Screening for depression in primary care with two verbally asked questions: cross sectional study

Bruce Arroll, Natalie Khin, Ngaire Kerse

Editorial by Del Mar and Glasziou

Department of General Practice and Primary Health Care, Faculty of Medical and Health Sciences, University of Auckland, PB 92019, Auckland, New Zealand

Bruce Arroll
associate professor

Ngaire Kerse
senior lecturer

Department of Psychiatry, University of Auckland

Natalie Khin
PhD student

Correspondence to: B Arroll b.arroll@auckland.ac.nz

BMJ 2003;327:1144–6

Abstract

Objective To determine the diagnostic accuracy of two verbally asked questions for screening for depression.
Design Cross sectional criterion standard validation study.
Setting 15 general practices in New Zealand.
Participants 421 consecutive patients not taking psychotropic drugs.
Main outcome measures Sensitivity, specificity, and likelihood ratios of the two questions compared with the computerised composite international diagnostic interview.
Results The two screening questions showed a sensitivity and specificity of 97% (95% confidence interval, 83% to 99%) and 67% (62% to 72%), respectively. The likelihood ratio for a positive test was 2.9 (2.5 to 3.4) and the likelihood ratio for a negative test was 0.05 (0.01 to 0.35). Overall, 37% (157/421) of the patients screened positive for depression.
Conclusion Two verbally asked questions for screening for depression would detect most cases of depression in general practice. The questions have the advantage of brevity. As treatment is more likely when doctors make the diagnosis, these questions may have even greater utility.

Introduction

Depression is a common and costly mental health problem seen often in general practice and general medicine.[1] In 2002 the US Preventive Services Task Force endorsed screening for depression but did not recommend a specific screening tool.[2] A systematic review found that screening for depression was not effective in improving psychosocial outcomes.[3] The US Preventive Services Task Force claims that its review is more extensive.

Many practitioners find the numerous case finding and screening questionnaires for depression too cumbersome and time consuming for routine use.[4] A feasible screening tool for use in general practice would comprise one or two questions, which, if positive, could be followed by further questions from the depression criteria. The primary care evaluation of mental disorders, designed to facilitate the diagnosis of common mental disorders in general practice, involved a screening questionnaire with 27 items and a follow up interview with a clinician.[5] The questionnaire included two questions about depressed mood: during the past month have you often been bothered by feeling down, depressed, or hopeless? and, during the past month have you often been bothered by little interest or pleasure in doing things? One study of these questions reported a sensitivity of 96% and a specificity of 57% compared with the quick diagnostic interview schedule.[6] We aimed to evaluate the questions when asked verbally, instead of in the written form, by general practitioners in the community.[5 6]

Participants and methods

From a database of Auckland general practices we randomly selected 15 general practices. Each general practitioner asked the two questions at any time during a consultation, and if either was positive, screening was considered positive. The general practitioners had access to the usual patient notes. They completed a form of the patient's responses and whether or not safety issues, such as suicidal thoughts, had been addressed. The study interviewer looked at the form after the patient had completed the mood module of the computer assisted composite international diagnostic interview.[7–9] Patients had no opportunity to start treatment before completing the composite interview. This interview takes the participant's answers, provided without any interpretation, probe, or explanation by the interviewer, as valid data for arriving at a diagnosis. It has been evaluated for test-retest reliability and compared with the schedules for clinical assessment in neuropsychiatry.[8 9]

The calculator on the University of Toronto website was used to determine the sensitivity, specificity, and likelihood ratios.[10–12] Our study was designed and analysed as recommended by the Standards for Reporting Diagnostic Accuracy Steering Group.[13]

Results

Overall, 670 consecutive patients were invited by their general practitioners to participate in our study. Of these, 476 took part (response rate 71.0%): 142 men, 330 women, and four had missing data (figure). The median age was 46 (range 16 to 90). We excluded 47 patients who were taking psychotropic drugs, 194

Table 1 Validity and positive predictive value for screening questions and physician diagnosis compared with composite international diagnostic interview as ideal screening tool for major depression

Screening question	Patients screened positive		Patients screened negative		Positive predictive value (%)
	True positive	False positive	True negative	False negative	
Both questions*	28	129	263	1	18
Depression question	25	111	281	4	18
Pleasure question	24	84	308	5	22

*Positive is yes to either question.

Table 2 Sensitivity, specificity, and likelihood ratios with composite international diagnostic interview as ideal screening tool for major depression

Screening question	Sensitivity % (95% CI)	Specificity % (95% CI)	Likelihood ratio	
			Positive test (95% CI)	Negative test (95% CI)
Both questions	97 (83 to 99)	67 (62 to 72)	2.9 (2.5 to 3.4)	0.05 (0.01 to 0.35)
Depression only question	86 (69 to 95)	72 (67 to 76)	3.0 (2.5 to 3.8)	0.19 (0.08 to 0.48)
Pleasure only question	83 (66 to 92)	79 (74 to 82)	3.9 (3.0 to 5.0)	0.22 (0.1 to 0.49)

declined, and eight were not asked the screening question. In total, 421 patients were asked the two screening questions. According to the composite interview, 28 of the 157 (18%) who screened positive were depressed, whereas only one of the 264 who screened negative was depressed.

Table 1 shows the raw data for both questions and each question and the positive predictive value when using the composite interview as the ideal screening tool. Table 2 shows the sensitivity, specificity, and likelihood ratios for both questions and the questions separately. A yes to either question was considered a positive response. The questions showed a sensitivity of 97% (95% confidence interval 83% to 99%) and a specificity of 67% (62% to 72%). The high sensitivity was accompanied by a high number of false positive results. This is reflected in the modest likelihood ratio for a positive test and the positive predictive value of 18%. On the other hand, the likelihood ratio for a negative test was low, and at the prevalence of 6% for major depression a negative test would almost always be a true negative (negative predictive value 99%).

What is already known on this topic

Screening for depression in general practice is effective at diagnosing depression and optimising treatment

Screening tests are usually in written form

What this study adds

Two questions verbally asked are potentially useful for screening for depression owing to reasonable validity and brevity

A reasonable trade-off exists between true and false positives

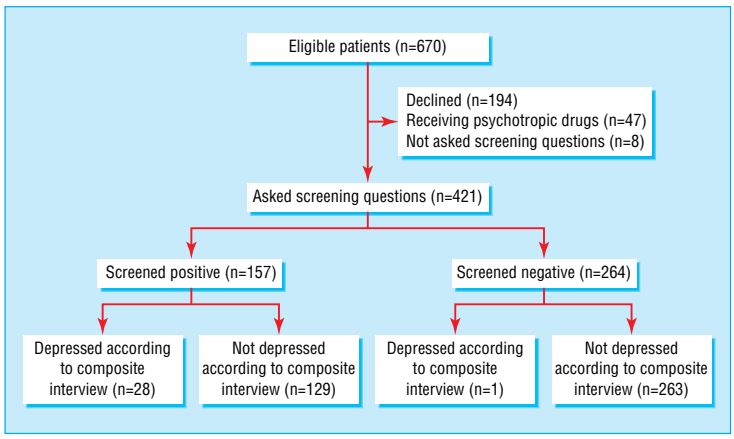
The questions detect most cases of depression

Discussion

Two verbally asked questions from the original primary care evaluation of mental disorders have good sensitivity and reasonable specificity for screening for depression. The 97% sensitivity we found is an improvement over the 29% to 35% often reported.[14] The post-test probabilities suggest about five false positives for every true positive when asking the questions alone. This is common in screening studies, which are in essence a diagnostic test performed in a "low prevalence" setting. This is not a major concern with depression, as further clarification can be obtained by asking more questions (the reference standard) or referral to another health professional.

Our study was conducted in a community setting by general practitioners and analysed after exclusion of patients taking psychotropic drugs. It is the first assessment of the questions administered verbally rather than in written form. A weakness of our study is that there was no non-screened group as a comparator.

The prevalence for screening studies for depression in general practice is usually low (8% for major depression); hence the likelihood ratio for a negative test does not need to be low to rule out depression when the test is negative (in this sample a patient with a negative test would have a 0.3% chance of being depressed). Also, when compared with the 41 studies evaluated by the US Preventive Services Task Force, the two questions (verbally asked) had a similar likelihood ratio for a positive test compared with most studies in that review.[15] The two questions were, however, considerably shorter than the shortest (seven questions)



Flow of participants through trial

screening questionnaire.[16] They are thus a good compromise between the time required to administer the screen and the likelihood ratio. The additional benefit is that general practitioners are more likely to prescribe drugs to patients in whom they have made the diagnosis.[17]

Contributors: All authors wrote the paper. AB had the original idea for the study and analysed the data; he will act as guarantor for the paper. NKhin assisted with the study design and funding. NKerse analysed the data. S Brighouse interviewed the patients.

Funding: Oakley Mental Health Foundation and Charitable Trust of the Auckland Faculty of the Royal New Zealand College of General Practitioners. The guarantor accepts full responsibility for the conduct of the study, had access to the data, and controlled the decision to publish.

Competing interests: None declared.

Ethical approval: Ethical approval was obtained from the Auckland ethics committees.

1 Katon W, Schulberg H. Epidemiology of depression in primary care. *Gen Hosp Psychiatry* 1992;14:237-47.
2 Pignone MP, Gaynes BN, Rushton JL, Burchell CM, Orleans CT, Mulrow CD, et al. Screening for depression in adults: a summary of the evidence for the US preventive services task force. *Ann Intern Med* 2002;136: 765-76.
3 Gilbody SM, House A, Shledon TA. Routinely administered questionnaires for depression and anxiety: a systematic review. *BMJ* 2001;322:406-9.
4 Andersen SM, Harthorn BH. The recognition, diagnosis, and treatment of mental disorders by primary care physicians. *Med Care* 1989;27: 869-86.
5 Spitzer RL, Williams JB, Kroenke K, Linzer M, deGruy III FV, Hahn SR, et al. Utility of a new procedure for diagnosing mental disorders in primary care. The prime-MD1000 study. *JAMA* 1994;14:1749-56.
6 Whooley MA, Avins AL, Miranda J, Browner WS. Case finding instruments for depression: two questions as good as many. *J Gen Intern Med* 1997;12:439-45.
7 World Health Organization. *Composite international diagnostic interview (CIDI).* Geneva: WHO.
8 Wittchen HU, Lachner G, Wunderlich U, Pfister H. Test-retest reliability of the computerized DSM-IV version of the Munich-composite international diagnostic interview. *Soc Psychiatry Psychiatr Epidemiol* 1998;33:568-78.
9 Andrews G, Peters L, Guzman AM, Bird K. A comparison of two structured diagnostic interviews: CIDI and SCAN. *Aust NZ J Psychiatry* 1995;29:124-32.
10 Centre for Evidence-Based medicine, Mount Sinai Hospital. www.cebm.utoronto.ca (accessed Dec 2002).
11 Jaeschke R, Guyatt G, Sackett D. Users' guides to the medical literature III. How to use an article about a diagnostic test. A. Are the results of the study valid? *JAMA* 1994;271:389-91.
12 Jaeschke R, Guyatt G, Sackett D. Users' guides to the medical literature III. How to use an article about a diagnostic test. B. What were the results and will they help me in caring for my patients? *JAMA* 1994;271:703-7.
13 Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig LM, et al. Towards complete and accurate reporting of studies of diagnostic accuracy: the STARD initiative. *BMJ* 2003;326:41-4.
14 Nease DE, Malouin JM. Depression screening: a practical strategy. *J Fam Pract* 2003;52:118-26.
15 US Preventive Services Task Force. www.ahrq.gov/clinic/uspstfix.htm (accessed May, 2003).
16 Steer RA, Cavalieri TA, Leonard DM, Beck AT. Use of the Beck depression inventory for primary care to screen for major depression disorders. *Gen Hosp Psychiatry* 1999;21:106-11.
17 Dowrick C. Does testing for depression influence diagnosis or management by general practitioners. *Fam Pract* 1995;12:461-5.

(Accepted 18 September 2003)

Curiouser and curiouser

Last week we buried my father in law, David. He was 92 and had had a full life. We put him to rest in a natural woodland site, and it was a joyful day.

His three sons each spoke beside his coffin. Curiosity was a strong theme. He had a never ending curiosity for how things worked—engines, boats, paints, clays, instruments. For about the past 20 years he had professed not to be able to see or hear, yet he skillfully made and mended clocks. Curiosity, apparently, overcame his disabilities. (Selective deafness was also a hypothesis.) He accepted what he could do and not do and went ahead and did what he could with great tenacity and perseverance.

He had a healthy irreverence, a naughtiness, which his grandchildren loved. Adults also benefited from it. A retired hill farmer friend who came to the funeral remembered how David had wanted to build a still on his land. "But that's illegal." "So?" He had wanted to know how it would work.

People's curiosities vary. This year's Reith lecturer, the neuroscientist Vilinor Ramachandran, speaking on "The emerging mind" showed a passion, not for the wiring of boat engines but for the wiring of the human brain. An intense curiosity about what goes wrong with the wiring to produce sensory anomalies, such as synaesthesia, drives him on to know more and more and to inspire others to know more and more.

Curiosity about people and how they work—physically, mentally, emotionally, spiritually—is at the heart of what doctors do. It drives us. Yes, other "C" words—communication, collaboration, consultation, computers even—are also important. But in a culture increasingly oriented towards data collection and management, this people curiosity is precious. There are nine general practitioners in my practice, and we spend a lot of time discussing chronic disease management protocols and care management screens. Nine doctors, nine views. The optimistic view is that all the energy and time diverted into this will create efficient systems that will liberate time for being curious about people, for being a family doctor. The less optimistic view is that, once the climate has been changed to the extent that it has, cultivating the curiosity strain of the species may become more difficult. Yet this is the strain the consumers want and the health service needs, the one most likely to produce good crops (of GPs) in the future. (Gardening was another of David's hobbies.)

Feeding the fire of curiosity, about engines and clocks and how machines work, gave my father in law a long and happy life. Sorry, another horticultural analogy: the healthy growth of general practice depends on keeping the balance of the soil right, on feeding the curiosity. And perhaps the naughtiness?

Lesley Morrison *general practitioner principal, Teviot Medical Practice, Hawick*

We welcome articles up to 600 words on topics such as *A memorable patient, A paper that changed my practice, My most unfortunate mistake,* or any other piece conveying instruction, pathos, or humour. Please submit the article on http://sdubmit.bmj.com Permission is needed from the patient or a relative if an identifiable patient is referred to. We also welcome contributions for "Endpieces," consisting of quotations of up to 80 words (but most are considerably shorter) from any source, ancient or modern, which have appealed to the reader.