T - 14



It is this emphasis on the importance of epidemiology for clinical decision making that has inspired this series of articles, entitled "Principles behind practice". However, it is important to point out that the scope of clinical epidemiology is much broader. Many of the clinicians who study epidemiology in depth (as for a Master degree) do so to learn research skills which can be applied to major health problems. The application by the clinician of research methods to study health problems of importance to the population is a way of bridging the gap between clinical practice and public health.

This approach has been adopted by some North American training programmes for clinicians, the best example of this being the Robert Wood Johnson Foundation-sponsored Clinical Scholars Program.[5] On the international scene, the Rockefeller Foundation has established an International Clinical Epidemiology Network (INCLEN).[4] The Centre for Clinical Epidemiology and Biostatistics in the Newcastle Faculty of Medicine is the only training centre in the INCLEN network outside North America. INCLEN has fostered the development of 26 units in teaching hospitals across the developing world where clinicians from various specialities are performing important research which is influencing health policy.

The series, "Principles behind practice", starts with an article on diagnostic tests in this issue (page 33) and continues with articles on clinical decision making, natural history and prognosis, how to determine causation, types of research study, health economics and prevention in the clinical setting. Each is intended to be of use as well as of interest. We hope that this series will not only help all clinicians with clinical decisions, but inspire some to develop further skills in research and apply these skills to the health problems of the country.

### References


1. Paul JR. Clinical epidemiology. *J Clin Invest* 1938; 17: 539-541.
2. Fletcher RH, Fletcher SW, Wagner EH. Clinical epidemiology: the essentials. Baltimore: Williams and Wilkins, 1988.
3. Cannon PJ, Connell PA, Stockley IH, et al. Prevalence of angina as assessed by a survey of prescriptions for nitrates. *Lancet* 1988; 7: 979-981.
4. Anonymous. Clinical epidemiology in the Third World [Editorial]. *Lancet* 1982; 1: 1448.
5. Shuster AL, Cluff LE, Haynes MA, et al. An innovation in physician training: the clinical scholars program. *J Med Educ* 1983; 58: 101-111.


# 2. Using the diagnostic test


Richard F Heller and Greg Whelan


In this paper we consider the interpretation of test results. For each of the examples given, some common principles apply and we need to obtain information about the way diagnostic tests work. For each case, we need to know **sensitivity, specificity,** and **positive** and **negative predictive value** of the test. In addition we must decide if the test and its result **make a difference** to our management plan for the patient.


**Centre for Clinical Epidemiology and Biostatistics, University of Newcastle, Royal Newcastle Hospital, Newcastle, NSW 2300.**
Richard F Heller, MD BS (Lond), FRCP, FFCM, FRACP, Professor of Community Medicine and Director, Centre for Clinical Epidemiology and Biostatistics.
**Department of Community Medicine, St Vincent's Hospital, 41 Victoria Parade, Fitzroy, VIC 3065.**
Greg Whelan, MB BS (Syd), MD (Melb), FRACP, Director.
Reprints: Professor Richard F Heller.


**Case 1**

*A 60-year-old man comes to see you complaining of chest pain. You arrange for a stress test (exercise ECG) and the result comes back "positive".*

How do you interpret this result?
Would a positive result have different implications if the patient with chest pain were a 30-year-old woman?

*You happen to listen to the man's neck and hear a carotid bruit.*

What test should you do for this? How can you interpret the results?

### Case I: Positive exercise ECG in a 60-year old man with "typical" chest pain

Before interpreting this positive result, we should make an estimate of the likelihood of this man having coronary artery disease (CAD). We estimate, on the basis of experience or published data, that a man of this age with "typical" angina-type chest pain has a 90% chance of having coronary artery disease. We are told that the exercise ECG has a **sensitivity** of 71% and a **specificity** of 73% — as compared with the "gold standard" of coronary angiography.[1] We now need to construct a table in which, of 100 such men, 90 will have coronary artery disease and the test will identify 71% of these as "positive" (this is the definition of **sensitivity**). Seventy three per cent of the 10 men without coronary artery disease will be identified as "negative" by the test (**specificity**) and we can make our table as follows:

| | | Coronary Artery Disease* | | |
|---|---|---|---|---|
| | | Yes | No | |
| Exercise ECG | Positive | 64 | 3 | 67 |
| | Negative | 26 | 7 | 33 |
| | | 90 | 10 | 100 |

*On angiography.

We see that of these 100 men, 67 will have a "positive" exercise ECG. Of these, 64 will actually have CAD ("true positive") but three will not have CAD ("false positive"). The chance of a positive test result reflecting the true presence of disease is called the **positive predictive value**, which here is 96%. Similarly, the **negative predictive value**, the chance of a negative test result reflecting true absence of disease, is 21% (7/33). A positive test result is fairly accurate, but a negative test result is quite unlikely to really mean that CAD is absent.

In this case, has the test **made a difference**? Not really. We already knew that there was a 90% chance of CAD being present. A positive result increases this likelihood to 96%, which, although a slight improvement, has not really improved things to such an extent that you would make a different decision about the next step in management of the patient. A negative result would be *highly confusing — so we would have been better* off not doing the test! Of course, if we had not known the likelihood of a 60-year-old man with typical chest pain having CAD (this is also called the pre-test probability) then a positive result would help a lot — but a negative result would still be very confusing. Unless we have some idea of the pre-test probability of the presence of disease, as well as of the sensitivity and specificity of the test, we really cannot interpret the result. What is very clear is that a positive or a negative test result does not necessarily mean that the answer is "correct".

#### *Positive exercise ECG in a 30-year-old woman with "atypical" chest pain*

Consider this quite different situation. We estimate a pre-test probability of the presence of CAD here of only 10% (quite different from the 90% quoted in the case of the 60-year-old man). Putting the figures in our table (with the same sensitivity and specificity) we find:

| | | Coronary Artery Disease | | |
|---|---|---|---|---|
| | | Present | Absent | |
| Exercise ECG | Positive | 7 | 24 | 31 |
| | Negative | 3 | 66 | 69 |
| | | 10 | 90 | 100 |

We end up with a **positive predictive value** of only 23% (7/31), suggesting that a positive test result is likely to be highly inaccurate. The **negative predictive value**, however, of 96% (66/69) suggests that a negative test result really does mean that CAD is unlikely to be present. Of course, we knew that anyway, but it may be of value for the patient to have "confirmation".

You could calculate some figures yourself. For example, if the pre-test probability was 50%, a positive result would increase the chance of disease being present (positive predictive value) to 71% and a negative result would increase the chance of disease being absent also to 71%.

#### *The asymptomatic neck bruit*

Duplex Doppler ultrasound is a test with good sensitivity and specificity (85% and 90% respectively).[2] We certainly could order the test, but should only do so if the result is going to influence our management or **make a difference**.

If a positive ultrasound test result is obtained and surgery is planned, we will need an angiogram. Serious complications occur in 1%–2% of people as a result of carotid angiography,[2] and the chance of a stroke or death as a result of surgery may be as much as 10%.[3] Since the natural history of an asymptomatic neck bruit is only a stroke rate of 2% per year,[4] we would be better off not operating. (This is quite different in the presence of transient ischaemic attacks, where surgery or aspirin may well improve the natural history of a carotid artery bruit.)[5,6] If we are

not going to intervene, why make a diagnosis?

It is thus essential to decide **before** ordering the test if management is to be changed as a result of the test findings — if the test will **make a difference**. If it will not, do not order the test.

**Case II**

*A 60-year-old woman complains of dysuria.*

**You arrange for an urgent microscopy and find $10^3$ bacteria per millilitre. Is this likely to be "significant"? Should you treat now or wait for a culture result?**

## Case II: 60-year-old woman with dysuria

The **sensitivity** of the mid-stream urine examination (as compared with a "gold standard" of suprapubic needle aspiration and culture) has been found to be 95% with a cut-off point of $10^2$ bacteria per millilitre and a specificity of 85%.[7] Our patient has $10^3$ bacteria per millilitre, and the result would thus be called "positive" if the $10^2$ cut-off point was used. If we use $10^5$ bacteria per millilitre as a cut-off point (which would exclude our patient), we pick up only 51% of those with true infection, but the specificity increases to 99%. Unfortunately, the $10^2$ cut-off point also produces many false positives, many more than when the $10^5$ cut-off point is used. There is always a trade-off between **sensitivity** and **specificity**; as one increases, the other decreases.

If we assume a pre-test probability of approximately 50% in such 60-year-old symptomatic patients (that is, 50% of women of this age presenting to a doctor with these symptoms will actually have an infection), our tables will look as follows:

| | | True Urinary Tract Infection | | |
|---|---|---|---|---|
| | | Yes | No | |
| $10^3$/mL bacteria | Yes | 47 | 8 | 55 |
| | No | 3 | 42 | 45 |
| | | 50 | 50 | 100 |

| | | True Urinary Tract Infection | | |
|---|---|---|---|---|
| | | Yes | No | |
| $10^5$/mL bacteria | Yes | 25 | 1 | 26 |
| | No | 25 | 49 | 74 |
| | | 50 | 50 | 100 |

You can calculate your own **positive** and **negative predictive values** and decide which cut-off you wish to use. The **makes a difference** criterion is vital here. You may feel it more important to "overtreat", that is, to use the $10^2$ point knowing that, if you treat on the basis of bacteria count and microscopy, you will treat nearly all of those who truly have infection as well as a number without infection. If you would rather avoid overtreating because of fears of antibiotic overuse, you would choose the $10^5$ point, knowing you will miss a proportion of those with true disease.

In a different clinical situation — for example, using the mid-stream urine examination as a screening test in asymptomatic people — the operating characteristics are quite different. Here, the pre-test probability (prevalence of disease) is only 2%, and our choice of cut-off point may be different. You may wish to perform the calculations yourself.

**Case III**

*A 15-year-old girl complains of a severe sore throat.*

**Should you treat or investigate? If you investigate, which tests provide "useful" information and how do you translate this information into a decision regarding treatment?**

## Case III: 15-year-old girl complaining of a severe sore throat

Would you treat or investigate and only treat those individuals with a positive throat culture?

Before examining the patient we estimate, on the basis of visits to emergency rooms, that an individual with a sore throat has about a 25% chance of having a β-haemolytic streptococcal pharyngitis.[8] We are told that physical examination (which includes recording the temperature and looking for erythema, exudate on the tonsils, lymphadenopathy and enlargement of the tonsils or swelling of the pharynx) has a sensitivity of 73% and a specificity of 69% as compared with the gold standard of bacterial culture.[8] If we now construct the table as we have in the previous examples, then, of 100 such individuals, 25 will have bacteriologically proven pharyngitis, and our clinical diagnosis will indicate that 73% of these patients with bacteriologically proven pharyngitis will be called "positive" by our physical examination (sensitivity), while 69% of patients who do not have bacteriologically proven pharyngitis will be called "negative" by our physical examination (specificity). Thus, as we can see in our table, of these 100 individuals, 41 will have a clinical diagnosis of bacterial pharyngitis. Of these, 18 will actually have bacteriologically proven pharyngitis

(true positive results), but 23 will not have bacteriologically proven pharyngitis (these are false positive results). The likelihood of a positive test result indicating the true presence of disease (the positive predictive value) is here 44% (18/41). Similarly the negative predictive value (the chance that a negative test result truly reflects the absence of disease) is 88% (52/59). Thus we can see that a clinical diagnosis of bacterial pharyngitis is similar to a coin flip. That is, about half the time it will be accurate and about half the time it will not. On the other hand, a negative physical examination result is helpful in that it indicates that such an individual is unlikely to have bacteriologically proven pharyngitis.

| | | β-Haemolytic Streptococcal Pharyngitis* | | |
|---|---|---|---|---|
| | | Present | Absent | |
| Clinical diagnosis | Yes | 18 | 23 | 41 |
| | No | 7 | 52 | 59 |
| | | 25 | 75 | 100 |

*On throat culture

In answering the question "Has the test — that is, the physical examination — made a difference?", we need to look at the gain and the setting. A positive physical examination result raises the post-test probability of disease when compared with the pre-test probability of disease from 25% to 44% — a gain of 19%. A negative physical examination result raises the pre-test probability of the absence of disease from 75% to 88%. Thus, initially, it is not clear whether we have gained much at all. To determine whether this does make a difference we need to look at the setting.

If a bacteriological test is easily performed and cheap, and the results are rapidly available, then physical examination has little benefit in this setting. If, on the other hand, laboratory tests are difficult to get, and the results are unreliable and not available for a period of time (such as in a rural setting), physical examination will be helpful in removing from further consideration a large number of individuals who have a low probability of disease. One would then treat those with positive physical signs. The costs and implications of therapy may need to be taken into account. If therapy is cheap with very few side effects, then logically it should be given to everybody with a sore throat with any physical signs. On the other hand, if the only available therapy is toxic and expensive, then therapy should be confined to those individuals with a positive clinical diagnosis.

### *Discussion*

Sore throat is a common problem in general practice. The majority of such individuals will not have bacterial pharyngitis or tonsillitis. However, there is good evidence that appropriate treatment of group A β-haemolytic streptococci reduces the likelihood of subsequent rheumatic fever, and thus it is important to recognise and treat such individuals. Studies have indicated that approximately 5% of all patients attending the adult medical section of an emergency room will complain of a sore throat and the prevalence of β-haemolytic pharyngitis among this group of individuals with pharyngitis is approximately 25%. Rheumatic fever is less common than it was many decades ago. Some physicians believe that widespread use of antibiotics in individuals with sore throat has been responsible for this reduction. There is no widespread consensus as to whether antibiotics should be used in this setting. Since the prevalence of disease is 25%, it is likely that in most individuals antibiotics will make no difference in the outcome in either the short or the long term.

Antibiotics are not without side effects. Some individuals will develop diarrhoea, sometimes seriously. A "wait and see" policy in other words "culture and return") adds an extra visit to the hospital for the individual, and needs to be done in a setting where the bacteriological support services are easily available and well coordinated with the physician's needs. Throat culture results need to be rapidly available, and there needs to be a regular mechanism within the hospital to ensure that physicians ordering laboratory tests receive the results of these tests.

Non-compliance amongst emergency room patients directed to return for follow-up care has been well documented and there is usually not an effective administrative service to help locate patients with positive throat cultures. Because of the rostering situation the physician may not be "on service" when the individual returns.

The picture in general practice may be quite different. The prevalence of β-haemolytic pharyngitis may also be different. The good clinician develops a working knowledge of the prevalence of disease in his or her setting — as we have seen, this is a vital step in interpreting the results of a diagnostic test.

## Some conclusions

1. To interpret a diagnostic test result you need to know: the sensitivity and specificity of the test; and the likely prevalence of disease in such

people in your clinical setting.
2. From the information you can calculate the chance of a positive (or negative) test result truly reflecting the presence (or absence) of disease. (A positive test result does not always mean the disease is present — and a negative test does not always mean the disease is absent.)
3. A diagnostic test should only be performed if the result is likely to make a difference to your management.

### References


1. Detrano R, Froelicher UF. Exercise testing: uses and limitations considering recent studies. *Prog Cardiovasc Dis* 1988; 94: 737-744.
2. Health and Public Policy Committee, American College of Physicians. Diagnostic evaluation of the carotid arteries. *Ann Intern Med* 1988; 109: 835-837.
3. Winslow CM, Solomon DH, Chassin MR, et al. The appropriateness of carotid endarterectomy. *N Engl J Med* 1988; 318: 721-727.
4. Chambers BR, Morris JW. Outcome in patients with asymptomatic neck bruits. *N Engl J Med* 1986; 315: 860-865.
5. Antiplatelet Trialists' Collaboration. Secondary prevention of vascular disease by prolonged antiplatelet treatment. *Br Med J* 1988; 296: 320-331.
6. Dyken ML. Transient ischaemic attacks and aspirin, stroke and death: negative studies and type II error [Editorial]. *Stroke* 1983; 14: 2-4.
7. Stamm WE, Counts GW, Running KR, et al. Diagnosis of coliform infection in acutely dysuric women. *N Engl J Med* 1982; 307: 463-468.
8. Fletcher SW, Hamann C. Emergency room management of patients with sore throats in a teaching hospital. *J Community Health* 1976; 1: 196-207.


## OTHER PEOPLE'S PRACTICES

# Prison doctor

W Pelham Sapsford

The Sir David Longlands' Correctional Centre is one of many correctional facilities in the outer western suburbs of Brisbane. It houses up to 264 adult male prisoners, for whom I, as a Government Medical Officer (GMO), have provided a general practitioner service for one year. There are very real challenges, both practical and ethical, in providing prisoners with health care which is both effective and humane.

There are two Government Medical Officers, one present in the prison for one hour every morning and one present for three hours every afternoon. The morning medical officer sees acutely ill patients, while the afternoon doctor examines new inmates. The morning doctor sees approximately 100 patients per month and the afternoon doctor sees between 250 and 300 patients per month.

**Suite 111, Silverton Place, 101 Wickham Terrace, Brisbane, QLD 4000.**
W Pelham Sapsford, MB BS, General Practitioner.
Reprints: Dr W Pelham Sapsford.

There are also five registered nurses who keep the prison surgery open for approximately 16 hours per day. These nurses attend to about 300 patients per month.

The prison hospital is near the entrance of the prison, but is centrally located for both work blocks and cell blocks. It covers 30 to 40 squares, and includes a nurse's station, with rooms for stores and drugs, and a large waiting room with a bathroom. There are consulting areas for a dentist and two doctors, as well as an operating room with full sterilisation equipment. In addition the building houses a security room, a staff room, a holding room, a charge nurse's office and an observation room which can take four beds.

There is a good range of diagnostic and consultant services, including x-ray facilities; a radiographer comes every day and takes 150 to 200 x-rays per month. These are reviewed by the GMO on the spot and also later by a radiologist at the Princess Alexandra Hospital. The dentist visits twice a week and has between 50 and 70 patients per month. There are two and occasion-