RGA®

Re-flections℠

reinventing reinsurance

Volume 9, May 2002

good judgment and a wide variety of skills drawn from many areas of insurance knowledge.

Years ago, blood testing for insurance purposes was relatively rare. Generally, SMA-12's or CBC's were limited to an as-needed basis or to the very large case. Those days are long gone.

Today, only small policies are not tested in some way. Alternatives to blood and urine tests are being developed. We see and are asked to evaluate other types of tests routinely. The growing number of tests, fluids, normal ranges and testing protocols makes the underwriter's job just a little more complicated each year.

Most papers that deal with underwriting tests have dealt specifically with blood, urine, or, more recently, saliva chemistry. The following reflections apply to those as well as other types of test: electrocardiograms, echocardiograms, treadmills, pathology reports, chest x-rays and others. Each type of test, regardless of method, system, or area studied, must have some predictive value.

Is it possible to simplify, at least a little, all the various normal ranges, predictive values, and varying interpretations of findings? Not really.

Is it possible to apply certain generalities to interpreting test results? I believe so.

Is it really necessary to get an advanced degree in biochemistry, statistics, or actuarial science to get to a reasonable decision based on test analysis? Absolutely not.

So where do we start to simplify a complicated environment of tests and interpretations? How about with a look to the purpose of those tests.

The Purpose of Testing

The medical profession tests for several reasons—to investigate, to diagnose, to monitor, to screen. Each purpose requires a different perspective on the results. For example, a borderline abnormal hemoglobin A1C result in an otherwise normal 45 year-old may not be of much consequence, but clinically may indicate the need for future follow-up and more testing. In a known diabetic the same result may indicate the diabetic is in good control.

On the other hand, underwriters are generally called upon to use test results only for screening and monitoring. We are not here to investigate symptoms or findings, nor should life insurers be called on to diagnose disease. Our role is to access life risks, so the screening and monitoring functions predominate our analysis of tests.

Sensitivity and Specificity

The concepts of sensitivity and specificity are very important for an underwriter to keep in mind.

Sensitivity is the degree to which a test will pick out a true abnormality when it exists. This is the percentage of true-positive results in people with a disease. Call it the confidence level that when the test is positive, you can say yes, there is a disease process here. Specificity is the degree to which a test will exclude true-negatives—the percentage of true-negatives in a healthy population. So a good test is one that finds those with disease and excludes the healthy.

Prevalence

Another important concept is prevalence—the frequency that a given disease is encountered in the population being tested. Prevalence is the number of existing cases divided by the total population being studied.

I should also point out that there is a distinction between prevalence and incidence. Incidence is the number of new cases of a disease in a defined period of time, divided by the total at-risk population.

Predictive Value

So when is a test result truly predictive? There are actually two types of predictive value: positive-predictive value, and negative-predictive value. Hence, it depends on the circumstances.

Positive-predictive value is the confidence level that a positive test result is truly positive, and therefore indicates a disease or abnormal state exists. Negative-predictive value is the confidence level that a negative result is truly negative, and therefore indicates the absence of a disease or abnormal state.

The table shows how the various concepts interrelate.

PREDICTIVE VALUE TABLE*

	Number with positive test result	Number with negative test result	Total
Number with disease	TP	FN	TP+FN
Number without disease	FP	TN	FP+TN

TP = True positives: the number of sick subjects correctly classified by the test.
FP = False positives: the number of subjects free of the disease who are misclassified by the test.
TN = True negatives: the number of subjects free of the disease who are correctly classified by the test.
FN = False negatives: the number of sick subjects misclassified by the test.
Prevalence = Percent of total subjects examined who are diseased.

$$\text{Sensitivity = positivity in disease} = \frac{TP}{TP+FN} \times 100 = \frac{TP}{\text{No. diseased}} \times 100$$

$$\text{Specificity = negativity in health} = \frac{TN}{TN+FP} \times 100 = \frac{TN}{\text{No. without disease}} \times 100$$

$$\text{Predictive value of a positive test} = \frac{TP}{TP+FP} \times 100 = \frac{TP}{\text{No. positive}} \times 100$$

$$\text{Predictive value of a negative test} = \frac{TN}{TN+FN} \times 100 = \frac{TP}{\text{No. negative}} \times 100$$

* *From Galen RS, Gambino SR:* Beyond Normality: The Predictive Value and Efficiency of Medical Diagnoses. *New York, John Wiley & Sons, Inc. , p. 124.*

COMMON CAUSES OF ERROR

LABORATORY TESTING: EIGHT POINTS TO REMEMBER

1. Laboratory reference ranges are generally established to include 95 percent of the normal population. That means that five percent of the population may have abnormal test results, even when nothing is wrong. In other words, if 20 different tests are run on a healthy individual, there is good possibility that at least one will fall outside the reference range.

2. Reference laboratory values vary widely due to testing methods. To know if a test is normal or abnormal you must have the reporting laboratory's reference values.

3. The further lab values are from their normal limits, the greater the likelihood of a true abnormality.

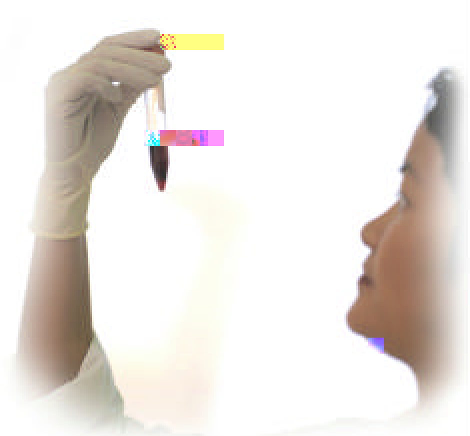
4. If diagnostically related tests are simultaneously abnormal, the probability is greater that a true abnormality is present.

5. If a current test reconfirms a historic test abnormality, the likelihood is greater that a true abnormality is present.

6. If you have not checked for possible prescription drug use and multiple drug interactions, mildly abnormal findings are of questionable value.

7. Use caution when referring to laboratory values from hospital reports. Patients in the hospital are usually in a state of physical (and often emotional) stress. Transient states such as: medication, trauma, diet, hydration, and previous tests done on the patient, can all effect the blood and urine chemistry.

8. Transient states don't just occur in the hospital. Other transient states include: dehydration, fasting and dieting, heavy carbohydrate intake, mild infections, muscle trauma, physical stress, emotional stress, and prescription drugs.



AVOID PRECONCEPTIONS

Good investigative skills include blinding oneself to preconceptions.

"Which is it to-day," I asked, "morphine or cocaine?" He raised his eyes languidly from the old black-letter volume which he had opened.

"It is cocaine," he said, "a seven per-cent solution. Would you care to try it?"

"No, indeed," I answered brusquely. "My constitution has not got over the ... campaign yet. I cannot afford to throw any extra strain upon it."

He smiled at my vehemence. "Perhaps you are right, Watson," he said. "I suppose that its influence is physically a bad one. I find it, however, so transcendently stimulating and clarifying to the mind that its secondary action is a matter of small moment."

"But consider!" I said earnestly. "Count the cost! Your brain may, as you say, be roused and excited, but it is a pathological and morbid process which involves increased tissue-change and may at least leave a permanent weakness. You know, too, what a black reaction comes upon you. Surely the game is hardly worth the candle. Why should you, for a mere passing pleasure, risk the loss of those great powers with which you have been endowed? Remember that I speak not only as one comrade to another but as a medical man to one for whose constitution he is to some extent answerable."

-- *Dr. Watson and Sherlock Holmes speaking in Sir Arthur Conan Doyle's* <u>The Sign of Four</u>

>>>

Some Commonly Encountered Preconceptions

> Problem drinkers will have elevated liver-function studies, and/or “alcohol” markers.

> There isn’t a significant alcohol problem if the liver-function studies and/or “alcohol” markers are normal.

> It’s best practice to average all the test results.

> The St. Louis Rams will win the Super Bowl. (Apologies to our international readers who may not follow American Football.)

IMPORTANT GENERALITIES FOR INTERPRETING LABORATORY RESULTS

“It is an old maxim of mine that when you have excluded the impossible, whatever remains, however improbable, must be the truth.”

-- *Sherlock Holmes, Sir Arthur Conan Doyle’s* The Beryl Coronet

Bayes’ Theorem & Bayes’ Rule

Bayes’ Theorem is the model for predictive value in the science of probability.

Simply put, any serious mathematician is going to have some objection to Sherlock Holmes’ over-simplification. Bayes’ Theorem, in an underwriting context, states that a good test result (high sensitivity and high specificity) is easier to obtain when the population has a high prevalence of the disease. For example, screening a general population of all males over the age of 20 for prostate cancer with PSA’s would not be as useful as only screening males over 50. The prevalence of prostate cancer for the general adult male population is significantly lower than a population specifically targeting older males. The probability that a positive test result in a younger male is a false-positive (it has a low positive-predictive value) is much higher than a positive result in an older male.

The formula used is known as Bayes’ Rule.

$$\text{Predictive Value} = \frac{\text{Prevalence (sensitivity)}}{[(\text{prevalence})(\text{sensitivity})+(1-\text{prevalence})(1-\text{specificity})]}$$

In the equation, ‘prevalence’ refers to the known or assumed degree that the disease exists in the population studied. For example, liver-function studies might be more prevalent in a population of treated alcoholics, than in a general insured population. One needs to know what group is being studied. ‘Sensitivity’ was previously defined as the degree to which a test will pick out a true abnormality. Specificity has also previously been defined as the degree to which a test will exclude true-negatives.

Both Rule

If two identical tests are performed, the first being abnormal, the second being normal, use the second result. There is a mathematical reason for this that goes far beyond the scope of this paper, but the reasoning is strong.

Coherence

If several similar tests are associated and related to the condition being studied, and they all are abnormal, there is a much greater chance that a true abnormality exists. Therefore, if you're analyzing liver-function tests, check them all. If more than one is out of range, your chances for something being a true abnormality increase significantly.

Regression Toward the Mean, The Central Tendency Theorem

Multiple similar tests, done over time, will tend toward either a normal result, or an abnormal one. Watch the trends. It doesn't matter as much that the most recent test dips slightly toward normal if the general trend has been toward abnormality. More likely, you will find that an abnormal result is returning to normal, but it can be helpful to write down results in date order to see if a trend exists. It also belies the concept of simple numeric, arithmetic averages (e.g. arithmetic means). They don't tell the whole story, and have lead many a detective down an alley without a magnifying glass.

SUMMARY

Given all the disclaimers regarding test results it's surprising that underwriters have come to rely on them so heav-