



## Evaluation of studies of assessment and screening tools, and diagnostic tests

Andrew Jull

*Evid. Based Nurs.* 2002;5;68-72  
doi:10.1136/ebn.5.3.68

---

Updated information and services can be found at:  
<http://ebn.bmj.com/cgi/content/full/5/3/68>

---

*These include:*

### References

This article cites 14 articles, 3 of which can be accessed free at:  
<http://ebn.bmj.com/cgi/content/full/5/3/68#BIBL>

### Rapid responses

You can respond to this article at:  
<http://ebn.bmj.com/cgi/eletter-submit/5/3/68>

### Email alerting service

Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article

---

### Notes

---

To order reprints of this article go to:  
<http://journals.bmj.com/cgi/reprintform>

To subscribe to *Evidence-Based Nursing* go to:  
<http://journals.bmj.com/subscriptions/>

## EBN users' guide

# Evaluation of studies of assessment and screening tools, and diagnostic tests

In nursing, diagnostic testing is commonly a regulated activity performed by nurse practitioners and midwives. However, assessment and screening of patients for further testing (case finding) are central elements of nursing. The number of studies evaluating assessments and tests is increasing, but overall, the methodological quality of these studies has been poor.<sup>1</sup> Thus, nurses should be able to critically appraise evidence from such studies to ensure that the highest quality assessment and screening tools are used. The tools of assessment, case finding (or screening), and diagnosis are evaluated using different criteria from those applied to studies investigating preventive or therapeutic interventions, although the 3 basic questions of critical appraisal are the same: are the results valid? What are the results? Will the results help me in caring for my patients? In this article, we outline a framework by Sackett *et al*<sup>2</sup> to critique studies that evaluate a screening tool to assess patients for depression. The same framework can also be applied to studies of assessment tools such as fall risk assessments or pressure sore risk scoring, as well as studies of diagnostic tests.

### Clinical scenario

You are a district nurse attending a 68 year old man with a diabetic ulcer in his home. He feels that his ulcer is taking forever to heal and that he will never be well again. You know from previous conversations that his wife died several years ago and his 2 children live outside of the area or overseas. When he could drive, he was socially active, but since becoming reliant on others to assist him, he doesn't get out much now. He reports he is eating okay, and that his glucose concentrations are kept within normal range with his hypoglycaemic medications. He also says he feels tired all the time. You have noticed he is not taking as much care with his appearance recently and that he seems much less interested in world events than when you first started dressing his ulcer. Although you are aware that malaise is a side effect of some hypoglycaemics, his drug regimen has not been changed recently. On further questioning, it seems unlikely your patient is anaemic, so you begin to consider whether he might be depressed. You wonder if there are simple assessments that could help you to screen patients for appropriate referral.

### The search

Unlike studies of preventive or therapeutic interventions, which are best answered by randomised controlled trials or systematic reviews of such trials, questions about the effectiveness of screening and assessment tools in clinical practice are best answered by cross sectional studies. Studies of diagnostic tests and screening tools can be quite difficult to locate. The keyword phrase *sensitivity and specificity* is useful, but studies cannot always be located with that phrase. If that is the case, then the subheading /di (for diagnosis) can be used. Specialist databases such as the *Cochrane Library* do not include studies of diagnostic testing, so you do the following search on Medline (OVID):

- (1) \*Depression/di
- (2) \*Depressive disorder/di
- (3) Questionnaires/
- (4) (1 OR 2) AND 3

303 articles are identified. Adding the text words primary care, and limiting the search to English language papers with abstracts published in the past 5 years produces a more manageable 24 citations, one of which is a study of screening tools, including a simple 2 question instrument ("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?").<sup>3</sup> Whooley *et al* tested 7 screening tools on 542 consecutive patients attending an urgent care clinic. 97% of participants were men, the majority of whom were unemployed. Prevalence of depression was 18%. All screening tools returned similar results, but the investigators recommended the 2 question instrument for use in primary care because of its simplicity. Before accepting the authors' conclusions, readers need to assure themselves that the study is valid. This requires answering 4 questions.

### Are the results of the study valid?

WAS THERE AN INDEPENDENT BLIND COMPARISON WITH A REFERENCE (GOLD) STANDARD OF DIAGNOSIS?

There are 2 aspects to this question. Firstly, the accuracy of any tool is best determined by comparing its results with those obtained from a widely accepted reference test. This is also referred to as a *gold standard* test and is often more invasive than the initial test. Thus palpation of a child's forehead for fever could be compared with a reading from a mercury thermometer to obtain a true estimation as to whether a child has a major fever. Similarly, the ankle-brachial pressure index (ABPI) could be compared with the gold standard of venography if testing ABPI as a screen for arterial disease in the leg. Readers need to be assured that the reference test is the gold standard test for the condition. Comparing palpation with a measure less accurate than a mercury thermometer (such as tympanic thermometry) would provide an inaccurate estimation of how many patients actually had a fever.<sup>4</sup> Even if the study used mercury thermometry or a device with similar accuracy, the reader must still be reassured that an acceptable technique of thermometry was applied. Thus, even if mercury thermometry was used, an axillary route is likely to provide an unreliable estimate of temperature. If > 1 assessor was used, the study should also provide an estimate of the level of agreement between assessors.

The second aspect of the question guards against *expectation bias*. Although in most clinical situations, healthcare workers have access to patient records, it is important, when evaluating an instrument, that clinicians form their own determinations of the patient's condition. Prior knowledge of the presence or absence of a disorder could influence a clinician's assessment. Therefore, it is imperative that clinicians making an assessment using the gold standard test are separate from those using the

other instrument, and that the 2 groups of clinicians are blinded to each other's assessments. A methodological study of evaluations of diagnostic tests has found that unblinded assessments overestimate correct diagnoses by as much as 30% compared with blinded studies.<sup>5</sup>

#### WAS THE DIAGNOSTIC TEST EVALUATED IN AN APPROPRIATE SPECTRUM OF PATIENTS (LIKE THOSE WE WOULD MEET IN CLINICAL PRACTICE)?

The main challenge when evaluating a case finding or screening instrument is to apply it to the indicated population.<sup>6</sup> Tests are often developed by the quick and dirty method of using an accessible population of patients known to have the target disorder and a group of healthy controls. If an instrument does not discriminate between those with and without the disorder at this stage of development, then it is unlikely to be clinically useful. But, the value of an assessment lies in its ability to distinguish the full spectrum of presenting patients with the disorder (as well as those who present with similar symptoms arising from different disorders) from those who do not have the condition. Diagnostically, it is easier to identify patients with florid presentation from those without the condition than it is to identify those with a mild presentation. Only if the instrument can differentiate those likely to have the disorder from those who do not in a real clinical population can it then be deemed useful. Evaluations of new tests often omit the essential developmental stage of evaluation in a real clinical population. For example, one use of ABPI is assessing patients with leg ulcers to screen for those with peripheral arterial disease, which would rule out treatment with high compression bandaging. In the late 1960s the normal values of an ABPI were established by testing 110 patients with known occlusive peripheral arterial disease and comparing their test values with those of 25 healthy controls.<sup>7</sup> It is only recently that the utility of the ABPI has been tested in community populations similar to those in which it is commonly used by district nurses.<sup>8</sup>

It generally accepted that it is good practice for studies to enrol consecutive patients who have agreed to participate (minimising the potential for selection bias), although non-consecutive enrolment has not been found to have any significant effect on study results.<sup>9</sup>

#### WAS THE REFERENCE STANDARD APPLIED REGARDLESS OF THE TEST RESULTS?

To avoid *verification* or *workup* bias, participants need to receive both tests regardless of the outcome of the first test. If the first test is negative and the participant does not receive the gold standard test to verify this result, then the study results will be distorted. In some instances, participants who have had a negative test may decide not to have the gold standard test, especially if the gold standard test is an invasive procedure such as venography. Rather than exclude these participants, investigators can follow them up over an appropriate time period and monitor for symptoms of the target disorder.

#### WAS THE TEST VALIDATED IN A SECOND, INDEPENDENT GROUP OF PATIENTS?

For a reader to be reassured that the study findings are accurate and not the result of idiosyncrasies in the initial cohort of participants or the individual skills of the assessors, the tool should be evaluated in a second independent group of patients.<sup>3</sup> If the findings are replicated, healthcare providers can have more confidence in the accuracy of the test results. For example, the combined use of gram staining and Acridine-Orange Leucocyte Cytospin testing to rapidly diagnose catheter related blood-stream infections without removing the central venous device

has been favourably reported, but the study did not evaluate the test on a second group of patients;<sup>9</sup> hence, the call for confirmation studies before the test is widely accepted.<sup>10</sup>

### Answering the original question

The study by Whooley *et al* probably met 3 of the 4 validity criteria. Firstly, the investigators compared the case finding instruments with an acceptable reference standard, a computerised version of the Diagnostic Interview Schedule (DIS). This is a 20 minute interview with a sensitivity of 80% and a specificity of 84% when compared with DSM III criteria for depression. Three trained interviewers who administered the reference test were blinded to the results of the screening tools. There was a high level of inter-rater agreement between the 3 interviewers with respect to the results of the reference test ( $\kappa=0.88$ ). Secondly, the study sample was a real clinical population in a primary care setting, with patients representing the full spectrum of depressive histories: recent episodes of depression, lifetime history of depression, and no history of depression. Cautious consideration needs to be given to some of the sample's features (such as the high prevalence of depression, the ratio of men to women, and the high number of unemployed people), but these can be addressed when considering the applicability of the study to your own patients. Thirdly, all 7 screening instruments and the reference test were administered to 542 consecutive participants attending an urgent care clinic at a Veterans Administration medical centre, although the results from 7 participants were excluded from analysis because of missing data. The results of the case finding instruments did not appear to influence whether the DIS was done. However, the study does not meet the fourth criterion for validity, as the study findings were not evaluated in a second group. No other evaluation of the instrument seems to have been done, although one is currently under way in general practice populations in New Zealand (personal communication, B Arroll).

### What are the results?

When patients present to healthcare providers, they have a probability of having particular disorders. This probability is the baseline prevalence of each disorder in the community. But each patient is different. Think about 2 patients, both presenting with a small ulcer involving the medial malleolus, with ankle flare, presence of haemosiderian pigmentation, and a history of varicose veins. One is a 71 year old woman who is otherwise healthy and the second is a 55 year old man with type 2 diabetes. Although venous aetiology accounts for up to 70% of all leg ulcers,<sup>11</sup> an experienced clinician knows that the baseline or pretest probability of the ulcer being venous for these 2 patients is different. For the first patient, who has an uncomplicated presentation, the pretest probability of having a venous ulcer is likely to be between 50% and 70%. Following simple assessments of her blood supply to rule out other causes, the experienced clinician is likely to recommend that the patient start compression treatment. However, the pretest probability for venous ulceration is likely to be considerably lower in the second patient. Venous disease only causes 6% to 9% of leg ulcers in patients with diabetes.<sup>12</sup> Simple assessments to rule out other causes of the ulcer may not convince the clinician that the ulcer is venous. Treatment for venous ulceration involves applying high compression bandaging to the patient's affected limb, but the bandaging can create an ischaemic leg if the patient has arterial insufficiency. The clinical hazard of misdiagnosis and ischaemia has increased the threshold for beginning treatment and the low pretest probability means that the clinician may

prefer to refer the patient for further testing before being reassured that compression is safe.

The above example illustrates that no matter what the outcome of an assessment or test is, it cannot tell a clinician whether or not the patient has the disorder. It can only reveal the probability of having or not having the disorder.<sup>13</sup> The ability to discriminate between people likely to have a disorder and those less likely to have a disorder is determined by a test's *likelihood ratio (LR)*. With respect to screening for depression, Whooley *et al* found several instruments with similar results, but the simplest instrument was the 2 question instrument. The reference test indicated that 97 patients had depression. The 2 question case finding instrument correctly identified 94 of these 97 patients (94/97 or 0.97) as likely to be depressed. However, the instrument also incorrectly classified 189 patients as likely to be depressed from the 439 patients (189/439 or 0.43) whom the reference test ruled out as not depressed (table 1). The ratio between these 2 likelihoods is the LR. When considering LRs, it is the percentages or proportions of patients that the test correctly and incorrectly identifies as having the disorder that is considered, not the actual numbers of patients. Thus, the ratio of true positive results (ie, those that the instrument correctly identifies as being depressed) to false positive results (those that the instrument incorrectly identifies as being depressed) is 0.97/0.43, or 2.25. This is the *likelihood ratio for a positive test result (+LR)* being correct. From the +LR 2.25, we can infer that a positive result from the 2 question instrument is only about 2 times more likely to be a true positive than a false positive result. If this instrument were used to diagnose depression, clinicians would be wrong quite often. Clearly, the 2 question case finding instrument is not very effective at diagnosing if a patient is depressed.

Just as the +LR can be calculated, the likelihood of the instrument being wrong when it returns a negative result can also be calculated. The 2 question instrument missed 3 of the 97 depressed patients (3/97 or 0.03), but correctly identified 250 patients as unlikely to be depressed out of the 439 patients (250/439 or 0.57) in which depression was absent. The ratio of false negative results (ie, those that the instrument incorrectly identifies as not being depressed) to true negative results (those that the instrument correctly identifies as not being depressed) is 0.03/0.57, or 0.05. This is the *likelihood ratio for a negative test result (-LR)* being wrong. From the -LR 0.05, we can infer that very few patients are likely to be depressed when the case finding instrument returns a negative result.

Table 1 Results of a 2 question tool as a case finding instrument for depression

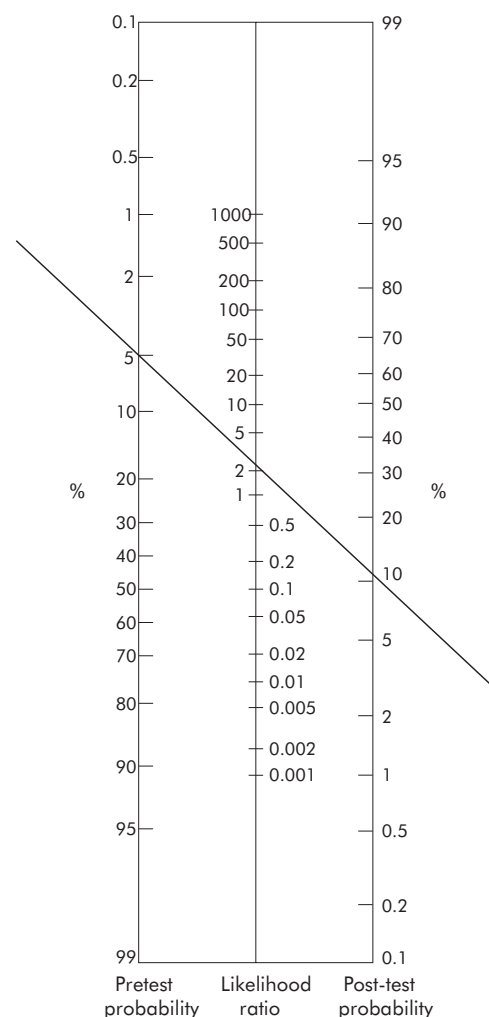
Test results	Depression		Likelihood ratio
	Present	Absent	
Present	0.97	0.43	2.25
Absent	0.03	0.57	0.05
Total	1.00	1.00	

Table 2 Size of likelihood ratios (LRs) and associated utility of changes in probability<sup>13</sup>

Size of LRs	Utility
+LR>10 or -LR<0.1	Generate large changes from pre to post-test probability
+LR 5–10 or -LR 0.1–0.2	Generate moderate shifts in pre to post-test probability
+LR 2–5 or -LR 0.5–0.2	Generate small changes in pre to post-test probability
+LR 1–2 or -LR 0.5–1	Generate little or no shift in pre to post-test probability

The usefulness of LRs is revealed when we look at their ability to shift a patient from a pretest probability to a post-test probability, and in doing so, help reduce the clinical uncertainty associated with case finding, screening, or diagnosis. A rough guide to the magnitude of LRs and their effect on post-test probability is shown in table 2.

The challenge in working out the changes in probability of a patient having a disorder after a test is eased by a simple nomogram (figure).<sup>14</sup> By running a straight line through the pretest probability (left hand column) and the LR (centre column), the post-test probability can be determined from the point at which the line intersects the right hand column. A pretest probability could simply be the prevalence of depression in the community, which has been estimated to be 5% of the adult population in Great Britain.<sup>15</sup> If the patient in our scenario answers yes to both questions, we can extend a line from a pretest probability of 5% through approximately 2 (+LR 2.25) to obtain a post-test probability of a little more than 10% that our patient actually has depression. However, if our patient answers no to both questions, we can extend a line from 5% through 0.05 (-LR) to obtain a posttest probability of approximately 0.03% of being wrong if we accept our patient is not depressed. Thus, we can be confident that if a patient answers no to the 2 questions, he or she is very unlikely to be depressed. On the other hand, a post-test probability of approximately 10% might not be high enough even to consider referral for further testing unless there



Nomogram for determining posttest probabilities.<sup>14</sup> Reproduced with permission from Fagan TJ. *N Engl J Med* 1975;293:257.



is no other likely explanation for the patient's symptoms. However, the tool may be useful at determining whether further testing is desirable in settings where the pretest probability is higher.

Whooley *et al* provided the LRs for each of the case finding instruments. Older studies often do not report LRs, but instead report the sensitivity and specificity of the tests. The *sensitivity* of a test is the proportion of patients with the target disorder who have a positive test result, whereas the *specificity* is the proportion of patients without the target disorder who have a negative test result. LRs are easily obtained if the sensitivity and specificity of a test are known. The sensitivity and specificity of the 2 question case finding instrument are 0.97 and 0.57, respectively. A +LR is obtained by the following formula:

$$\frac{\text{sensitivity}}{1-\text{specificity}} \text{ or } \frac{0.97}{1-0.57} = \frac{0.97}{0.43} = 2.25$$

Similarly -LR can be obtained using a slightly different formula:

$$\frac{1-\text{sensitivity}}{\text{specificity}} \text{ or } \frac{1-0.97}{0.57} = \frac{0.03}{0.57} = 0.05$$

Sometimes sensitivity and specificity are presented as percentages (ie, 97% and 57%). The same formulas can be used, substituting 100 for 1 when subtracting. For further explanation of how sensitivity and specificity are calculated, see Sackett *et al*<sup>2</sup> or any text on clinical epidemiology.

### Can I apply this test to my patient?

We have determined that the study by Whooley *et al* is probably valid and decided that the results indicate that the instrument (1) may be useful for identifying patients who may benefit from referral for further testing when the patient responds positively to the questions, and (2) is useful for ruling out depression as a possibility when the patient responds negatively to the questions. The next step is to determine whether it can be used with your presenting patient or group of patients. Answering 3 questions will assist this decision.

#### IS THE TEST AVAILABLE, AFFORDABLE, ACCURATE, AND PRECISE IN YOUR SETTING?

Obviously if a test is not available, or the costs are similar to equally accurate and usable alternatives, then it is unlikely to be used. Similarly, we need to be assured that a test will maintain its accuracy in the clinical setting in which we work. LRs can be stable, but they are derived from selections of patients, and thus may not be as accurate for patients who are selected in different ways. In an earlier question about validity, we needed to be assured that the instrument was tested in patients with mild, moderate, and severe conditions as well as those without the disorder. Now we need to be assured of the similarity of the study population to that in our own setting. It is uncommon to find a report that exactly describes a population of patients like our own, so we need to examine the demography of the study participants to decide whether they are so dissimilar from our own to rule out using the study. Another concern about the accuracy of a test is that many instruments are reported as having only one +LR and one -LR, although a test can behave differently depending on the severity of the disorder. Higher LRs are found with florid conditions and lower ones with earlier presentations of the disorder. Some tests make this distinction by reporting LRs for different presentations of the disorder, but this would be unusual for screening or case finding tools.

#### CAN WE GENERATE CLINICALLY SENSIBLE ESTIMATES OF PATIENTS' PRETEST PROBABILITIES?

Pretest probability is the probability that a presenting patient has a particular disorder. Sackett *et al* identify 5 different sources for estimating pretest probability: clinical experience, prevalence statistics, practice databases, studies specifically focused on determining pretest probabilities, and the original study itself.<sup>2</sup> Clinical experience will generate what is essentially a "guesstimate", and several false heuristics may influence such an estimate. However, in the absence of other sources, this method can still be useful. Prevalence statistics can be drawn from regional or national morbidity data, or from studies investigating the prevalence of a disorder, but these estimates are only as good as the sources of the data or the settings of the prevalence studies. Databases that rely on voluntary reporting can have inaccurate data. If the prevalence study is set in an acute care setting, the results can be misleading if applied to primary care settings. Practice databases, whether local, regional, or national, are also only as good as their data sources. Studies investigating pretest probabilities are few in number and difficult to retrieve from databases. Finally, the prevalence of the disorder in the study being critically appraised can be used.

#### WILL THE RESULTING POST-TEST PROBABILITIES AFFECT PATIENT MANAGEMENT?

The major concern here is whether the results will move a patient across a threshold that would stop further testing for the suspected disorder. This would occur when a disorder has been ruled out, when a referral for further testing or treatment is made, or when treatment is initiated. For example, if the pre-test probability for depression is 5%, and the patient response to the 2 question case finding instrument is negative, the post-test probability would be so low that depression could be abandoned as an explanation for the patient's symptoms. However, if the patient response was positive (and remembering that the post-test probability was slightly > 10%), it would still be too low to move the patient over a treatment threshold, or perhaps even to referral for further testing. On the other hand, if the pretest probability was higher, perhaps because of a higher prevalence of depression in people with diabetes, then referral for further investigation might be warranted.

#### Resolution of the scenario

Although you accept that the study by Whooley *et al* has reasonably strong validity, you have reservations about using the 2 question case finding instrument with all of your patients. The study sample was primarily men (97%), 71% were unemployed, and the setting was an urgent care medical clinic where patients had a high prevalence of depression (18%) rather than a community population where the prevalence is probably much lower. Thus, you decide to continue looking for a case finding instrument that has been evaluated in different populations and may be more generally applicable to your patients. Until you find such a tool, the 2 question instrument could be useful with this particular patient. You have a feeling that depression is more frequent in people with diabetes than in the general population, and a quick search of the literature reveals a systematic review of the prevalence of depression people with diabetes that confirms this view.<sup>16</sup> The mean rate of current depression (as opposed to lifetime history of depression) in controlled studies is reported as 14%, almost 3 times that of the general population. You decide to use this as your pretest probability. A quick check of the nomogram shows that if the patient answers yes to both questions, then the post-test probability of depression will be approximately 27%, which is high enough to

suggest the need for further testing. Given that the 2 question tool is no less an invasion of privacy than obtaining a blood sample for testing, you decide to discuss with your patient the possibility of a clinical cause for his symptoms the next time you visit to change his ulcer dressing. If he is willing, you resolve to use the 2 question case finding instrument to screen for depression.

ANDREW JULL, RN, MA

*Clinical Trials Research Unit, University of Auckland  
Auckland, New Zealand*

- 1 Reid MC, Lachs MS, Feinstein AR. Use of methodological standards in diagnostic test research. Getting better but still not good. *JAMA* 1995;**274**:645–51.
- 2 Sackett DL, Straus SE, Richardson WS, *et al.* *Evidence-Based Medicine. How to practice and teach EBM*. 2nd edition. London: Churchill Livingstone, 2000.
- 3 Whooley MA, Avins AL, Miranda J, *et al.* Case-finding instruments for depression. Two questions are as good as many. *J Gen Intern Med* 1997;**12**:439–45.
- 4 Jensen BN, Jensen FS, Madsen SN, *et al.* Accuracy of digital tympanic, oral, axillary, and rectal thermometers compared with standard rectal mercury thermometers. *Eur J Surg* 2000;**166**:848–51.
- 5 Lijmer JG, Mol BW, Heisterkamp S, *et al.* Empirical evidence of design-related bias in studies of diagnostic tests. *JAMA* 1999;**282**:1061–6.
- 6 van der Schouw YT, Verbeek AL, Ruijs SH. Guidelines for the assessment of new diagnostic tests. *Invest Radiol* 1995;**30**:334–40.
- 7 Yao ST, Hobbs JT, Irvine WT. Ankle systolic pressure measurements in arterial disease affecting the lower extremities. *Br J Surg* 1969;**56**:676–9.
- 8 Stoffers HE, Kester AD, Kaiser V, *et al.* The diagnostic value of the measurement of the ankle-brachial systolic pressure index in primary health care. *J Clin Epidemiol* 1996;**49**:1401–5.
- 9 Kite P, Dobbins BM, Wilcox MH, *et al.* Rapid diagnosis of central-venous-catheter-related bloodstream infection without catheter removal. *Lancet* 1999;**354**:1504–7.
- 10 Farr BM. Accuracy and cost-effectiveness of new tests for diagnosis of catheter-related bloodstream infections. *Lancet* 1999;**354**:1487–8.
- 11 Nelzen O, Bergqvist D, Lindhagen A. Venous and non-venous leg ulcers: clinical history and appearance in a population study. *Br J Surg* 1994;**81**:182–7.
- 12 Baker SR, Stacey MC, Jopp-McKay AG, *et al.* Epidemiology of chronic venous ulcers. *Br J Surg* 1991;**78**:864–7.
- 13 Jaeschke R, Guyatt G, Lijmer J. Diagnostic tests. In: Guyatt G, Rennie D, editors. *Users' guides to the medical literature. A manual for evidence-based clinical practice*. Chicago: American Medical Association, 2002:121–140.
- 14 Fagan TJ. Nomogram for Bayes theorem. *N Eng J Med* 1975;**293**:257.
- 15 Ohayon MM, Priest RG, Guilleminault C, *et al.* The prevalence of depressive disorders in the United Kingdom. *Biol Psychiatry* 1999;**45**:300–7.
- 16 Gavard JA, Lustman PJ, Clouse RE. Prevalence of depression in adults with diabetes. An epidemiological evaluation. *Diabetes Care* 1993;**16**:1167–78.

# Want to extend your search?

## Cross journal searching

If you can't find what you are looking for in *Evidence Based Nursing* you can extend your search across many of the more than 200 journals available for selection. You can restrict your search to specific subject areas (eg, clinical medicine, basic research), or select specific journals, or search all available titles.

**[www.evidencebasednursing.com](http://www.evidencebasednursing.com)**

**The virtual journal Paediatric Asthma, produced by the American Academy of Pediatrics and the BMJ Publishing Group, is now available free of charge to EBN subscribers at [www.pediatric-asthma.org](http://www.pediatric-asthma.org)**