The WHO-5 Wellbeing Index performed the best in screening for depression in primary care

Henkel V, Mergl R, Kohnen R, et al. **Identifying depression in primary care: a comparison of different methods in a prospective cohort study.** BMJ. 2003;326:200-1.

QUESTION

In the setting of primary care, what is the accuracy of screening questionnaires in identifying depression compared with the accuracy of clinical diagnosis without the aid of questionnaires?

DESIGN

Blinded comparison of unaided clinical diagnosis and 3 screening questionnaires with a standardized psychiatric interview.

SETTING

18 primary care practices in Germany.

PATIENTS

431 patients {18 to 88 years of age (mean age 52.6 y, 63% women)}* who were attending the practices on 1 given day and agreed to complete 3 screening questionnaires.

DESCRIPTION OF TESTS AND DIAGNOSTIC STANDARD

Before being seen by a physician, patients completed 3 screening questionnaires: the depression module of the Brief Patient Health Questionnaire (PHQ-9), the General Health Questionnaire, and the World Health Organization (WHO) Wellbeing Index (WHO-5). Physicians treating the patients completed an encounter form to record a clinical assessment, blinded to the screening results. Within 6 days of their visit, patients were contacted via telephone by a psychologist, who was blinded to the screening results and physician assessment, for a standardized psychiatric interview using the Composite International Diagnostic Interview.

MAIN OUTCOME MEASURES

Sensitivity, specificity, and predictive values of the 3 depression screening tools and physicians' unaided clinical diagnosis.

MAIN RESULTS

The diagnostic performances of the tests are in the Table. The WHO-5 had the greatest sensitivity and negative predictive value of the tests and clinical assessment.

CONCLUSIONS

In primary care, the World Health Organization Wellbeing Index (WHO-5) performed better than 2 other questionnaires and unaided clinical diagnosis as a depression screening tool. More cases of depression could be identified by using the WHO-5.

For correspondence: Dr. V. Henkel, Ludwig-Maximilians-University Munich, Munich, Germany. E-mail verena.henkel@psy.med.uni-muenchen.de. ∎

*Information provided by author.

Diagnostic characteristics of 3 screening questionnaires and unaided clinical diagnosis for detecting depression†

Tests	Sensitivity (95% CI)	Specificity (CI)	PPV	NPV	+LR	−LR
WHO-5	93% (85 to 98)	64% (59 to 69)	34%	98%	2.58	0.11
GHQ-12	85% (74 to 92)	62% (57 to 67)	31%	95%	2.24	0.24
PHQ-9	78% (66 to 87)	85% (81 to 89)	51%	95%	5.20	0.26
Clinical diagnosis	65% (53 to 76)	74% (69 to 79)	34%	91%	2.50	0.47

†WHO-5 = World Health Organization Wellbeing Index; GHQ-12 = General Health Questionnaire; PHQ-9 = Brief Patient Health Questionnaire, 9 items; PPV = positive predictive value; NPV = negative predictive value. Diagnostic terms defined in Glossary; LRs calculated from data in article.

COMMENTARY

In May 2002, the U.S. Preventive Services Task Force recommended screening all adults for depression. In this study, Henkel and colleagues provide a timely and useful model for evaluating different screening measures. They find that unaided primary care providers detect depression with only 65% sensitivity, emphasizing the need for depression screening tests.

Of the 3 tests evaluated, it is not surprising that the WHO-5 had the best sensitivity and negative predictive value, since it is perhaps the broadest of the measures (1). Participants are asked to agree or disagree with such statements as "I feel calm and relaxed." Such general statements improve sensitivity and negative predictive value at the cost of specificity and positive predictive value.

However, others still prefer the PHQ-9 as an overall screening test. It has been found to have sensitivity and specificity as high as 88% each and, despite its brevity, to be useful in grading the severity of depression (2).

Other screening tests not used in this trial merit consideration for further study. One clinically useful measure is the BDI-PC (Beck Depression Inventory for Primary Care) (3). It consists of 7 questions and can be completed in a few minutes. Since brevity is indeed important in the primary care arena, the 2-question PRIME-MD (Primary Care Evaluation of Mental Disorders) depression screen, which has been shown to be useful when combined with a 4-question follow-up screen, should also be further evaluated (4).

Finally, it is essential to compare the utility of the various depression screening measures in different populations, because patients of different ages, sex, and cultural backgrounds respond differently to such screens (5).

Brian A. Primack, MD, EdM
University of Pittsburgh School of Medicine
Pittsburgh, Pennsylvania, USA

References

1. **Bonsignore M, Barkow K, Jessen F, Heun R.** Validity of the five-item WHO Well-Being Index (WHO-5) in an elderly population. Eur Arch Psychiatry Clin Neurosci. 2001;251 Suppl 2:II27-31.
2. **Kroenke K, Spitzer RL, Williams JB.** The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med. 2001;16:606-13.
3. **Sharp LK, Lipsky MS.** Screening for depression across the lifespan: a review of measures for use in primary care settings. Am Fam Physician. 2002;66:1001-8.
4. **Brody DS, Hahn SR, Spitzer RL, et al.** Identifying patients with depression in the primary care setting: a more efficient method. Arch Intern Med. 1998;158:2469-75.
5. **Kerr LK, Kerr LD, Jr.** Screening tools for depression in primary care: the effects of culture, gender, and somatic symptoms on the detection of depression. West J Med. 2001;175:349-52.