Does Suicide Risk Screening Improve the Identification of Primary Care Patients Who Will Attempt Suicide Versus Depression Screening Alone?

Craig. J. Bryan PsyD, ABPP , Michael H. Allen MD , AnnaBelle O. Bryan MS , Cynthia J. Thomsen PhD , Justin C. Baker PhD, ABPP , Alexis M. May PhD

PII: S1553-7250(23)00206-4
DOI: https://doi.org/10.1016/j.jcjq.2023.08.008
Reference: JCJQ 738

To appear in: *The Joint Commission Journal on Quality and Patient Safety*

Please cite this article as: Craig. J. Bryan PsyD, ABPP , Michael H. Allen MD , AnnaBelle O. Bryan MS , Cynthia J. Thomsen PhD , Justin C. Baker PhD, ABPP , Alexis M. May PhD , Does Suicide Risk Screening Improve the Identification of Primary Care Patients Who Will Attempt Suicide Versus Depression Screening Alone?, *The Joint Commission Journal on Quality and Patient Safety* (2023), doi: https://doi.org/10.1016/j.jcjq.2023.08.008

Does Suicide Risk Screening Improve the Identification of Primary Care Patients Who Will Attempt Suicide Versus Depression Screening Alone?

Craig. J. Bryan, PsyD, ABPP [a]
Michael H. Allen, MD [b]
AnnaBelle O. Bryan, MS [a]
Cynthia J. Thomsen, PhD [c]
Justin C. Baker, PhD, ABPP [a]
Alexis M. May, PhD [d]

[a] Department of Psychiatry and Behavioral Health, The Ohio State University College of Medicine, 3650 Olentangy River Rd, Suite 310, Columbus, OH 43214
[b] Department of Psychiatry, University of Colorado Anschutz Medical Campus, 13199 E. Montview Blvd, Suite 330, MS F550, Aurora, CO 80045
[c] Naval Health Research Center, 140 Sylvester Road, San Diego, CA 92106
[d] Department of Psychology, Wesleyan University, Judd Hall, Room 201D, 207 High Street, Middletown, CT 06459

Corresponding Author: Craig J. Bryan, PsyD, ABPP, Department of Psychiatry & Behavioral Health, The Ohio State University Wexner Medical Center, 3650 Olentangy River Rd, Suite 310, Columbus, OH 43214; email: craig.bryan@osumc.edu

Lab report 22-74 was supported by the Office of the Assistant Secretary of Defense for Health Affairs, through the Defense Medical Research and Development Program under Award No. W81XWH-14-1-0272 (PI: Bryan) and work unit no. N1426. The views expressed in this article are those of the authors and do not necessarily reflect the official policy or position of the Department of the Navy, Department of Defense, nor the U.S. Government.

The study protocol was approved by the Naval Health Research Center's Institutional Review Board in compliance with all applicable Federal regulations governing the protection of human subjects. Research data were derived from an approved Naval Health Research Center Institutional Review Board protocol, number NHRC.2014.0046.

ABSTRACT

Objective: The effectiveness of suicide risk screening relative to depression screening alone among primary care patients has not been tested rigorously. This study compared the performance of multiple depression screening methods (Patient Health Questionnaire [PHQ]-2, PHQ-8, and PHQ-9) and multiple suicide risk screening methods (PHQ-9 item 9 and suicide-focused screening of "thoughts of killing yourself" during the entire lifespan, within the past month, and within the past week) in a convenience sample of primary care patients.

Methods: A total of 2,475 patients (military personnel, family members, and retirees) from 6 military primary care clinics completed the PHQ-9 and screening for suicidal ideation (SI) during routine clinic visits. Follow-up phone interviews were conducted for one year post-baseline to assess the incidence of suicide attempts, our primary outcome. Sensitivity, specificity, accuracy, and F1 statistics were calculated for each screening method for identifying patients who attempted suicide.

Results: More than 65% of patients who screened positive for SI also screened positive for depression on the PHQ-9. Depression screening with the PHQ-9 correctly identified more patients who attempted suicide during follow-up than the PHQ-2, past-week SI, and past-month SI. The PHQ-9 correctly identified more patients who attempted suicide within 3 months than lifetime SI, but lifetime SI correctly identified more patients who attempted suicide within 6 and 12 months.

Conclusions: Depression screening with the PHQ-9 was the most effective strategy for identifying patients who attempted suicide in the near term. Universal suicide risk screening is

unlikely to meaningfully improve identification of higher risk patients beyond PHQ-9 depression screening.

Keywords: suicide; screening; primary care

From 1999 to 2021, the US suicide rate increased by more than 35%.[1] Approximately half of suicide decedents visit a primary care or general medical provider in the month preceding their deaths.[2] Suicide decedents are more than twice as likely to visit primary care than a mental health clinic.[2] According to The Joint Commission, screening for suicidal ideation (SI) can identify patients at risk for suicide, thereby facilitating further evaluation and guiding interventions that could avert suicidal behavior.[3] The US Preventive Services Task Force (USPSTF) currently does not recommend suicide risk screening in primary care settings, owing to insufficient evidence to assess its benefits and harms.[4,5] However, the USPSTF recommends screening of primary care patients for depression when appropriate systems are in place to ensure adequate diagnosis, treatment, and follow-up.[6,7]

Although multiple depression screening tools exist, the nine-item Patient Health Questionnaire (PHQ-9)[8] and the shortened two-item version of this scale (PHQ-2)[9] are commonly used for depression screening in primary care. In some clinics, a two-stage screening approach is used wherein patients are first administered the PHQ-2 and, if positive, are subsequently administered the remaining seven items of the full PHQ-9. The PHQ-9's ninth item, which assesses the frequency of "thoughts that you would be better off dead, or thoughts of hurting yourself in some way" during the past two weeks, is often used to screen for elevated suicide risk. Empirical support for this approach comes from multiple studies showing that PHQ-

9 item 9 scores significantly predict suicide death and nonfatal suicidal behaviors.[10-12] Based on these findings, The Joint Commission recommends the PHQ-9 as a validated/evidence-based screening tool for use in primary care settings.[13]

Some researchers have critiqued the PHQ-9 as a suicide risk screening tool due to its poor sensitivity. Prior studies have found, for example, that 13.2% to 62.9% of patients who report SI on other measures of SI screen negative on PHQ-9 item 9[14-16] and 37.0% to 47.2% of patients who attempt suicide or die by suicide also screen negative.[11,17] Studies have also found that approximately 60.9% to 80.2% of patients who endorse PHQ-9 item 9 deny active SI on other measures of SI,[18,19] suggesting PHQ-9 item 9 has limited validity. PHQ-9 item 9's poor sensitivity is comparable to other SI screening tools, however.[20] Like PHQ-9 item 9, suicide risk screening tools such as the Columbia Suicide Severity Rating Scale (CSSRS)[21] miss approximately half of patients who subsequently attempt suicide or die by suicide.[22-25] Similar findings with other suicide risk screening and assessment methods[10,11,17,20,26-36] suggest poor sensitivity is not a property unique to the PHQ-9 but rather characterizes most (if not all) such tools and methods.

The very low sensitivity of SI for identifying patients who will eventually attempt suicide has been attributed to multiple factors, including the weak correlation between SI and suicidal behavior,[24] the highly dynamic and episodic nature of SI,[37,38] and inconsistencies in reporting across screening and assessment tools.[39,40] Specific to the latter point, research shows that concordance rates across SI screening and assessment tools are low when patients are administered multiple scales concurrently. In a study of patients who were administered three separate SI scales, for example, 74.3% of patients endorsed SI on at least one scale but only 28.8% endorsed SI on all three scales.[39] In a separate study of 153 primary care patients

administered three separate SI scales, 56.1% of patients endorsed SI on at least one scale but only 7.7% reported SI across all three.[40] Cases missed by PHQ-9 item 9 may therefore reflect a more general pattern of response variability across measures of SI rather than a deficiency of PHQ-9 item 9 itself.

Owing to these issues, the incremental value of expanded suicide risk screening relative to depression screening in primary care remains uncertain. The present study aims to address this knowledge gap. Our primary aim was to examine the performance of multiple screening approaches for identifying primary care patients who subsequently attempt suicide. To achieve this aim, we administered the PHQ-9 and suicide risk screening items that directly assessed SI to adult primary care patients recruited as part of the multisite PRImary care Screening Methods (PRISM) study, a one-year prospective cohort study conducted in multiple Department of Defense (DoD) primary care clinics with military personnel, family members, and reitirees.[41]

METHODS

Participants and Procedures

Participants included 2,744 adult patients (aged 18 to 89 years) recruited from the waiting rooms of six primary care clinics located at five military installations around the United States during routine visits. Interested patients spoke with a member of the research team to learn about the study. After providing consent, patients completed self-report scales on tablet or laptop. Participants were compensated with one of several possible small incentives (for example, $5 gift card to local coffee shop, T-shirt). Consistent with clinic policies and procedures, participants who endorsed thoughts of death or self-harm on the PHQ-9 were asked follow-up questions about the nature of these thoughts (for example, frequency, intensity, planning) and were referred to an integrated behavioral health provider, if indicated. Participants were

contacted 6 and 12 months later by phone to assess SI and suicidal behaviors since baseline. Participants received a $50 electronic gift card for each follow-up interview. This study was approved by the Naval Health Research Center's Institutional Review Board.

Instruments

Patient Health Questionnaire-9 (PHQ-9). All participants completed the PHQ-9,[8] a self-report scale that assesses the frequency of symptoms of major depressive disorder within the preceding two weeks using a four-point scale. Three different screening methods have been validated using PHQ-9 items: the PHQ-2[42] uses the scale's first two items to assess the frequency of depressed mood and anhedonia only, the PHQ-8[43] uses the scale's first eight items to assess the frequency of all major depressive disorder symptoms except for "thoughts that you would be better off dead or of hurting yourself in some way," and the full PHQ-9 assesses all nine symptoms of major depressive disorder, including thoughts of death and self-harm. All three scales are validated for the purposes of depression screening and monitoring.[8,9,35,42,43] The PHQ-9's ninth item assessing the frequency of "thoughts that you would be better off dead or of hurting yourself in some way" during the past two weeks is recommended by The Joint Commission for suicide risk screening in primary care,[13] with a score of one or higher indicating a positive suicide risk screen.

Suicidal Ideation (SI) Screening. All participants were administered the following item at baseline to screen for SI: "Have you ever had thoughts of killing yourself?" Participants endorsing the item were subsequently asked to report if they had experienced these thoughts within the past month or within the past week. The item mirrors the language and timeframe used in multiple validated suicide risk screening recommended by The Joint Commission, including the Ask Suicide-Screening Questions[44] (ASQ; "In the past week, have you been having thoughts about killing yourself?" and "Are you having thoughts of killing yourself right now?"),

the Patient Safety Screener 3 (PSS-3; "Over the past 2 weeks, have you had thoughts of killing yourself?"), and the Columbia Suicide Severity Rating Scale[21] (CSSRS; "Have you actually had any thoughts about killing yourself?").

Self-Injurious Thoughts and Behaviors Interview (SITBI). The SITBI[45] is a clinician-administered interview that assesses a range of self-injurious behaviors, consistent with the Centers for Disease Control and Prevention's Self-Directed Violence Classification System.[46] The SITBI was administered during follow-up via phone interview by trained graduate clinical psychology students and licensed social workers at 6 and 12 months post-baseline to assess the incidence of multiple forms of suicide attempts: aborted ("Have you been close to killing yourself and at the last minute decide not to kill yourself?"), interrupted ("Have you been very close to killing yourself and at the last minute someone or something else stopped you?"), and actual ("Have you made an actual attempt to kill yourself in which you had at least some intent to die?") suicide attempts. Patients reporting suicide attempts were asked to report the date of occurrence and describe the circumstances surrounding each behavior.

Data Analysis

All patients were administered the PHQ-9 and the SI screening item at baseline. Of 2,744 enrolled patients, follow-up data were missing from 952 (34.7%) participants due to dying of natural causes (n = 1, 0.04%), withdrawing from the study because they no longer wished to be contacted (n = 186, 6.8%), or not responding to follow-up contacts (n = 765, 27.9%). Participants with missing follow-up data were significantly younger (42.4 vs. 37.0 years, $t(1946) = 6.7$, $p < 0.001$) and were less likely to identify as white (70.2% vs. 62.5%, $\chi^2(2) = 16.4$, $p < 0.001$). Missingness was not related to PHQ-9 total score or PHQ-9 item 9 response, however. We conducted analyses using the observed data and then conducted a sensitivity analysis using the

pooled results from 10 multiply imputed datasets using an expectation-maximization algorithm.[47] The pattern of results did not differ between the observed data and multiply imputed data, however; we therefore report here the results from the observed data only.

Patients were classified as positive depression screens using three methods: a total score ≥ 3 on the PHQ-2,[9,35,42] a total score ≥ 10 on the PHQ-8,[43] and a total score ≥ 10 on the PHQ-9.[43] Patients were classified as positive suicide risk screens using four methods: item score ≥ 1 on PHQ-9 item 9, endorsement of lifetime SI, past month SI, and past week SI on the SI screening item. To examine the performance of each scale for identifying primary care patients who would subsequently attempt suicide within 1, 3, 6, and 12 months, we calculated sensitivity, specificity, accuracy, and F1 statistics. Sensitivity is the proportion of patients correctly identified by the screening method as having suicide attempts during follow-up, specificity is the proportion of patients correctly identified by the screening method as having no suicide attempts during follow-up, accuracy is the proportion of screening results that were correct (for both positive and negative screens), and F1 is the harmonic mean of sensitivity and specificity. F1 was calculated because it is useful for examining the performance of screening tools when the outcome is imbalanced.

RESULTS

The demographic and clinical characteristics of the enrolled patients are summarized in Table 1. With respect to depression screening results, at baseline 537 (19.6%) screened positive on the PHQ-2, 699 (25.5%) screened positive on the PHQ-8, and 651 (23.7%) screened positive on the PHQ-9. With respect to suicide risk screening results, at baseline 261 (9.5%) screened positive for "thoughts that you would be better off dead or of hurting yourself in some way" during the past 2 weeks on PHQ-9 item 9, 773 (28.2%) screened positive for lifetime SI, 159 (5.8%)

screened positive for past month SI, and 103 (3.8%) screened positive for past week SI. Five (0.2%) patients attempted suicide within 1 month, 16 (0.6%) attempted suicide within 3 months, 39 (1.4%) attempted suicide within 6 months, and 57 (2.1%) attempted suicide within 12 months of baseline. There were no known suicide deaths.

Overlap of Depression and Suicide Risk Screening Results at Baseline

The overlap of depression screening and suicide risk screening is depicted in Figure 1 and reported in Table 2. The number of patients screening positive for both depression and suicide risk at baseline was lowest for the PHQ-2 and highest for the PHQ-9 (Table 2, columns D and G). Specific to the PHQ-9, 87.0% of patients screening positive for suicide risk on PHQ-9 item 9, 45.2% screening positive for lifetime SI, 67.7% screening positive for past month SI, and 78.4% screening positive for past week SI also screened positive for depression on the PHQ-9.

Case Identification Performance of Depression and Suicide Risk Screening

In the short term, depression screening with the PHQ-9 correctly identified the most patients who subsequently attempted suicide (Table 2, columns F and G): 5 of 5 (100%) patients who attempted within 1 month and 10 of 16 (62.5%) patients who attempted within 3 months. Over the longer term, however, lifetime SI correctly identified the most patients who subsequently attempted suicide (Table 2, columns E and G): 23 of 39 (52.6%) patients who attempted within 6 months and 30 of 57 (62.7%) patients who attempted within 12 months. Suicide risk screening tools that assessed longer timeframes for SI correctly identified more patients: lifetime SI correctly identified more patients than past month SI and past month SI correctly identified more patients than past week SI and PHQ-item 9. Sensitivity, specificity, accuracy, and F1 statistics for each screening tool across follow-up timeframes are reported in Table 3. The PHQ-8 and

PHQ-9 had the highest sensitivity and F1 scores during the first 1 and 3 months. Lifetime SI had the highest sensitivity and F1 scores during the 6-month and 12-month timeframes.

Incremental Utility of Supplementing PHQ-9 Depression Screening with SI Screening

Of those patients who screened positive for suicide risk and then attempted suicide during follow-up (Table 2, columns E and G), more than 70% also screened positive for depression on the PHQ-9 (Table 2, column G). Suicide risk screening with PHQ-9 item 9, past week SI, and past month SI correctly identified very few additional patients who attempted suicide but were missed by the PHQ-9 (Table 2, column E): 0 additional patients who attempted within 1 month, 0 to 1 additional patients who attempted within 3 months, 1 to 3 additional patients who attempted within 6 months, and 1 to 3 additional patients who attempted suicide within 12 months. Lifetime SI performed somewhat better over the longer term, correctly identifying 0 additional patients who attempted suicide within 1 month, 2 additional patients who attempted suicide within 3 months, 6 additional patients who attempted within 6 months, and 9 additional patients who attempted within 12 months.

Limitations

Conclusions based on our findings should be made cautiously considering several limitations. First, our study examined depression as the only clinical indication for suicide risk screening. These findings may not extrapolate to other potential indications of screening in primary care such as substance use disorders, postpartum depression, and other psychiatric conditions. Additional research is needed to examine the incremental (or lack of) benefit of depression and suicide risk screening among primary care patients with these indications and other patient subgroups. Second, owing to our convenience sampling design, we were unable to assess how representative our sample was of the full patient population at participating clinics. It is possible

that the patients who volunteered for this study systematically differed from those who did not; our results therefore may not be applicable to all patients who could have participated. Third, this study was conducted within the military medical system with adult patients only, which may limit generalizability to other healthcare systems and populations. Because military selection and retention standards exclude certain psychiatric and medical conditions and military personnel tend to be younger than the US population, this sample may represent a relatively healthier population. Although this concern is mitigated somewhat by the inclusion of military family members and retirees, two beneficiary groups who are not subject to these age or medical restrictions, additional research in other populations should be conducted to determine the extent to which the present results apply more broadly. Fourth, although we screened for suicide risk using an item that mirrors several suicide-specific screening tools recommended by The Joint Commission (for example, ASQ, CSSRS, PSS-3), we did not administer the full scales, which were undergoing initial validation efforts themselves when the present study started. The similarities in language and assessment timeframes across the items used in this study and these other measures suggest our findings should have some applicability to these other scales, however. Nonetheless, further research using these scales is warranted to definitively determine incremental benefit (or lack thereof) associated with these specific measures. Fifth, follow-up data were missing from approximately one-third of participants. To assess the impact of missingness, we conducted sensitivity analyses using multiple imputation, a robust method for handling missingness. Although our results and conclusions did not change after these sensitivity analyses, caution in extending conclusions beyond this sample is nonetheless warranted. It is also possible that one or more participants with missing follow-up data died by suicide. Finally, although our sample was sufficiently large to achieve our study aims,[41] it was small enough that

relatively few patients reported suicidal behavior during follow-up, as would be expected in a primary care context. Our conclusions should therefore be considered preliminary until they can be replicated in larger studies.

CONCLUSION

In this study, we compared the effectiveness of multiple depression and suicide risk screening methods for the purpose of identifying primary care patients who subsequently attempted suicide. Our findings indicated depression screening with the PHQ-9 outperformed suicide risk screening under most conditions. In this study, depression screening with the PHQ-9 correctly identified more patients who attempted suicide in the next 1 to 3 months than every suicide risk screening method examined. The PHQ-9 also correctly identified more patients who attempted suicide over the longer term (6 to 12 months postbaseline) than screening for recent (past week to past month) SI, assessed in this study as "thoughts of killing yourself." Only lifetime SI screening correctly identified more patients than PHQ-9 depression screening over the longer term. The incremental utility of lifetime SI screening relative to PHQ-9 depression screening alone during these longer timeframes was only modest, however, for two reasons. First, more than 70% of the patients who screened positive for lifetime SI also screened positive for depression on the PHQ-9. Second, PHQ-9 depression screening correctly identified some patients who were missed by lifetime SI. To illustrate, of the 39 patients who attempted suicide within 6 months, 17 (43.6%) screened positive for both depression and lifetime SI, 6 (16.2%) screened positive for lifetime SI but not depression, and 2 (5.1%) screened positive for depression but not lifetime SI. A similar pattern was observed when extending the follow-up timeframe to 12 months. Owing to lifetime SI's own limited sensitivity, the net increase in

correct case identification associated with lifetime SI screening was limited. Lifetime SI's sensitivity was therefore only modestly higher than the PHQ-9's during the longer timeframes.

In contrast to lifetime SI screening, screening for recent (past week to past month) SI performed the relative worst, correctly identifying the fewest of primary care patients who subsequently attempted suicide: only 40.0% to 60.0% of patients who attempted within 1 month and 32.6% to 38.9% of patients who attempted within 12 months. Of the three recent SI screening methods examined, PHQ-9 item 9 performed the relative best, especially in the short term. This finding is noteworthy considering recent concerns that PHQ-9 item 9 is insufficient as a suicide risk screening method due to ambiguous and imprecise language and is not specific to SI (in other words, "thoughts that you would be better off dead or of hurting yourself in some way").[14,16] Contrary to this view, in this study PHQ-9 item 9 had only marginally lower specificity than past week SI and past month SI but had better sensitivity, suggesting the item's poor sensitivity is probably attributable to the poor sensitivity of SI more broadly, a finding that aligns with prior studies[10,11,17,26-36] rather than being a specific property of the item itself. Because assessing "thoughts of killing yourself" is a common feature of multiple suicide risk screening tools recommended by The Joint Commission (such as the ASQ, CSSRS, and PSS-3), this finding may hold important implications for suicide risk screening approaches more broadly, highlighting the need for additional prospective research in multiple healthcare settings.

Our results also suggest that supplementing the PHQ-9 with additional SI screening items did not meaningfully improve the identification of primary care patients who attempted suicide, especially in the near-term. On the contrary, our results suggest the possibility that additional SI screening—especially screening that focuses on current or recent SI—may have the unintended effect of negatively influencing clinical decision-making. In this study, for example, 13.3% to

60.0% of patients who attempted suicide after screening positive for suicide risk on PHQ-9 item 9 ("thoughts that you would be better off dead or of hurting yourself in some way") also screened negative for past week SI or past month SI ("thoughts of killing yourself"). Conflicting results between PHQ-9 item 9 and other suicide risk screening methods could create confusion for clinicians and/or prompt clinicians to underestimate the risk of many patients based on the faulty assumption that supplemental suicide risk screening methods are more effective than the PHQ-9 and PHQ-9 item 9.

The limited incremental utility of suicide risk screening in this study is also evident when considering results specific to the PHQ-8, which excludes the scale's ninth item assessing thoughts about death and self-harm. In this study, excluding the ninth item had negligible impact on case identification as compared to the full scale, correctly identifying the same number of patients within 1 and 3 months and reducing the number of correctly identified patients by only 2 within 6 and 12 months. This pattern aligns with previous research showing that nearly all the measured variance in recent SI is encompassed by depression but very little measured variance in depression is encompassed by SI.[48] Clinically, these results support the utility of depression screening with the full PHQ-9 for the purposes of identifying primary care patients who will attempt suicide in the near term and suggest that using the full PHQ-9 is superior to the PHQ-2 for this objective. Our results suggest that universal suicide risk screening is unlikely to meaningfully improve identification of higher-risk patients beyond PHQ-9 depression screening, especially in the short term, a conclusion that aligns with current USPSTF recommendations.[6,7] Because the present study was conducted in the military's health system with military beneficiaries, though, additional research using similar designs in diverse primary care settings and populations is needed to further test this possibility.

Despite the limitations, our results indicate depression screening was superior to suicide risk screening under most conditions for the specific purpose of identifying patients who would subsequently attempt suicide, with the PHQ-9 performing the best and screening for "thoughts of killing yourself" within the past week or past month performing the worst. Use of the PHQ-9 instead of the PHQ-2 could markedly improve the identification of at-risk patients in primary care, especially those who are most likely to attempt suicide in the short term. Supplementing depression screening with suicide-focused screening is unlikely to meaningfully improve case identification in primary care, however, and may have the unintended effect of redirecting clinician attention and focus away from the highest risk patients.

References

1. Xu J, Murphy SL, Kochanek KD, et al. Deaths: Final Data for 2019. Journal Issue. *National Vital Statistics Reports*. 2021;70(8):1-86.

2. Luoma JB, Martin CE, Pearson JL. Contact with mental health and primary care providers before suicide: a review of the evidence. *Am J Psychiatry*. 2002;159(6):909-916.

3. The Joint Commission. National Patient Safety Goal for Suicide Prevention. The Joint Commmission. Accessed June 6, 2022. https://www.jointcommission.org/-/media/tjc/documents/standards/r3-reports/r3_18_suicide_prevention_hap_bhc_cah_11_4_19_final1.pdf

4. Viswanathan M, Wallace I, Cooke Middleton J, et al. *Screening for Depression, Anxiety, and Suicide Risk in Children and Adolescents: An Evidence Review for the U.S. Preventive Services Task Force*. 2022.

5. LeFevre ML. Screening for suicide risk in adolescents, adults, and older adults in primary care: US Preventive Services Task Force recommendation statement. *Annals of internal medicine*. 2014;160(10):719-726.

6. Siu AL, on behalf of the USPSTF. Screening for Depression in Children and Adolescents: US Preventive Services Task Force Recommendation Statement. *Pediatrics*. 2016;137(3):e20154467.

7. US Preventive Services Task Force. Screening for Depression and Suicide Risk in Children and Adolescents: US Preventive Services Task Force Recommendation Statement. *JAMA*. 2022;328(15):1534-1542.

8. Kroenke K, Spitzer RL, Williams JB. The PHQ- 9: validity of a brief depression severity measure. *Journal of general internal medicine*. 2001;16(9):606-613.

9. Löwe B, Kroenke K, Gräfe K. Detecting and monitoring depression with a two-item questionnaire (PHQ-2). *Journal of psychosomatic research*. 2005;58(2):163-171.

10. Louzon SA, Bossarte R, McCarthy JF, et al. Does suicidal ideation as measured by the PHQ-9 predict suicide among VA patients? *Psychiatr Serv*. 2016;67(5):517-522.

11. Simon GE, Rutter CM, Peterson D, et al. Does response on the PHQ-9 Depression Questionnaire predict subsequent suicide attempt or suicide death? *Psychiatr Serv*. 2013;64(12):1195-1202.

12. Yarborough BJH, Stumbo SP, Ahmedani B, et al. Suicide behavior following PHQ-9 screening among individuals with substance use disorders. *Journal of addiction medicine*. 2021;15(1):55-60.

13. The Joint Commission. Suicide prevention resources to support Joint Commission accredited organizations implementation of NPSG 15.01.01. The Joint Commission. Accessed

May 27, 2019. https://www.jointcommission.org/-/media/tjc/documents/resources/patient-safety-topics/suicide-prevention/pages-from-suicide_prevention_compendium_5_11_20_updated-july2020_ep2.pdf

14. Horowitz LM, Mournet AM, Lanzillo E, et al. Screening pediatric medical patients for suicide risk: is depression screening enough? *Journal of Adolescent Health*. 2021;68(6):1183-1188.

15. Kemper AR, Hostutler CA, Beck K, et al. Depression and suicide-risk screening results in pediatric primary care. *Pediatrics*. 2021;148(1)

16. Mournet AM, Smith JT, Bridge JA, et al. Limitations of Screening for Depression as a Proxy for Suicide Risk in Adult Medical Inpatients. *Journal of the Academy of Consultation-Liaison Psychiatry*. 2021;62(4):413-420.

17. Bryan CJ, Allen MH, Thomsen CJ, et al. Improving suicide risk screening to identify the highest risk patients: Results from the PRImary care Screening Methods (PRISM) study. *The Annals of Family Medicine*. 2021;19(6):492-498.

18. Razykov I, Ziegelstein RC, Whooley MA, et al. The PHQ-9 versus the PHQ-8—is item 9 useful for assessing suicide risk in coronary artery disease patients? Data from the Heart and Soul Study. *Journal of psychosomatic research*. 2012;73(3):163-168.

19. Walker J, Hansen CH, Butcher I, et al. Thoughts of death and suicide reported by cancer patients who endorsed the "suicidal thoughts" item of the PHQ-9 during routine screening for depression. *Psychosomatics*. 2011;52(5):424-427.

20. Berman AL. Risk factors proximate to suicide and suicide risk assessment in the context of denied suicide ideation. *Suicide and Life- Threatening Behavior*. 2018;48(3):340-352.

21. Posner K, Brown GK, Stanley B, et al. The Columbia–Suicide Severity Rating Scale: initial validity and internal consistency findings from three multisite studies with adolescents and adults. *Am J Psychiatry*. 2011;168(12):1266-1277.

22. Simpson SA, Loh RM, Goans CR. New data on suicide risk assessment in the emergency department reveal the need for new approaches in research and clinical practice. *Psychol Med*. 2021:1-2.

23. Bjureberg J, Dahlin M, Carlborg A, et al. Columbia-Suicide Severity Rating Scale Screen Version: initial screening for suicide risk in a psychiatric emergency department. *Psychol Med*. 2021:1-9.

24. Franklin JC, Ribeiro JD, Fox KR, et al. Risk factors for suicidal thoughts and behaviors: A meta-analysis of 50 years of research. *Psychol Bull*. 2017;143(2):187.

25. Carter G, Milner A, McGill K, et al. Predicting suicidal behaviours using clinical instruments: systematic review and meta-analysis of positive predictive values for risk scales. *The British Journal of Psychiatry*. 2017;210(6):387-395.

26. Appleby L, Shaw J, Amos T, et al. Suicide within 12 months of contact with mental health services: national clinical survey. *Bmj*. 1999;318(7193):1235-1239.

27. Barraclough B, Bunch J, Nelson B, et al. A hundred cases of suicide: clinical aspects. *The British Journal of Psychiatry*. 1974;125(587):355-373.

28. Busch KA, Fawcett J, Jacobs DG. Clinical correlates of inpatient suicide. *Journal of Clinical Psychiatry*. 2003;64(1):14-19.

29. Chavan B, Singh GP, Kaur J, et al. Psychological autopsy of 101 suicide cases from northwest region of India. *Indian journal of psychiatry*. 2008;50(1):34.

30. DeLong WB, Robins E. The communication of suicidal intent prior to psychiatric hospitalization: A study of 87 patients. *Am J Psychiatry*. 1961;117(8):695-705.

31. Denneson LM, Basham C, Dickinson KC, et al. Suicide risk assessment and content of VA health care contacts before suicide completion by veterans in Oregon. *Psychiatr Serv*. 2010;61(12):1192-1197.

32. Hall RC, Platt DE, Hall RC. Suicide risk assessment: a review of risk factors for suicide in 100 patients who made severe suicide attempts: evaluation of suicide risk in a time of managed care. *Psychosomatics*. 1999;40(1):18-27.

33. Isometsä ET, Heikkinen ME, Marttunen MJ, et al. The last appointment before suicide: is suicide intent communicated? *The American journal of psychiatry*. 1995;

34. Simon GE, Coleman KJ, Rossom RC, et al. Risk of suicide attempt and suicide death following completion of the Patient Health Questionnaire depression module in community practice. *The Journal of clinical psychiatry*. 2016;77(2):20461.

35. Arroll B, Goodyear-Smith F, Crengle S, et al. Validation of PHQ-2 and PHQ-9 to screen for major depression in the primary care population. *The annals of family medicine*. 2010;8(4):348-353.

36. Smith EG, Kim HM, Ganoczy D, et al. Suicide risk assessment received prior to suicide death by Veterans Health Administration patients with a history of depression. *The Journal of clinical psychiatry*. 2013;74(3):20465.

37. Kleiman EM, Turner BJ, Fedor S, et al. Examination of real-time fluctuations in suicidal ideation and its risk factors: Results from two ecological momentary assessment studies. *Journal of abnormal psychology*. 2017;126(6):726.

38. Coppersmith DD, Ryan O, Fortgang R, et al. Mapping the Timescale of Suicidal Thinking. *Proceedings of the National Academy of Sciences*. 2022;120:e2215434120.

39. Valtonen HM, Suominen K, Sokero P, et al. How suicidal bipolar patients are depends on how suicidal ideation is defined. *J Affect Disord*. 2009;118(1-3):48-54.

40. Vuorilehto M, Valtonen H, Melartin T, et al. Method of assessment determines prevalence of suicidal ideation among patients with depression. *European Psychiatry*. 2014;29(6):338-344.

41. Bryan CJ, Allen MH, Thomsen CJ, et al. The PRImary care Screening Methods (PRISM) study: Rationale and design considerations. *Contemporary clinical trials*. 2019;84:105823.

42. Manea L, Gilbody S, Hewitt C, et al. Identifying depression with the PHQ-2: a diagnostic meta-analysis. *J Affect Disord*. 2016;203:382-395.

43. Shin C, Lee S-H, Han K-M, et al. Comparison of the usefulness of the PHQ-8 and PHQ-9 for screening for major depressive disorder: analysis of psychiatric outpatient data. *Psychiatry investigation*. 2019;16(4):300.

44. Horowitz LM, Snyder DJ, Boudreaux ED, et al. Validation of the Ask Suicide-Screening Questions (ASQ) for adult medical inpatients: a brief tool for all ages. *Psychosomatics*. 2020;

45. Nock MK, Holmberg EB, Photos VI, et al. Self-Injurious Thoughts and Behaviors Interview: development, reliability, and validity in an adolescent sample. 2007;

46. Crosby A, Ortega L, Melanson C. Self-directed violence surveillance; uniform definitions and recommended data elements. 2011;

47. MOLENBERGHS GVG, Verbeke G. Longitudinal and incomplete clinical studies. *Metron*. 2005;63(2):143-176.

48. Rogers ML, Stanley IH, Hom MA, et al. Conceptual and Empirical Scrutiny of Covarying Depression Out of Suicidal Ideation. *Assessment*. 2018;25(2):159-172.

Table 1: Sample Characteristics (N = 2744)

Sample Size	**n (%)**	**Mean (SD)**
Age in years	--	40.4 (19.6)
Gender		
Male	1380 (51.3)	--
Female	1279 (47.5)	--
Other	9 (0.3)	--
Prefer not to Answer	17 (0.6)	--
Unknown / Missing	59 (2.2)	--
Race		
White / Caucasian	1811 (67.3)	--
Black / African American	506 (18.8)	--
Asian	115 (4.3)	--
Native American / Alaska	123 (4.6)	--
Native	44 (1.6)	--
Pacific Islander / Native	272 (10.1)	--
Hawaiian		
Other		
Hispanic / Latino Ethnicity		

Yes	415 (15.4)	--
No	2199 (81.7)	--
Other	20 (0.7)	--
Prefer not to Answer	51 (1.9)	--
Military Service		
Yes, current member	1652 (61.4)	--
Yes, in the past	451 (16.8)	--
No	580 (21.6)	--
Branch of Service		
Air Force	236 (11.2)	--
Army	408 (19.4)	--
Coast Guard	3 (0.1)	--
Marines	283 (13.5)	--
Navy	1171 (55.7)	--

SD, standard deviation.

Table 2: Relationships Among Positive Screens for Depression and Positive Screens for Suicide Risk Among Primary Care Patients Who Attempted Suicide Within 1, 3, 6, and 12 Months of Screening

	1 Month								**3 Months**						
Screening Tool	**A**	**B**	**C**	**D**	**E**	**F**	**G**		**A**	**B**	**C**	**D**	**E**	**F**	**G**
PHQ-2															

PHQ-9 item 9	176	46	1	72	1	1	2		175	46	5	68	1	2	6
Lifetime SI	129	342	0	119	2	1	2		128	340	2	115	4	2	6
Past month SI	190	48	1	58	1	2	1		188	47	4	55	2	4	4
Past week SI	209	22	1	39	1	2	1		207	22	5	36	1	4	4
PHQ-8															
PHQ-9 item 9	251	26	0	92	0	2	3		250	26	4	88	0	3	7
Lifetime SI	172	290	0	171	0	1	4		171	288	2	169	2	2	8
Past month SI	277	40	0	66	0	3	2		275	39	3	63	1	5	5
Past week SI	298	16	0	45	0	3	2		296	16	4	42	0	5	5
PHQ-9															
PHQ-9 item 9	251	18	0	100	0	2	3		250	18	4	96	0	3	7
Lifetime SI	177	287	0	174	0	1	4		176	285	2	170	2	2	8
Past month	284	39	0	67	0	3	2		28	38	3	64	1	5	5

SI									2						
Past week SI	306	16	0	45	0	3	2		304	16	4	42	0	5	5
	6 Months								**12 Months**						
PHQ-2															
PHQ-9 item 9	188	62	16	83	2	3	17		170	46	23	61	1	7	13
Lifetime SI	137	372	6	146	12	2	18		126	330	10	105	14	4	16
Past month SI	196	46	12	51	4	5	11		183	45	20	48	4	9	11
Past week SI	218	25	14	36	3	5	10		202	21	22	29	2	9	11
PHQ-8															
PHQ-9 item 9	273	33	13	112	0	6	19		242	26	19	81	0	11	14
Lifetime SI	180	311	6	207	7	2	23		169	281	10	154	9	4	21
Past month SI	282	37	9	60	3	8	12		267	37	16	56	3	13	12
Past week SI	312	18	11	43	1	8	12		288	15	18	35	1	13	12
PHQ-9															

PHQ-9 item 9	273	25	13	120	0	6	19		242	18	19	89	0	11	14
Lifetime SI	175	308	6	204	7	2	23		174	278	10	157	9	4	21
Past month SI	279	36	9	60	3	8	12		274	36	16	57	3	13	12
Past week SI	310	18	11	42	1	8	12		296	15	18	35	1	13	12

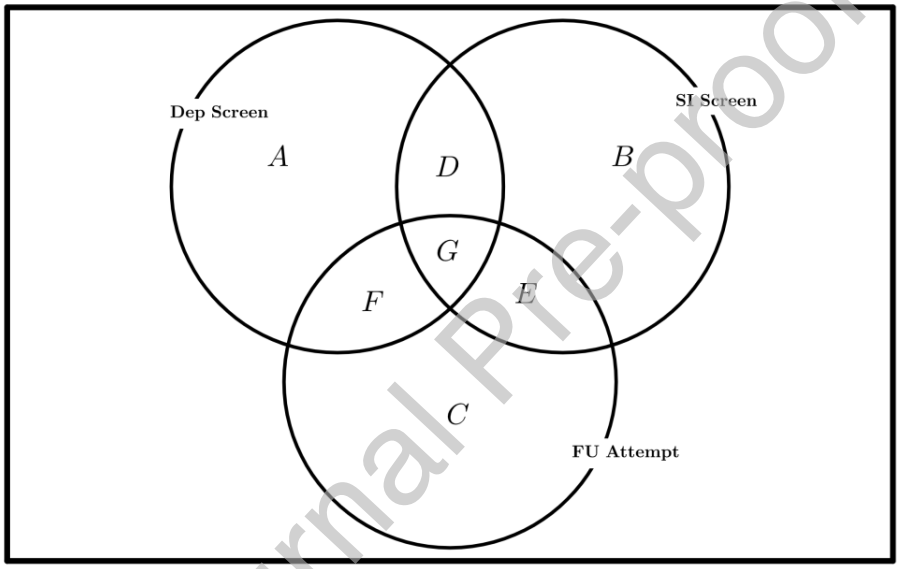
Note: Letters in column headers correspond to the following segments from Figure 1: A, patients without suicide attempts who screened positive for depression and negative for suicide risk; B, patients without suicide attempts who screened negative for depression and positive for suicide risk; C, patients with suicide attempts who screened negative for both depression and suicide risk; D, patients without suicide attempts who screened positive for both depression and suicide risk; E, patients with suicide attempts who screened negative for depression and positive for suicide risk; F, patients with suicide attempts who screened positive for depression and negative for suicide risk; G, patients with suicide attempts who screened positive for both depression and suicide risk. PHQ, Patient Health Questionnaire; SI, suicidal ideation.

Table 3: Depression and Suicide Risk Screening Performance for the Identification of Primary Care Patients Who Attempted Suicide Within 1, 3, 6, and 12 Months of Screening

Screening Tool	SE	SP	ACC	F1		SE	SP	ACC	F1		SE	SP	ACC	F1		SE	SP	ACC	F1
Depression																			
PHQ-2	0.60	0.81	0.81	0.69		0.63	0.81	0.81	0.71		0.54	0.82	0.80	0.65		0.51	0.82	0.81	0.63
PHQ-8	1.00	0.74	0.74	0.85		0.75	0.75	0.75	0.75		0.67	0.76	0.75	0.71		0.61	0.75	0.75	0.68
PHQ-9	1.00	0.76	0.76	0.86		0.73	0.76	0.76	0.75		0.66	0.77	0.77	0.71		0.59	0.77	0.76	0.67
Suicide risk																			
PHQ-9 item 9	0.60	0.91	0.91	0.72		0.53	0.91	0.91	0.67		0.50	0.91	0.90	0.65		0.39	0.92	0.90	0.55
Lifetime SI	0.80	0.68	0.68	0.74		0.73	0.68	0.68	0.71		0.79	0.69	0.68	0.74		0.73	0.69	0.69	0.71
Past month SI	0.40	0.93	0.93	0.56		0.43	0.93	0.93	0.59		0.47	0.94	0.93	0.63		0.36	0.94	0.92	0.52
Past week SI	0.40	0.96	0.95	0.56		0.36	0.96	0.95	0.52		0.41	0.96	0.95	0.57		0.33	0.96	0.95	0.49

SE, sensitivity (the proportion of patients correctly identified by the screening method as having suicide attempts during follow-up); SP, specificity (the proportion of patients correctly identified by the screening method as having no suicide attempts during follow-up); ACC, accuracy (the proportion of screening results that were correct); F1, F1 statistic (the harmonic mean of sensitivity and specificity); PHQ, Patient Health Questionnaire; SI, suicidal ideation

Figure 1: Illustrated here are relationships among positive screens for depression, positive screens for suicide risk, and follow-up suicide attempts among 2,475 primary care patients. Patients who screened positive for depression at baseline are represented by the top left circle labeled "Dep Screen" (segments A, D, F, and G), patients who screened positive for suicide risk at baseline are represented by the top right circle labeled "SI Screen" (segments B, D, E, and G), and patients who attempted suicide during follow-up are represented by the bottom center circle labeled "FU Attempts" (segments C, E, F, and G).



Declaration of interests

☐ The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☒The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

All authors report grant funding from the Department of Defense. CJ Bryan, AO Bryan, and JC Baker report income for paid workshops and consultant fees related to suicide prevention through Anduril LLC.