn for a specific hypothesis, so no power calculation was performed. The analytic sample of 17,901 adults yields a 95% CI of approximately ± 0.9 percentage points on each condition's diagnosis rate and ± 0.8 percentage points on the difference, which is the precision relevant to the primary comparison.

Rounding and confidence intervals. All percentages and percentage-point differences are reported to one decimal place from full precision; because each rate is rounded independently, a difference computed from displayed rates may differ by 0.1 pp from the displayed difference. Confidence intervals use the design degrees of freedom (t distribution). Odds-ratio intervals are computed on the log scale and back-transformed, so they are asymmetric on the odds-ratio scale.

3. Results

3.1 Sample

The analytic sample comprised 17,901 adults; 17,838 had both diagnosis items observed. Weighted characteristics are shown in Table 1.

Table 1
Characteristics of the analytic sample, n = 17,901.

Characteristic	Percent	n
Age 18–29	26.5	3,053
Age 30–44	25.3	4,301
Age 45–59	23.3	4,382
Age 60+	24.9	6,165
Female	59.1	10,898
Uninsured	10.4	1,505
No usual source of care	2.4	415
Anxiety symptoms daily	68.2	12,103
Depression symptoms daily	40.6	7,582
Anxiety diagnosis	62.1	10,944
Depression diagnosis	73.8	13,230
Anxiety treatment	61.5	11,126
Depression treatment	61.2	11,130

Weighted percentages of adults with weekly-or-more symptoms of both anxiety and depression, NHIS 2019–2025. The two diagnosis rows use all 17,901 respondents; the within-person contrasts in Table 2 use the 17,838 with both diagnosis items observed, so the anxiety diagnosis rate differs by 0.1 pp between the two tables.

3.2 Diagnosis differs but treatment does not

Within persons with both symptom types, the lifetime anxiety diagnosis rate was 62.2% (95% CI 61.3 to 63.0) versus 73.8% (73.0 to 74.7) for depression, a difference of **− 11.7 pp** (− 12.5 to − 10.9, $p < 10^{-15}$). In contrast, any-treatment rates were 61.5% (60.6 to 62.4) for anxiety and 61.2% (60.3 to 62.1) for depression, a difference of **0.25 pp** (− 0.2 to 0.7, p = 0.275).

Formal within-person tests. Of the 17,838 adults, 660 had an anxiety diagnosis without a depression diagnosis and 2,931 the reverse; 10,277 had both and 3,970 neither. McNemar's test gave χ^2=1434.9 (df = 1, p < 0.001). Conditional logistic regression among the 3,591 discordant respondents gave **OR = 4.44 (4.08–4.83)** for depression relative to anxiety, equivalent to **OR = 0.225 (0.207–0.245)** for anxiety relative to depression. The same specification applied to treatment gave OR = 1.02 (0.91–1.14, p = 0.797).

Equivalence testing. A null treatment difference could reflect either true equivalence or insufficient power. We conducted two one-sided tests (TOST) at margins of ± 1 and ± 2 pp (see Methods). The any-treatment difference was 0.25 pp (90% CI − 0.12 to 0.62), inside the margin (TOST $p < 0.001$ at both margins), and medication alone gave the same conclusion. Treatment rates are therefore formally equivalent, the precondition for interpreting the diagnosis gap as a labeling difference. The same test applied to diagnosis is nowhere near equivalence (TOST $p = 1.00$), confirming that the two contrasts genuinely differ.

Table 2
Core within-person contrasts and four-way decomposition.

Contrast	Anxiety	Depression	Estimate	Detail	p
Core within-person contrasts					
Diagnosis rate	62.2%	73.8%	−11.7 pp	95% CI − 12.5 to − 10.9	< 0.001
Treatment rate	61.5%	61.2%	0.25 pp	95% CI − 0.20 to 0.69	0.275
Diagnosis difference	—	—	−11.7 pp	95% CI − 12.5 to − 10.9	< 0.001
Treatment difference	—	—	0.25 pp	90% CI − 0.12 to 0.62	0.275
Formal within-person tests					
McNemar chi-square (df = 1)	—	—	1434.9	p < 0.001	< 0.001
Discordant pairs (n)	—	—	3,591	anxiety-only 660; depression-only 2,931	—
clogit OR, depression vs anxiety	—	—	4.44	95% CI 4.08–4.83	< 0.001
clogit OR, anxiety vs depression	—	—	0.225	95% CI 0.207–0.245	< 0.001
clogit OR, treatment	—	—	1.02	95% CI 0.91–1.14	0.797
TOST equivalence (± 1 pp)	—	—	equivalent	90% CI − 0.12–0.62	< 0.001
TOST equivalence (± 2 pp)	—	—	equivalent	90% CI − 0.12–0.62	< 0.001
Four-way decomposition					
Diagnosed and treated	49.2%	55.7%	—	—	—
Diagnosed, untreated	13.0%	18.2%	—	—	—
Undiagnosed, treated	12.3%	5.6%	—	—	—
Undiagnosed, untreated	25.4%	20.5%	—	—	—

Core within-person contrasts (n = 17,901): diagnosis and treatment rates by condition with the within-person difference. Formal within-person tests: McNemar's test (continuity-corrected) and conditional logistic regression (n = 17,838, both diagnoses observed) and TOST equivalence testing (n = 17,843, both

treatment items observed). Four-way decomposition (n = 17,783, all four items observed). Confidence intervals use design degrees of freedom; odds-ratio intervals are on the log scale. clogit = conditional logistic regression.

3.3 Four-way decomposition

The four-way decomposition (Table 2, Group 3) localizes where the gap sits:

Adults treated for anxiety without an anxiety diagnosis outnumbered those treated for depression without a depression diagnosis (12.3% vs 5.6%). Within the former group, 63.6% had a depression diagnosis and 59.3% were taking antidepressant medication.

3.4 Treatment without a label

The proportion of treated adults lacking the matching diagnosis was higher for anxiety than for depression in every one of the seven waves (difference + 13.5, + 13.0, + 11.8, + 11.1, + 9.3, + 10.2 and + 7.9 pp from 2019 to 2025). Among adults treated for anxiety without an anxiety diagnosis (n = 2,358), 63.6% carried a depression diagnosis; conversely, among adults treated for depression without a depression diagnosis (n = 1,058), only 25.1% carried an anxiety diagnosis. This asymmetry suggests a labeling preference: clinicians appear to record depression when both are present, and anxiety disproportionately when depression is absent.

Recall bias. If anxiety diagnoses were merely less memorable, then among people who are certainly in contact with the system—those currently taking psychotropic medication for a named condition—the missing-label rate should be similar for the two conditions. It is not: among the 8,994 adults taking anxiety medication, 15.4% (95% CI 14.5 to 16.3) had no anxiety diagnosis, against 5.2% (4.7 to 5.7) of the 9,035 adults taking depression medication.

An adult taking anxiety medication is 3.0 times as likely to lack an anxiety diagnosis as an adult taking depression medication is to lack a depression diagnosis (difference 10.2 pp). Because both groups have equivalent evidence of having been in care, the asymmetry is more consistent with an anxiety label never having been recorded than with patients forgetting it. This does not exclude recall bias, but it constrains it.

3.5 The gap is not a consequence of comorbidity

Because the analytic sample is defined by having both symptom types, a natural objection is that the diagnosis gap is a mechanical consequence of comorbidity: when a patient presents with both, clinicians may record only the depression diagnosis, so anxiety would appear under-labeled by construction. The observed co-diagnosis pattern is consistent with this reading (58.6% both, 15.3% only depression, 3.7% only anxiety, 22.5% neither).

This pattern is not unprecedented: in UK cohort data, adults with anxiety–anxiety comorbidity were 60% less likely to self-report having received an anxiety or depressive diagnosis than those with anxiety–

depression comorbidity, despite similarly severe symptoms (Davies et al., 2023). Conversely, we tested the objection directly by moving outside the comorbid sample. If the gap were confined to comorbid presentations, adults with only anxiety symptoms should be diagnosed at rates closer to those with only depression symptoms. The pattern does not support that prediction.

Anxiety is diagnosed at roughly half the rate of depression among adults with a single symptom type. The direction matches the comorbid sample. If anything, the comorbid setting raises both rates (62.2% and 73.8% vs 33.2% and 62.8%) without closing the difference. The gap is therefore not created by conditioning on comorbidity.

Two caveats apply. The single-symptom comparison uses different denominators and is descriptive rather than a formal test; and the depression-only group is an order of magnitude smaller (3,012 vs 37,269), so its estimate is less precise.

3.6 Severity-matched comparison (2019, 2022, 2025)

In the three waves that administered the GAD-7 and PHQ-8, computed scores agreed with the official categorical variables in 100.0% of 83,863 records. Within severity bands, the diagnosis gap was not constant, and the pattern was consistent across waves (Table 3; Fig. 1):

Table 3
Severity-matched diagnosis and treatment by wave.

Severity band / stratum	n	Anxiety dx	Depression dx	Gap (pp)	95% CI	p
Four severity bands (pooled across waves)						
None (0–4)	61,287	8.1%	7.8%	+ 0.3	0.1 to 0.6	0.017
Mild (5–9)	3,799	45.0%	46.3%	−1.3	−3.1 to 0.4	0.139
Moderate (10–14)	1,005	62.3%	69.7%	−7.3	−10.4 to − 4.3	< 0.001
Severe (15+)	1,142	75.2%	81.3%	−6.1	−8.5 to − 3.7	< 0.001
Two strata (pooled across waves)						
Low (0–9)	65,086	10.4%	10.2%	+ 0.2	−0.1 to 0.5	0.110
High (10+)	2,147	69.2%	75.8%	−6.7	−8.6 to − 4.7	< 0.001
High severity, by wave						
2019	863	67.2%	74.7%	−7.5	−10.8 to − 4.1	< 0.001
2022	675	71.2%	79.7%	−8.5	−11.8 to − 5.2	< 0.001
2025	609	69.2%	73.4%	−4.2	−7.6 to − 0.8	0.015
Low severity, by wave						
2019	25,153	8.1%	9.1%	−0.9	−1.3 to − 0.5	< 0.001
2022	21,387	11.1%	10.6%	+ 0.5	0.0 to 1.0	0.056
2025	18,546	12.0%	10.9%	+ 1.1	0.6 to 1.6	< 0.001

Four blocks: (i) the within-person diagnosis gap by GAD-7/PHQ-8 severity band, pooled across 2019, 2022 and 2025; (ii) the same comparison for two pooled strata (low 0–9 vs high 10+); (iii) the high-severity gap in each wave separately; (iv) the low-severity gap in each wave separately. dx = diagnosis.

The severity dependence is stable across survey years. Because the GAD-7 and PHQ-8 were administered in three waves rather than one, we could replicate the severity analysis in each. The high-

severity gap (both scales ≥ 10) was present and statistically significant in all three years (Table 3, “High severity, by wave”), while the low-severity gap was near zero throughout. The severity dependence therefore does not rest on a single wave or a single small cell. Formal heterogeneity testing supported pooling: Cochran’s Q was 3.4 (df = 2, p = 0.18, I^2=41.5%).

The gap is severity-dependent, and its direction is informative. At low severity the two conditions are labeled at nearly the same rate. From the moderate band upward, however, the anxiety label is systematically withheld: +0.3 pp in the none band, − 1.3 pp in the mild band, and − 7.3 and − 6.1 pp in the moderate and severe bands. From the lowest to the highest band the anxiety diagnosis rate rises by 67.1 pp (8.1% to 75.2%) and the depression rate by 73.5 pp (7.8% to 81.3%). Depression therefore gained 6.4 points more over the same range, close to the severe-band gap. Under the strictest matching (identical GAD-7 and PHQ-8 totals), the gap is − 3.0 pp (n = 1,315, totals ≥ 5) and − 4.5 pp (n = 393, totals ≥ 10).

This pattern is consistent with a threshold account. When symptoms are mild, neither condition is labeled much and there is little differential to explain. Once symptoms are severe enough that the encounter is likely to produce any diagnosis, the choice of which condition to record becomes consequential, and the depression label is preferred. Under this reading the main analysis’s − 11.7 pp gap is an average over a severity distribution rather than a uniform penalty. The gradient is not strictly monotonic (the severe-band gap is slightly smaller than the moderate-band gap), and the moderate band contains only 1,005 respondents, so its fine structure should be interpreted with caution.

Treatment comparison. Matching on severity, anxiety medication was more common than depression medication in three of the four bands: none 6.8% vs 4.8% (+ 2.0 pp, 95% CI 1.8 to 2.2), mild 32.9% vs 28.2% (+ 4.7 pp, 3.4 to 5.9), and moderate 43.4% vs 40.0% (+ 3.2 pp, 0.8 to 5.5). The two were indistinguishable in the severe band (53.8% vs 52.8%, + 1.1 pp, − 1.1 to 3.2, p = 0.328). The direction of this medication contrast does not track the diagnosis gap: medication favoring anxiety is present at every severity level, whereas the diagnosis gap is absent at low severity and only emerges at moderate-to-severe severity. Anxiety is therefore more likely to be medicated wherever a difference exists, even though it is less likely to be labeled.

Bar chart of the within-person difference (anxiety minus depression diagnosis rate) across four severity bands defined by matching GAD-7 and PHQ-8 totals. Error bars show 95% confidence intervals; n shown below each band.

3.7 Sensitivity analyses

Diagnosis rates rose with symptom frequency for both conditions, but the anxiety rate remained below the depression rate at every level (Table 4, symptom frequency group):

These rates are estimated in the full adult population, not the 17,901-person analytic sample; the anxiety and depression rates at a given level come from different sub-populations, so this is a comparison of marginal distributions. When all adults are counted, with missing symptom responses treated as “never”, 27.7% reported weekly-or-more anxiety symptoms versus 9.9% for depressive symptoms.

Notably, adjusting the within-person comparison for symptom frequency **widened** rather than narrowed the gap: the conditional logit OR for depression versus anxiety rose from 4.44 (4.08 to 4.83) unadjusted to 5.47 (4.95 to 6.05) after adjustment, equivalent to 0.183 (0.165 to 0.202) for anxiety versus depression. The gap therefore cannot be attributed to the two domains having different symptom-frequency distributions.

Treatment measurement asymmetry. Medication items carry no recall window; psychotherapy refers to the past 12 months. Rebuilding the treatment comparison from medication items alone gave the same result: anxiety 48.9% versus depression 48.6%, difference 0.22 pp.

Daily symptom threshold. Restricting to adults with daily symptoms of both types (n = 6,500) preserved the pattern: diagnosis 69.8% vs 82.1% (12.3 pp).

By survey year. The diagnosis gap narrowed monotonically across the seven waves: **−14.1 pp (2019), −13.2 pp (2020), − 12.4 pp (2021), − 11.5 pp (2022), − 10.9 pp (2023), − 10.4 pp (2024), − 9.8 pp (2025)**; a linear trend in the difference was + 0.72 pp per year (SE 0.031, $p < 0.001$). (Fig. 2) The narrowing trend may reflect increased anxiety awareness after the COVID-19 pandemic, expanded screening recommendations (US Preventive Services Task Force, 2023), or shifting diagnostic preferences as anxiety disorders gained visibility in public discourse. NHIS data cannot distinguish these mechanisms, but the acceleration after 2020 is consistent with heightened attention to anxiety.

Line plot of the within-person difference in diagnosis rate (anxiety minus depression) among dual-symptom adults (n = 17,838), by survey year. Point labels show the gap in percentage points; error bars show 95% confidence intervals; the dotted line is the ordinary least-squares trend (+ 0.72 pp per year, SE 0.031, $p < 0.001$).

Excluding 2020. Because the 2020 wave moved from in-person to telephone interviewing, we repeated the analysis without it (n = 15,247). The overall gap was − 11.5 pp (95% CI − 12.3 to − 10.6), essentially identical to the full-sample estimate (− 11.7 pp), and the narrowing trend was unchanged when 2020 was excluded.

Subgroup consistency. We repeated the diagnosis contrast in ten pre-specified subgroups (sex, age, insurance, usual source of care, symptom frequency). The gap was present and significant in all ten (range − 15.0 to − 7.1 pp), and all ten survived Benjamini-Hochberg and Bonferroni correction (all $p < 0.001$). The direction never reversed; the gap was larger among adults aged 45+ (− 15.0 pp) than 18–44 (− 8.7 pp) and among men (− 12.8 pp) than women (− 10.9 pp).

Usual source of care. The label gap is not primarily driven by access to care. Restricting to adults with a usual source of care (97.6% of symptomatic adults) gave essentially the same gap, and stratifying all outcomes shows a similar contrast in the two access strata (Table 4, usual source of care group):

The diagnosis gap was present in both strata and the treatment gap near zero in both. Adults without a usual source of care received somewhat less treatment, but the relative labeling of the two conditions

was unchanged; if anxiety were under-labeled merely because people lack access, the gap should shrink in the group with established care. It does not, although the no-usual-care stratum is small (n = 412).

Table 4
Sensitivity analyses.

Analysis	Anxiety	Depression	Conclusion
Symptom frequency (full population)			
Frequency: daily	53.1%	79.9%	Unchanged
Frequency: weekly	32.6%	67.5%	Unchanged
Frequency: monthly	17.9%	48.2%	Unchanged
Frequency: a few times a year	7.7%	18.6%	Unchanged
Frequency: never	2.1%	2.6%	Unchanged
Frequency-adjusted OR	OR (anx vs dep) 0.183 (0.165–0.202)	OR (dep vs anx) 5.47 (4.95–6.05)	Unchanged, gap widens
Usual source of care			
Usual care: has (n = 17,418)	62.1%	73.9%	Unchanged
Usual care: none (n = 412)	66.8%	73.8%	Unchanged
Diagnosis gap, has vs none	+ 11.8 pp	+ 7.1 pp	Unchanged
Treatment gap, has vs none	+ 0.2 pp	+ 0.5 pp	Unchanged
Other sensitivity analyses			
Daily symptom threshold	69.8%	82.1%	Unchanged
Excluding 2020 (n = 15,247)	−11.5 pp	—	Unchanged, trend + 0.72/yr
Ten pre-specified subgroups	−15.0 to − 7.1 pp	all significant	Unchanged
Multiplicity correction	all 10 survive BH and Bonferroni	all $p < 0.001$	Unchanged
Severity bands pooled	−6.7 pp	$p < 0.001$ all waves	Unchanged
Strict severity match	−3.0 pp	n = 1,315	Unchanged

Symptom frequency (full population): diagnosis rates by symptom frequency and the frequency-adjusted contrast. Usual source of care: outcomes stratified by access stratum. Other sensitivity analyses: remaining analyses. The final column states whether the main conclusion was unchanged. BH = Benjamini-Hochberg.

3.8 Population context

Population context. The estimates in this section cover all adults reporting weekly-or-more anxiety symptoms, not the dual-symptom group used in Sections 3.2–3.6. Among US adults 2019–2025, 28.3% had weekly-or-more anxiety symptoms. Within this symptomatic group, 42.2% had ever been diagnosed and 43.7% were receiving treatment; at the population level this corresponds to 11.9% of all adults with both symptoms and a diagnosis, 12.4% with both symptoms and treatment, and 9.0% with both. Notably, 3.4% of all adults—12.1% of symptomatic adults—were taking treatment for anxiety without ever being told they had it. This gap was consistent across all seven years. In 2025, an estimated 9.1 million US adults (95% CI 8.3 to 9.8 million) were taking anxiety medication without a recorded diagnosis. Care intensity rose steeply with symptom burden: anxiety treatment rates were 9.4% among adults with neither symptom type, 35.9% among those with one symptom type, and 61.5% among those with both.

4. Discussion

4.1 Principal findings

Among US adults with both anxiety and depressive symptoms, anxiety was diagnosed 11.7 percentage points less often than depression (62.2% vs 73.8%), while treatment rates were formally equivalent (difference 0.25 pp, TOST equivalence at ± 1 pp). The diagnosis gap was confirmed by McNemar's test and conditional logistic regression (OR = 0.225 for anxiety relative to depression), was severity-dependent and largest at moderate symptom levels, and was present in all three waves that administered the GAD-7 and PHQ-8. That the gap has narrowed only modestly over seven years is consistent with evidence that clinical detection of anxiety has not improved over decades: across 24 primary-care studies and 34,902 patients, pooled diagnostic sensitivity for anxiety disorders was 44.5%, with no change over time (Olariu et al., 2015). Decomposition localized the gap to a large group treated for anxiety without an anxiety diagnosis (12.3%).

Four analyses address the main alternatives. First, among adults taking psychotropic medication for a named condition, the missing-label rate was 3.0 times higher for anxiety than depression (15.4% vs 5.2%), more consistent with a label never recorded than with unequal recall. Second, the gap is present among adults with only anxiety symptoms (33.2% vs 62.8%). Third, pooling the moderate and severe bands leaves the gap highly significant (− 6.7 pp, $p < 0.001$) in each wave. Fourth, the medication contrast runs in the opposite direction to the labeling contrast (+ 2.0 to + 4.7 pp), so under-labeling coexists with equal or greater pharmacological treatment.

Relationship to prior work. Earlier primary-care studies documented low recognition and treatment of mood and anxiety disorders (Lecrubier, 2007). Two findings bear directly on our results. First, and closest to our question, a primary-care chart review found generalized anxiety undetected more often than major depression (71.0% vs 65.9%), and near-parity between the two when any recorded mental-health diagnosis was accepted (55.0% vs 55.5%; Vermani et al., 2011). That study compared different patients with different disorders, so the observed asymmetry could reflect who presents for care as much as how clinicians label. Second, the treatment gap conventionally reported for anxiety—lower psychotropic use among people with anxiety disorders than among people with mood disorders (32.4% vs 45.6% in the ESEMeD surveys; Lecrubier, 2007)—also rests on different people with different disorders and could reflect the same confounders. Our paired design isolates the clinician's response directly: within the same individual, and matched on GAD-7 and PHQ-8 severity, anxiety was diagnosed 11.7 percentage points less often than depression yet treated at an equivalent rate (0.25 pp; TOST $p < 0.001$). The excess of untreated anxiety in cross-person data is therefore better described as a deficit in labeling than as a deficit in treatment. Our treatment measure—medication or psychotherapy in the past 12 months—remains distinct from the World Mental Health Surveys, which examined 12-month DSM-IV anxiety disorders (27.6% any treatment; Alonso et al., 2018). The poor agreement between insurance-physician ICD-10 codes and CIDI-based DSM-IV classifications among disability claimants (Cornelius et al., 2014) shows that under-recognition can arise in particular assessment settings, although that population and process differ from ours. The severity-dependent pattern we observe has not, to our knowledge, been reported in prior paired designs.

4.2 Why this matters

The identification gap for anxiety is a specific feature of clinical practice, not a compositional artifact. The gap is a **labeling asymmetry**, not a barrier to access: patients may be receiving anxiety treatment without a formal diagnosis on record. This has consequences for treatment planning, patient education, and the interpretation of treatment-effectiveness research. The severity gradient suggests that the pattern is consistent with a preferential attribution of the diagnostic label to depression when both are present and severe enough to warrant any label, rather than indifference to anxiety. A chart-based primary-care study is instructive here because it separates recognition from labeling. Among adults who met MINI criteria for a disorder, clinicians had noted some mental-health problem in 45.0% of generalized-anxiety cases and 44.5% of depression cases—nearly identical. But the specific matching diagnosis appeared in only 29.0% and 34.1% respectively: the gap between any recognition and an exact diagnosis was 16.0 percentage points for anxiety and 10.4 for depression (Vermani et al., 2011). Distress is noticed; the anxiety diagnosis is less often written down. That sample was recruited from clinic waiting rooms and may over-represent patients with more severe symptoms. For patients, an unlabeled diagnosis has practical consequences: medication adherence, self-management strategies, and insurance documentation all depend on having an explicit diagnosis on record.

The larger gap among adults aged 45+ (− 15.0 pp) than 18–44 (− 8.7 pp) and among men (− 12.8 pp) than women (− 10.9 pp) may reflect generational differences in mental health literacy or gender-specific

help-seeking patterns. Older adults and men may be less likely to articulate anxiety symptoms, giving clinicians fewer cues to record the anxiety label. Community evidence is consistent with an upstream recognition pathway: respondents classified depressive symptoms more accurately than anxiety symptoms (49.5% versus 44.2%), and older adults were less accurate overall at identifying psychiatric symptoms (Wetherell et al., 2009). That study assessed lay symptom recognition rather than clinician labeling, so it does not explain the observed gap directly. The subgroup analysis also covered race and ethnicity; the gap was present in all strata, suggesting the labeling asymmetry is not confined to any single racial or ethnic group. However, NHIS sample sizes within some strata limit precision, and the intersection of race, ethnicity, and diagnostic labeling warrants dedicated study.

Plausible mechanisms and consistency with observed patterns. We tested three mechanism-specific predictions using available NHIS data:

1. Clinician preference: if depression is preferentially recorded when both criteria are met, the four-way decomposition should show excess “diagnosed depression + treated anxiety.” This pattern is present (12.3% vs 5.6%), consistent with but not specific to preferential recording.
2. Differential recall: if anxiety diagnoses are less memorable, the medication-user analysis should show higher missing-label rates. This is supported (3.0× higher for anxiety), consistent with patients being less able to recall anxiety labels.
3. Measurement threshold: if diagnosis depends on severity thresholds, the gap should emerge only at moderate+ severity. This is partially supported: the gap emerges at moderate severity (+ 0.3 → −7.3 pp) but does not increase further at severe (− 6.1 pp), indicating a threshold model with non-monotonic detail.

All three predictions receive support, but these patterns are consistent with rather than specific to any single mechanism. Prospective clinical observation, chart review, or clinician survey would be required to attribute the observed asymmetry to a particular mechanism.

Several mechanisms could produce this asymmetry. First, depression has a longer diagnostic tradition in primary care: the PHQ-2/PHQ-9 screening cascade is widely embedded in US quality metrics (e.g., HEDIS; NCQA, 2023), whereas anxiety screening is not yet similarly incorporated, creating institutional momentum toward the depression label. Second, the diagnostic hierarchy in DSM-5 may favor depression (American Psychiatric Association, 2022): when both criteria are met, clinicians may record the “primary” mood diagnosis. Third, antidepressants are marketed for both conditions, so treatment does not force the diagnostic choice, whereas the anxiety label may carry less reimbursement weight. These mechanisms are not mutually exclusive and cannot be separated with NHIS data.

4.3 Strengths and limitations

Strengths. The within-person design removes person-level confounding; the two symptom items use identical wording and response scales; the severity-matched analysis spans three waves with 100%

agreement between computed and official GAD-7/PHQ-8 scores, and the formal within-person tests (McNemar, conditional logit, TOST) confirm the marginal contrast.

Limitations. (i) Diagnosis is self-reported and lifetime; recall bias is possible, and its direction is likely to **overstate** the gap, but the magnitude cannot be identified from these data. The analysis of medication users (§ 3.4) constrains but does not eliminate this bias: within that group the missing-label rate is three times higher for anxiety than for depression, which is hard to reconcile with pure recall.

1. Depression is a comparator, not a gold standard. The labeling interpretation rests on the premise that depression is screened at least as often as anxiety in primary care, which we did not observe. Subject to this premise, over-diagnosis of either condition could bias the absolute gap in either direction. However, the primary finding is the within-person asymmetry rather than the absolute gap size: for the same individual with the same symptom burden, the depression label is recorded more often than the anxiety label. This asymmetry is robust to common-mode biases that affect both conditions equally, though not to biases that affect the two conditions differently. The direction of any misclassification bias is toward under-reporting, though its magnitude cannot be quantified (see Methods).
2. The psychotherapy item has a 12-month window whereas medication items have none, and the therapy item does not distinguish anxiety from depression; the treatment comparison rests on the medication channel.
3. The GAD-7 and PHQ-8 share somatic items (Spitzer et al., 2006; Kroenke et al., 2009) so severity-matched groups are not matched on clinical severity beyond the total score.
4. The GAD-7 and PHQ-8 were administered in only three waves (2019, 2022, 2025), so the severity analysis cannot be trended across the full 2019–2025 period.
5. The no-usual-care stratum contains only 412 respondents, limiting the precision of the access-stratified comparison.
6. Symptom frequency reflects current status whereas diagnosis is lifetime; this time-window mismatch may bias the diagnosis gap upward. However, the treatment comparison—our primary evidence for equivalence—uses current medication, which shares the symptom time window; the TOST equivalence result is therefore unaffected by this mismatch. The E-value for the diagnosis gap (OR 4.44 for depression vs anxiety) was 3.6, indicating that an unmeasured confounder would need to be associated with both symptom type and diagnosis by a risk ratio of 3.6-fold each to explain away the observed gap. Sensitivity analyses showed E-values ranged from 3.5 (lower CI bound) to 3.8 (upper CI bound), and this magnitude exceeds that of typical common confounders (age, sex, insurance: RR 1.2–2.0), suggesting the finding is robust to plausible unmeasured confounding.

4.4 Conclusions

Among US adults with equivalent symptom burdens, anxiety is diagnosed less often than depression while being treated at the same rate. The asymmetry is severity-dependent, stable across survey years, robust to multiple sensitivity analyses, and not a consequence of comorbidity or measurement artifacts.

Closing the identification gap therefore requires interventions that target labeling—prompting clinicians to record the anxiety diagnosis when warranted—rather than access alone. Future work should test whether structured diagnostic prompts in primary care reduce this asymmetry, and whether the gap varies across health systems with different screening incentives. Concrete interventions could include: (i) adding an explicit anxiety documentation field to electronic health records when GAD-7 ≥ 10; (ii) including anxiety diagnosis rates in quality metrics alongside the existing PHQ-2/PHQ-9 depression cascade; and (iii) clinician education on documenting comorbid anxiety when both criteria are met.

Declarations

Data availability

The NHIS public-use data are available from the National Center for Health Statistics (https://www.cdc.gov/nchs/nhis/). Analysis code and the automated verification suite will be deposited in a public repository upon acceptance.

Ethics

This study used publicly available, de-identified NHIS data and required no ethics approval or informed consent.

Declaration of competing interest

The author declares no competing interests.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

CRediT authorship contribution statement

Yueqi Cui: Conceptualization, Methodology, Software, Formal analysis, Data curation, Writing – original draft, Writing – review and editing, Visualization. The author has read and approved the final manuscript.

Declaration of generative AI and AI-assisted technologies in the manuscript preparation process

During the preparation of this work the author used DeepSeek-V4.1-Flash and Kimi-K3 for language polishing, translation, reference formatting, and manuscript consistency checking. The same tools were also used to provide guidance on writing the analysis code used in this study, and to assist with that code's optimization and review. After using these tools, the author reviewed and edited the content as needed and takes full responsibility for the content of the published article.

Acknowledgments

The author thanks the National Center for Health Statistics for making the National Health Interview Survey public-use files available. This study builds on a broader body of prior research on the epidemiology and clinical recognition of anxiety and depressive disorders, to which the author is indebted.

References

1. Alonso J, Liu Z, Evans-Lacko S et al (2018) Treatment gap for anxiety disorders is global: results of the World Mental Health Surveys in 21 countries. Depress Anxiety 35:195–208. https://doi.org/10.1002/da.22711
2. American Psychiatric Association (2022) Diagnostic and Statistical Manual of Mental Disorders, fifth edn. American Psychiatric Association, Washington, DC
3. van der Cornelius B, Brouwer S, Groothoff JW (2014) Under-recognition and under-treatment of DSM-IV classified mood and anxiety disorders among disability claimants. Disabil Rehabil 36:1161–1168. https://doi.org/10.3109/09638288.2013.833310
4. Daly M, Robinson E (2021) Anxiety reported by US adults in 2019 and during the 2020 COVID-19 pandemic: population-based evidence from two nationally representative samples. J Affect Disord 286:296–300. https://doi.org/10.1016/j.jad.2021.02.054
5. Davies MR, Glen K, Mundy J, ter Kuile AR, Adey BN, Armour C, Assary E, Coleman JRI (2023) Factors associated with anxiety disorder comorbidity. J Affect Disord 323:280–291. https://doi.org/10.1016/j.jad.2022.11.051
6. Evans-Lacko S, Aguilar-Gaxiola S, Al-Hamzawi A et al (2018) Socio-economic variations in the mental health treatment gap for people with anxiety, mood, and substance use disorders: results from the WHO World Mental Health (WMH) surveys. Psychol Med 48:1560–1571. https://doi.org/10.1017/S0033291717003336
7. Hosmer DW, Lemeshow S (2000) Applied Logistic Regression, second edn. Wiley, New York. (Chap. 7)
8. Kroenke K, Strine TW, Spitzer RL, Williams JBW, Berry JO, Mokdad AH (2009) The PHQ-8 as a measure of current depression in the general population. J Affect Disord 114:163–173. https://doi.org/10.1016/j.jad.2008.06.026
9. Lakens D (2017) Equivalence tests: a practical primer for t tests, correlations, and meta-analyses. Soc Psychol Personal Sci 8:355–362. https://doi.org/10.1177/1948550617697177
10. Lecrubier Y (2007) Widespread underrecognition and undertreatment of anxiety and mood disorders: results from 3 European studies. J Clin Psychiatry 68(Suppl 2):36–41
11. Leikauf J, Federman AD (2009) Comparisons of self-reported and chart-identified chronic diseases in inner-city seniors. J Am Geriatr Soc 57:1219–1225. https://doi.org/10.1111/j.1532-5415.2009.02313.x

12. Mojtabai R, Jorm AF (2015) Trends in psychological distress, depressive episodes and mental health treatment-seeking in the United States: 2001–2012. J Affect Disord 174:556–561. https://doi.org/10.1016/j.jad.2014.12.039
13. National Center for Health Statistics, 2020–2026. National Health Interview Survey, 2019–2025. Public-use data files. NCHS, Hyattsville MD Available from: https://www.cdc.gov/nchs/nhis/
14. National Committee for Quality Assurance (NCQA) (2023) HEDIS measures. NCQA, Washington, DC
15. Olariu E, Forero CG, Castro-Rodriguez JI, Rodrigo-Calvo MT, Álvarez P, Martín-López LM, Sánchez-Toto A, Adroher ND, Blasco-Cubedo MJ, Vilagut G, Fullana MA, Alonso J (2015) Detection of anxiety disorders in primary care: a meta-analysis of assisted and unassisted diagnoses. Depress Anxiety 32:471–484. https://doi.org/10.1002/da.22360
16. Penninx BW, Pine DS, Holmes EA, Reif A (2021) Anxiety disorders. Lancet 397:914–927. https://doi.org/10.1016/s0140-6736(21)00359-7
17. Sanchez-Villegas A, Schlatter J, Ortuño F, Lahortiga F, Pla J, Benito S, Martinez-Gonzalez MA (2008) Validity of a self-reported diagnosis of depression among participants in a cohort study using the Structured Clinical Interview for DSM-IV (SCID-I). BMC Psychiatry 8:43. https://doi.org/10.1186/1471-244X-8-43
18. Sarkar AA, Alam MF, Ahmed HU, Alam MT, Khan NM (2026) Prevalence, symptom profile, associated factors, and treatment gap of depressive disorders among adults: findings from a nationwide household survey in Bangladesh. J Affect Disord 402:121370. https://doi.org/10.1016/j.jad.2026.121370
19. Schuirmann DJ (1987) A comparison of the two one-sided tests procedure and the power approach for assessing the equivalence of average bioavailability. J Pharmacokinet Biopharm 15:657–680. https://doi.org/10.1007/BF01068419
20. Spitzer RL, Kroenke K, Williams JBW, Löwe B (2006) A brief measure for assessing generalized anxiety disorder: the GAD-7. Arch Intern Med 166:1092–1097. https://doi.org/10.1001/archinte.166.10.1092
21. US Preventive Services Task Force (2023) Screening for anxiety disorders in adults: US Preventive Services Task Force recommendation statement. JAMA 329:2163–2170. https://doi.org/10.1001/jama.2023.9301
22. VanderWeele TJ, Ding P (2017) Sensitivity analysis in observational research: introducing the E-value. Ann Intern Med 167:268–274. https://doi.org/10.7326/m16-2607
23. Vermani M, Marcus M, Katzman MA (2011) Rates of detection of mood and anxiety disorders in primary care: a descriptive, cross-sectional study. Prim Care Companion CNS Disord 13. https://doi.org/10.4088/pcc.10m01013. PCC.10m01013
24. Wang PS, Aguilar-Gaxiola S, Alonso J et al (2007) Use of mental health services for anxiety, mood, and substance disorders in 17 countries in the WHO world mental health surveys. Lancet 370:841–850. https://doi.org/10.1016/S0140-6736(07)61414-7

25. Wetherell JL, Petkus AJ, McChesney K, Stein MB, Judd PH, Rockwell E, Sewell DD, Patterson TL (2009) Older adults are less accurate than younger adults at identifying symptoms of anxiety and depression. J Nerv Ment Dis 197:623–626. https://doi.org/10.1097/NMD.0b013e3181b0c081
26. Wu C-S, Chen K-Y, Chen P-Q, Lin M-C, Lin Y-F, Fan C-C, Huang W-L, Wang S-H (2027) Self-reported and recent claims-based major depressive disorder identifies distinct risk profiles in a nationwide population study. J Affect Disord 416:122561. https://doi.org/10.1016/j.jad.2026.122561
27. Zablotsky B, Weeks JD, Terlizzi EP, Madans JH, Blumberg SJ (2022) Assessing anxiety and depression: a comparison of National Health Interview Survey measures. Natl Health Stat Rep 172:1–17. https://doi.org/10.15620/cdc:117491

Figures

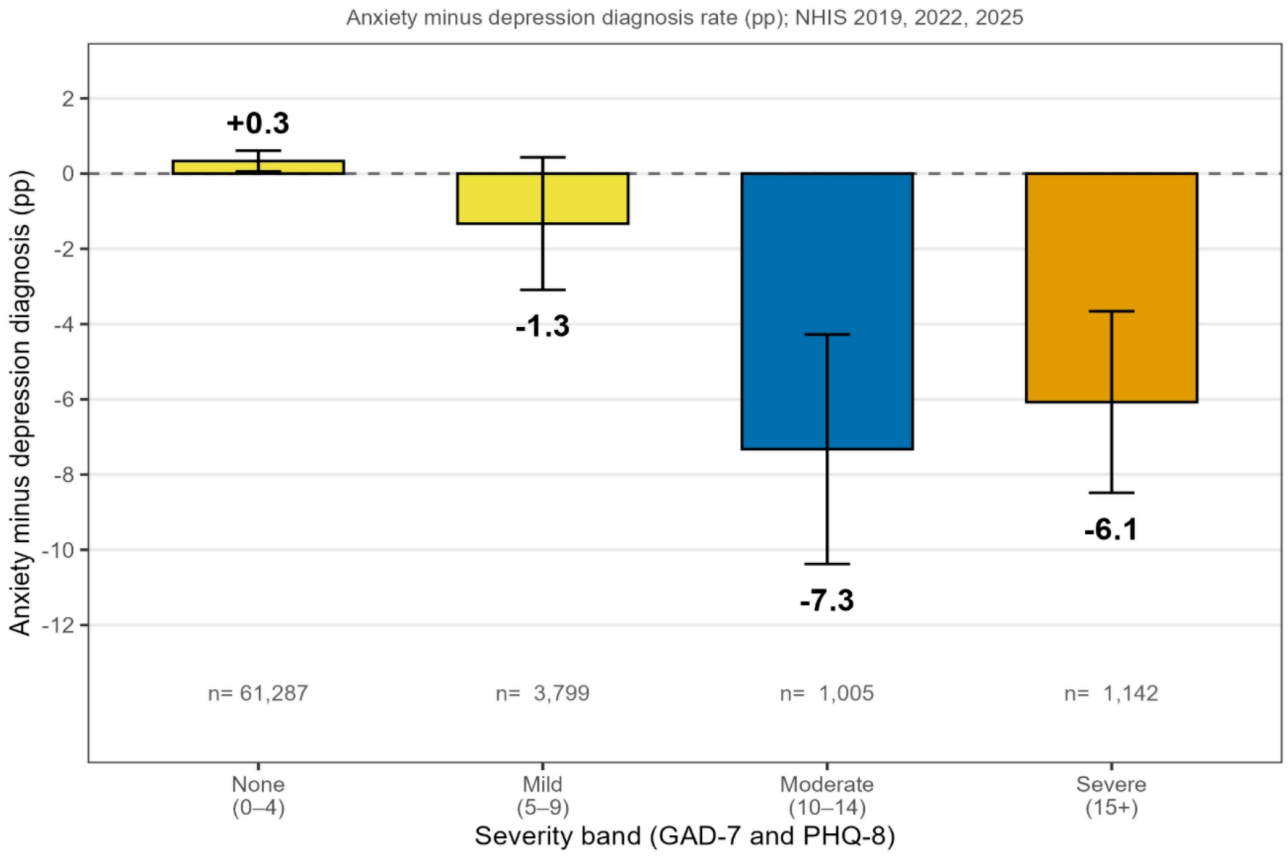


Figure 1

The diagnosis gap is severity-dependent (NHIS 2019, 2022, 2025).

Bar chart of the within-person difference (anxiety minus depression diagnosis rate) across four severity bands defined by matching GAD-7 and PHQ-8 totals. Error bars show 95% confidence intervals; n shown

below each band.

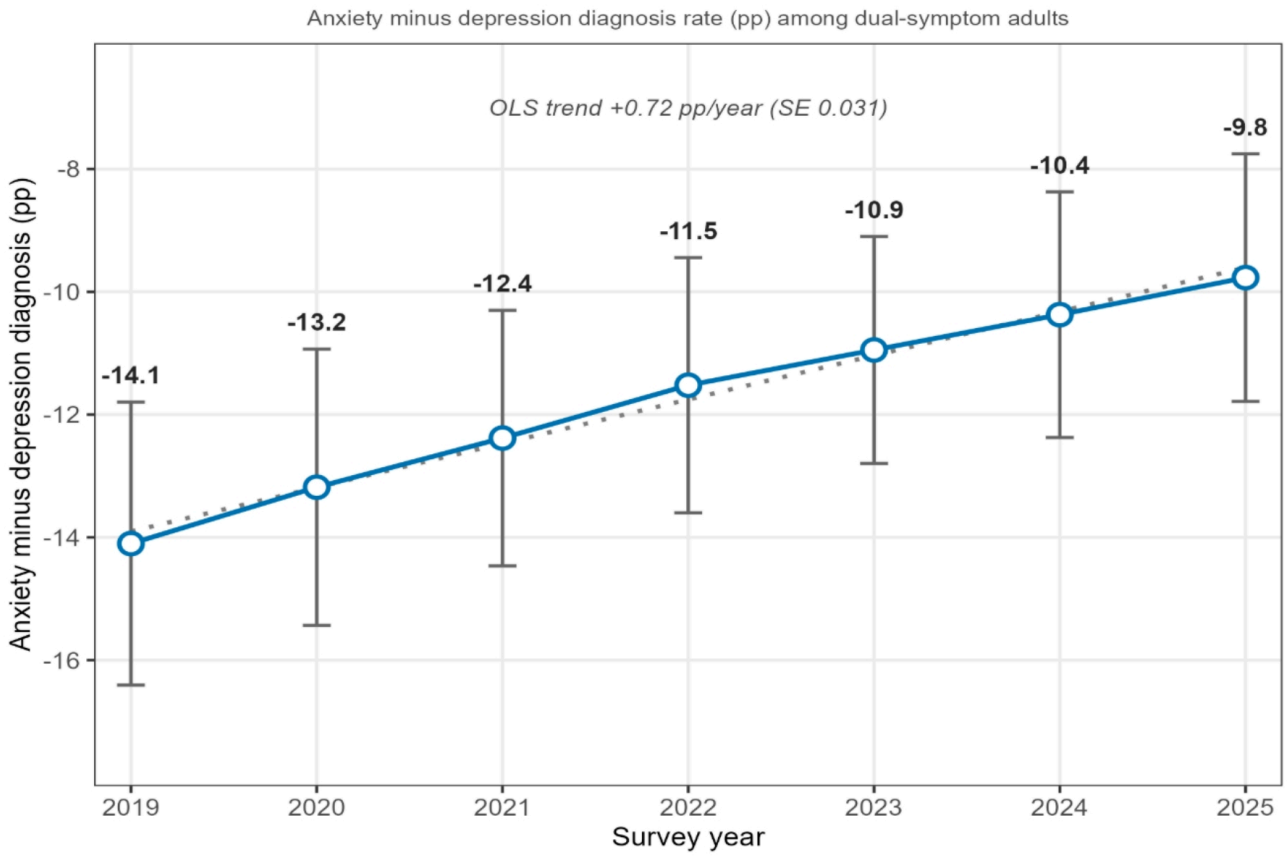


Figure 2

The diagnosis gap narrows over time, 2019–2025.

Line plot of the within-person difference in diagnosis rate (anxiety minus depression) among dual-symptom adults (n=17,838), by survey year. Point labels show the gap in percentage points; error bars show 95% confidence intervals; the dotted line is the ordinary least-squares trend (+0.72 pp per year, SE 0.031, p<0.001).