

Best tests?

The general principles of laboratory investigations in primary care

Laboratory investigations are essential for the diagnosis and management of many conditions. However, laboratory tests do not provide clinical value in every scenario, and in some cases, they may even cause harm. Before a laboratory test is requested, clinicians should consider the aim of the test and have a clear understanding of how the result will be interpreted and how the patient's management will be affected by the result. Understanding the clinical situations where laboratory testing may be problematic can help to improve the overall approach to testing.

“Testing, testing: one, two, three”

- 1. Think twice before you test**
- 2.. Select the right test, at the right time, for the right patient**
- 3.. Ask yourself: can I improve my testing?**

Think twice before you test

Laboratory tests are generally requested in primary care for one of the following reasons:

1. **Diagnosis:** to either include or exclude a disease, e.g. thyroid stimulating hormone (TSH) levels in a patient with suspected thyroid dysfunction
2. **Establishing a baseline prior to treatment initiation:** e.g. liver function test (LFT) before commencing methotrexate
3. **Monitoring:**
 - a. To ensure a medicine is within a therapeutic range, e.g. patients taking lithium where the serum lithium concentration relates to clinical effect
 - b. To detect early signs of an adverse effect to treatment, e.g. full blood count in patients taking clozapine
 - c. To monitor or predict the response to treatment, e.g. INR assessment in patients taking warfarin, serum urate monitoring in patients taking allopurinol, or antimicrobial susceptibility of a pathogen
 - d. To monitor long-term conditions for disease control and associated complications, e.g. the monitoring of HbA_{1c} and albumin creatinine ratio (ACR) in people with diabetes
4. **Targeted testing,** e.g. antenatal screening for rubella status, lipid levels as part of a cardiovascular assessment

In each of these situations the test result will benefit the patient and the clinician by allowing better decisions to be made about future management.

When considering laboratory investigation it is important to acknowledge that testing is not always beneficial, and that in certain situations the balance may shift from benefit to harm. Understanding the clinical situations which may lead to a poor outcome can provide insight into when to be more cautious in deciding if laboratory investigation is needed, or if a request for tests should be deferred or delayed.

Considerations before testing

Before requesting a laboratory test it may be helpful for clinicians to consider their answers to the following questions:

- What is my reason for requesting this test?
- Will the test improve patient (or in some cases, family or partner) care?
- Is this the right test or combination of tests for the clinical situation?
- How will the test result be interpreted?
- How will the test result influence patient management?
- Are there potential harms of doing this test?

The following examples demonstrate potentially problematic scenarios when considering laboratory investigations.

Laboratory tests may reveal incidental findings

The early discovery of dormant conditions or incidental findings that have little or no long-term consequences to the patient's health can be unveiled by laboratory investigations. Once a condition is identified, it can sometimes be difficult for the patient to understand and accept that treatment is not necessary.

Over-diagnosis and over-treatment are the most important adverse effects of screening programmes. Estimates of over-diagnosis of indolent cancers (slow growing and low-grade) in PSA screening populations are 27% at age 55 years and 56% at age 75 years.¹ The risk of patients receiving a diagnosis and treatment for a cancer that would not have affected their long-term outcome is one of the reasons why PSA screening is controversial, especially in older patients with co-morbidities.

A discussion with patients about the potential risks of testing, and consideration of what a positive or negative result will mean for the patient's management, can help when making an informed decision about whether to test or not.

Some symptoms are medically unexplained

In some cases, patients with underlying emotional distress or psychological illness present with a complex pattern of

medically unexplained symptoms, leading to a degree of diagnostic uncertainty. An increasing level of uncertainty about the patient's presenting symptoms and signs, leads to an increasing number of laboratory tests requested.² Rather than clarifying the situation, sometimes this can lead to "digging an even deeper hole", and emphasises the importance of thoughtful test requesting and interpretation.

Normal laboratory results are often not helpful in reassuring such patients. In addition, multiple test requests are also likely to eventually result in a value being identified outside the normal reference range, regardless of whether it is clinically significant (Page 9). This ultimately leads to dissatisfaction for both patients and clinicians, and increased health costs.

This situation provides the "perfect storm" of clinical uncertainty. Instead of requesting a battery of tests, it may be more appropriate to identify any psychological or environmental stressors, or administer a formal depression screening tool. In this scenario, providing an explanation for the patient's symptoms in relation to the psychological problems is likely to have more benefit than a series of laboratory tests.³

Patients ask for tests themselves

Patients often ask for laboratory tests based on their own research, or following consultation with friends or family. A common scenario is for a patient to be concerned about possible dietary deficiencies. Trace element testing, e.g. zinc, copper and selenium, has been increasing in New Zealand in recent years.⁴ However, in most cases, patients are unlikely to have a deficiency, and borderline low levels are a non-specific finding, with low predictive value of organic disease (see Pages 8 and 9 for further discussion of these terms).⁴ For example, transient inflammation is a common cause of low levels of iron and zinc, but can also result in raised copper levels (due to an increase in its binding protein ceruloplasmin).

Education and evidence-based discussions can be helpful in explaining to patients why testing is not always appropriate. Patients need to be aware that they may need to pay for some tests themselves, if they are not clinically justified.

Selecting the right test, at the right time, for the right patient

There are approximately 200 standard laboratory tests in New Zealand, with many more available on request. It has been estimated that if one patient undergoes 12 biochemical tests there is almost a 50% chance that one or more of the results will be outside the normal reference range, even in a healthy person.⁵

Once the decision is made to request laboratory investigation, selecting the right test at the right time for the right patient can sometimes be a challenge. This decision may be influenced by many factors including patient and family/whānau expectations, emerging evidence, changing guidelines, clinical experience and individual clinical, social and cultural factors. All of which are combined with the need to identify the problem within the consultation time, and the natural concern of the clinician not to get it wrong and miss a diagnosis.

In the search for a diagnosis, some patients may undergo numerous investigations, some of which, only in hindsight, will be unnecessary. This does not mean that a carefully chosen test with a normal result was wasted, rather it may have redirected the investigation to another likely explanation. However, in the worst case scenario, a “shotgun” approach to laboratory testing can lead to misdiagnosis and patient harm. By having a clear purpose when selecting a test and selecting the right test, in the right circumstances, with a clear understanding of how results will be interpreted, clinicians can improve patient outcomes while making the best use of tests.

Selecting the most appropriate test

It is estimated that over half of all errors that occur in the process of laboratory investigation, take place during the test selection process.⁶ To reduce the likelihood of errors clinicians should be careful not to request tests that are likely to cause confusion or false reassurance. For example, a faecal occult blood test (FOBT) is inappropriate in an older patient with anaemia, where there is a high suspicion of bowel cancer, as a negative result is not sufficiently reassuring to avoid definitive investigation and a positive result could be due to other causes.

Sometimes it is clear that an investigation is required, but there may be uncertainty as to what test to use. For

example, the routine use of laboratory microscopy and culture is inappropriate for testing for microscopic (non-visible) haematuria. Dipstick analysis alone is sensitive enough to determine the presence of haematuria in patients with suspected renal disease.

The usefulness of some tests depends on the clinical setting. For example, tumour markers are useful tests when used in the appropriate clinical context, e.g. patients receiving cancer treatments, but as a first-line rule in/rule out test for cancer, they have a limited diagnostic value in the large majority of clinical circumstances. In a United Kingdom-based study, requests for tumour marker tests from General Practitioners were studied retrospectively over a 34 month period, and the appropriateness of each test reviewed. Comparison with best-practice guidelines suggested that 84% of the tests requested were inappropriate.⁷

The timing of laboratory tests is an important consideration

Even if a test is appropriate, it needs to be requested at the right time for the patient, and with the right preparation, where necessary.

Some tests require certain factors to be present (or not present) in order to produce a meaningful result. For example, the measurement of antibodies to tissue transglutaminase (TTG) in a patient with suspected coeliac disease may be falsely-negative if the patient has already removed gluten from their diet. A patient undergoing skin prick testing for allergies needs to avoid antihistamine medicines for at least 72 hours prior to undertaking the procedure, in order not to mask any response.

Other tests must be undertaken at specific times. For example, a patient undergoing therapeutic drug monitoring must have samples taken at certain intervals to measure the drug concentration relative to dosing, in order to maximise clinical effect, while avoiding toxicity. Similarly, tests such as cortisol and iron should be measured in the morning, as diurnal variation leads to a fall in levels later in the day.

Some tests must be timed to coincide with a certain stage of the disease cycle. For example, if an HIV serology test is requested too early, seroconversion may not have occurred, and therefore a false-negative result is possible.

In another example, raised serum urate levels are the most important risk factor for gout, however, ideally levels should not be measured during an acute episode of gout, as they may be misleadingly normal during this time in 11–49% of people.⁸

Can I improve my testing?

There are several examples of ways in which clinicians can use laboratory investigations in a more effective way.

Use serial testing rather than parallel testing

Serial testing is when subsequent tests are requested, based on the results of initial tests, rather than testing all at once (i.e. parallel testing). For example if a patient presents with feeling “tired all the time” a clinician may consider a full blood count, ferritin and TSH as first-line tests. Based on these results, the clinician can then instruct the laboratory to further analyse the sample for other tests, such as B12/folate and electrophoresis, if there is unexplained anaemia.

Manage test ordering forms

Electronic laboratory test ordering forms can be customised to reduce the temptation for “checkboxitis”, i.e. routinely selecting certain tests with every laboratory request. Consider moving the position of tests that are frequently ticked, such as antinuclear antibody (ANA) and serum magnesium, to another tab. In a large population of 3000 physicians, reformatting a computerised ordering form resulted in a 36% – 53% decrease in requests for vitamin B12, folic acid and ferritin tests after two months.⁹

 You can find out the number of tests ordered by your practice from your “2012 Annual Report: Pharmaceutical & laboratory test utilisation”.

Be aware of standing orders for tests

Repetition of unnecessary tests can occur when regular tests are automatically repeated, without checking that the clinical justification for testing is still present. For example, continuing to test INR levels in patients who are no longer receiving warfarin or testing lipids in patients no longer receiving lipid-lowering medicines.

Consider if treatment can commence without testing

Vitamin D testing is an example of a laboratory test that is frequently unjustified in New Zealand. Due to seasonal variation in sunlight, most people’s vitamin D levels fluctuate through the year, making interpretation of vitamin D test results difficult. It is therefore recommended that instead of testing people for vitamin D deficiency, clinicians should focus on treating individuals who have a high likelihood of deficiency, e.g. older people in residential care, dark-skinned people or people with evidence of osteoporosis.¹⁰

Despite this guidance, the rate of vitamin D testing in New Zealand has been increasing. Between 2000 and 2010, the number of vitamin D tests in the Auckland region increased almost four-fold from 8 500 to 32 800.¹⁰ At a cost of over \$31 per test, more than \$1 million was spent on vitamin D testing in the Auckland region in 2010.¹⁰ The vast majority of this testing did not reveal a vitamin D deficiency, and in some individuals multiple testing was required before a deficiency was detected - one individual was tested 13 times before a deficiency was found, at a total cost of \$404.¹⁰ This compares to the treatment cost of \$10 per year of vitamin D supplementation.¹⁰

Consult with the laboratory

When in doubt about what test to order, or how to interpret the results, phone the laboratory. Laboratory staff, including pathologists, are available to provide expert assistance, and this resource should be utilised.



ACKNOWLEDGEMENT: Thank you to **Dr Cam Kyle**, Clinical Director of Biochemistry and Immunology, Diagnostic Medlab, Auckland for expert guidance in developing this article.

References

1. Greene KL, Albertsen PC, Babaian RJ, et al. Prostate specific antigen best practice statement: 2009 update. *J Urol* 2009;182(5):2232–41.
2. Davis P, Gribben B, Scott A, Lay-Yee R. The 'supply hypothesis' and medical practice variation in primary care: testing economic and clinical models of inter-practitioner variation. *Soc Sci Med* 2000;50(3):407–18.
3. Salmon P. The potentially somatizing effect of clinical consultation. *CNS Spectr* 2006;11(3):190–200.
4. Song D, Palmer B, du Toit S, Davidson JS. Inappropriate trace element testing in the Auckland region. *N Z Med J* 2010;123(1327):24–34.
5. Deyo RA. Cascade effects of medical technology. *Annu Rev Public Health* 2002;23:23–44.
6. Stroobants AK, Goldschmidt HMJ, Plebani M. Error budget calculations in laboratory medicine: linking the concepts of biological variation and allowable medical errors. *Clin Chim Acta* 2003;333(2):169–76.
7. Walker P, Crook M. Tumour marker requesting in primary care and the role of the laboratory. *J Clin Pathol* 2011;64(5):443–6.
8. Leiszler M, Ahlin S, Fletcher A. Clinical inquiry. Are serum uric acid levels always elevated in acute gout? *J Fam Pract* 2011;60(10):618–20.
9. Kahan NR, Waitman D-A, Vardy DA. Curtailing laboratory test ordering in a managed care setting through redesign of a computerized order form. *Am J Manag Care* 2009;15(3):173–6.
10. Bolland MJ, Grey A, Davidson JS, et al. Should measurement of vitamin D and treatment of vitamin D insufficiency be routine in New Zealand? *N Z Med J* 2012;125(1349):83–91.
11. Hickner J, Graham DG, Elder NC, et al. Testing process errors and their harms and consequences reported from family medicine practices: a study of the American Academy of Family Physicians National Research Network. *Qual Saf Health Care* 2008;17(3):194–200.
12. Wells P, Owen C, Doucette S, et al. Does this patient have deep vein thrombosis? *JAMA* 2006;295(2):199–207.
13. Ruf M, Morgan O. Statistical aspects of screening tests, including knowledge of and ability to calculate, sensitivity, specificity, positive and negative predictive values, and the use of ROC curves. 2008. Available from: www.healthknowledge.org.uk (Accessed Jan, 2013).

Deciding when a test is useful: how to interpret the jargon

The usefulness of any laboratory test is determined by the clinical context. For example, a study of diagnostic tests ordered by 87 General Practitioners for over 1200 patients found that when a test was ordered purely for patient reassurance, approximately 66% of results outside the reference range were interpreted as normal, however, when a test was ordered to confirm a suspected diagnosis, only 28% of results outside the reference range were interpreted as normal.¹¹

To determine the likelihood that a patient has a specific condition, based on a test result, the clinician must first consider:

1. How likely is it that the patient has this condition? This is termed the pre-test probability, and is based on the clinical characteristics of the patient, the local prevalence of the diseases being considered, and the clinician's personal experience.
2. How accurate is this diagnostic test? This is determined by the sensitivity and specificity of the test.

Pre-test probability is defined as the probability that the condition being tested for is the cause of the symptoms, before a diagnostic test result is known. The pre-test probability helps clinicians to decide whether it is worthwhile requesting a diagnostic test. This probability may be altered during the consultation as symptoms and signs are weighted as being "somewhat more suggestive" or "somewhat less suggestive" of the suspected medical condition.

The sensitivity of a test is defined as the proportion of people with the disease who have a "positive" result (above or below the diagnostic threshold used), i.e. the ability of the test to correctly identify patients with the condition. Because the number of false-negatives decreases as the sensitivity of the test increases, a highly sensitive test is useful for "ruling out" a disease if the patient tests negative. Highly sensitive tests, with deliberate use of an appropriate diagnostic threshold for follow-up, are used when the consequences of missing a particular disease are

potentially very serious, such as for an acute myocardial infarction.

The specificity of a test is defined as the proportion of people without the disease who have a "negative" result, i.e. the ability of the test to correctly identify patients without the condition. Because the number of false-positives decreases as the specificity of the test increases, a test with a high specificity is useful in "ruling in" a disease if a person tests positive. As with sensitivity, the specificity of a test will vary somewhat depending on the diagnostic threshold chosen.

Unfortunately, almost no test is perfect with complete (100%) sensitivity and specificity. The choice of what threshold is used depends on the parameters of the test and what the purpose is when using it. Deliberately setting the threshold for optimum sensitivity can result in increased numbers of false positives (above or below the threshold) as well, resulting in reduced specificity. Conversely, in other circumstances optimising specificity may be more relevant, at the cost of reduced sensitivity.

Performing several tests serially increases the overall specificity for detecting a particular disease, with each test being sequentially more specific than the previous one.

Positive predictive value

The positive predictive value is defined as the probability that a patient with a positive test result really does have the condition for which the test was requested. Unlike sensitivity and specificity which are independent of the population being tested, the positive predictive value of a test changes depending on the prevalence of the disease in the population being tested.

For example, a theoretical ELISA test for HIV may have a sensitivity and specificity of 99.9%. Among 1000 intravenous drug users with an HIV prevalence of 10%, the test will correctly detect approximately 100 (99.9) people with the disease, but incorrectly label one person (0.9)

without the disease as being HIV-positive. This is a positive predictive value of 99%.¹³ However, in a population of blood-donors (already screened for HIV) the prevalence of HIV would be much lower, closer to 0.1%.¹³ For every 1000 blood-donors screened for HIV the test would correctly detect one person (0.9) with HIV, but incorrectly label one person (0.9) as being falsely-positive for HIV. In this second population the positive predictive value of the test falls to 50%.¹³

The negative predictive value is defined as the probability that a patient with a negative test result really is free of the condition for which the test was conducted.

The probability of an abnormal result increases when the number of tests increases

The risk of a healthy individual having a result outside the reference interval increases as the number of tests selected increases. This is because the normal reference interval for most biochemical tests is defined as being two standard deviations from the mean of a healthy population.⁵ Therefore, an average of 5% of all test results from healthy patients will fall outside the normal range and be recorded as abnormal (Table 1).⁵

False-positive results are more likely when people with a low probability of a condition undergo testing. Although

false positive results can cause significant anxiety to the patient, false-negative results can often have more serious health consequences. Test results should always be interpreted in the context of other information gained from the clinical history and physical examination. Results which are borderline need to be interpreted with caution as the inter-test variability could mean the result is either normal or abnormal, so may need to be repeated after a period of time. If there is doubt, consultation with a pathologist about the test results can be helpful.

Table 1: Probability of a healthy person returning an abnormal biochemical test result, adapted from Deyo (2002)⁵

Number of tests	Probability of at least one abnormal test (%)*
1	5
6	26
12	46
20	64
100	99.4

*Assuming each test outcome is independent

An example of pre-test probability, sensitivity and specificity

A D-dimer test can be used in conjunction with the Wells Rule or Primary Care Rule to determine the probability of a patient having a deep vein thrombosis (DVT). The sensitivity of the D-dimer test is 88% and the specificity is 72%.¹² Because of the low specificity, D-dimer is most useful as a “rule-out” test for DVT, i.e. a patient with a low or normal D-dimer level, whose symptoms and signs suggests a low pre-test probability of DVT, is unlikely to have a DVT. A patient with a high pre-test probability of DVT should be referred for ultrasound irrespective of the results of the D-dimer test.

 For further information see: “The role of thrombophilia testing in general practice”, Best Tests (Mar, 2011).