UNIVERSITY OF TEXAS
HEALTH SCIENCE CENTER AT DALLAS
SOUTHWESTERN MEDICAL SCHOOL

MEDICAL GROUND ROUNDS

April 14, 1977

# CARDIAC DIAGNOSIS

Kirk Lipscomb, M.D.

*"Medicine is a science of uncertainty and an art of probability."*

*Sir William Osler*

TABLE OF CONTENTS

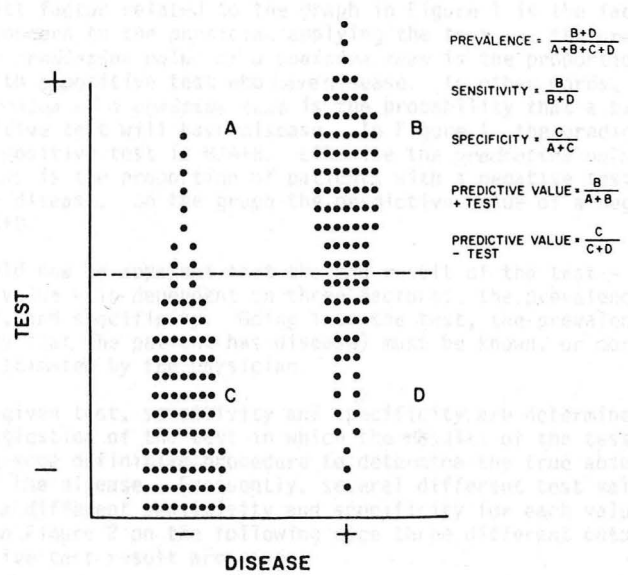
## INTRODUCTION

Diagnosis is a complex process that is poorly understood. It is largely an intuitive process and it is likely that we will never understand what goes on in the mind of a good diagnostician. There are however some concrete principles that are basic to the proper evaluation and use of diagnostic tests. Since the diagnostic process in cardiology now uses many tests, it is imperative that the physician interested in cardiology both understand the basic principles of testing and the tests used. Today I will begin by discussing some basic principles of testing and then apply these principles to some tests used in cardiology.

## PRINCIPLES OF TESTING

I will begin with an example. In the graph in Figure 1 below, points are plotted according to whether the patient does or does not have disease, and according to the value of some test designed to diagnose the disease.

Figure 1



The *prevalence* of disease is the proportion of persons in the population who have disease. In Figure 1, prevalence is B+D/A+B+C+D. The probability that a single person in the population has disease is equal to the prevalence of disease for the whole population. In essence to say that a person has $x$ probability of disease is to say that if there were many identical person, the prevalence would be $x$. For example, if in a population of 100 people, 30 have disease, the prevalence is .30. Likewise the probability that any single one of the 100 has disease is .30.

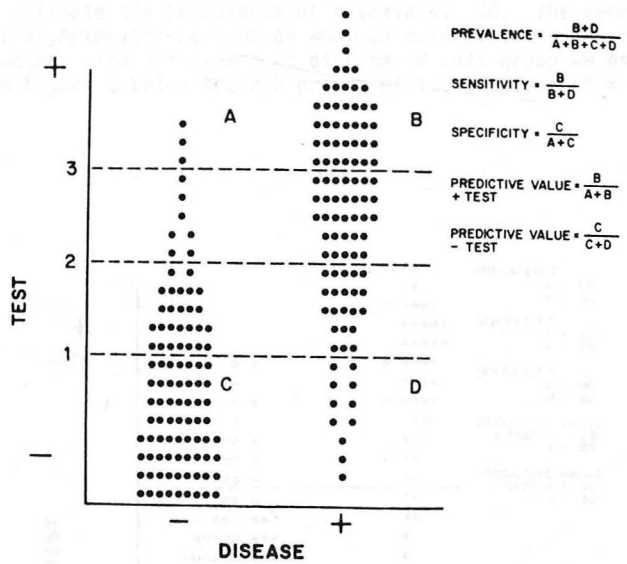
If a test value such as the horizontal dotted line in Figure 1 is then chosen, it will divide both the diseased and non-diseased patients into two groups each depending on whether they have a positive or negative test. There is now a total of 4 groups in a matrix consisting of patients with and without disease, and with and without a positive test. On the graph these groups are labeled A, B, C, and D. The ratio determined by the test result in patients *with the disease* is known as the *sensitivity*, and the ratio determined by the test result in patients *without the disease* is known as the *specificity*. *Sensitivity* is the proportion of patients with disease who have a positive test. On the graph sensitivity is B/B+D. *Specificity* is the proportion of patients without disease who have a negative test. On the graph specificity is C/A+C. Remember --- *sensitivity* deals only with people who have the disease and *specificity* deals only with people who do not have the disease. As with prevalence, sensitivity and specificity refer to the probabilities in individual patients as well as the ratios of populations.

The last factor related to the graph in Figure 1 is the factor of immediate concern to the physician applying the test --- the *predictive value*. The *predictive value of a positive test* is the proportion of patients with a positive test who have disease. In other words, the *predictive value of a positive test* is the probability that a patient with a positive test will have disease. In Figure 1, the predictive value of a positive test is B/A+B. Likewise the *predictive value of a negative test* is the proportion of patients with a negative test who do not have disease. On the graph the predictive value of a negative test is C/C+D.

It should now be apparent that the end result of the test - the predictive value - is dependent on three factors: the prevalence, sensitivity, and specificity. Going into the test, the prevalence (probability that the patient has disease) must be known, or more commonly, estimated by the physician.

For a given test, sensitivity and specificity are determined by prior investigation of the test in which the results of the test are compared to some definitive procedure to determine the true absence or presence of the disease. Frequently, several different test values are given with a different sensitivity and specificity for each value. In the graph in Figure 2 on the following page three different cutoff values for a positive test result are given.

Figure 2



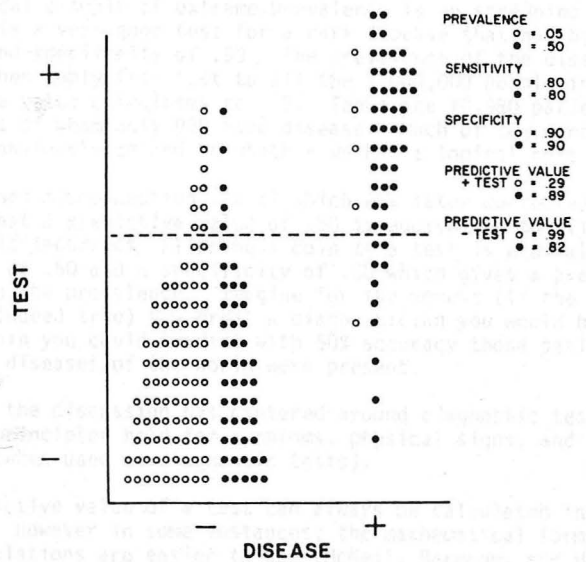

At test value 1, sensitivity is high, but specificity is low. At test value 2, sensitivity is decreased, but specificity is increased. At test value 3, sensitivity is quite low, but specificity is very high. In general high sensitivity is most desirable when the benefit of treating diseased patients is high and the cost of inappropriately treating non-diseased patients is low, such as in tuberculosis. On the other hand high specificity is most desirable when the benefit of treating diseased patients is low and the cost of inappropriately treating non-diseased patients is high, such as in lung cancer.

In essence, the sequence in using the testing procedure outlined is to estimate the probability that the patient has the disease (prevalence), then utilize a test with known sensitivity and specificity to derive a new probability (predictive value).

The impact of sensitivity and specificity on the outcome of a test is generally better appreciated than the impact of prevalence. Yet, prevalence is clearly just as important as specificity or sensitivity. An example will illustrate this. Assume that we have a new test to diagnose coronary artery disease, which to be in vogue, we will call the echoisotopogram. We know that the sensitivity of this test is .80 and the specificity is .90. We will now use this test to predict the presence of coronary disease in 2 groups of 100 people each. The first group consists

of patients admitted to the hospital for the evaluation of chest pain in whom we estimate the prevalence of disease as .50. The second group consists of asymptomatic persons on whom an epidemiologic survey is being conducted. The prevalence of disease in this group we estimate at .05. In Figure 3 below these 2 groups of 100 people each are shown.

Figure 3




The closed circles represent the chest pain patients of whom 50 have disease and 50 do not have disease. The open circles represent the asymptomatic persons of whom 5 have disease and 95 do not have disease. Since the sensitivity is .80, 40 of the 50 chest pain patients with disease will have a positive test and 10 will have a negative test. Of the 5 asymptomatic persons with disease, 4 will have a positive test and 1 will have a negative test. Since specificity is .90, 45 of the 50 chest pain patients who do not have disease will have a negative test and 5 will have a positive test. Of the 95 asymptomatic persons who do not have disease, 85 will have a negative test and 10 will have a positive test. To determine the predictive value of a positive test we now take

the patients with a positive test and see how many have disease. In the chest pain patients (closed circles) 45 patients have a positive test of whom 40 have disease. This gives a predictive value of 40/45 or .89. In the asymptomatic persons 14 have a positive test of whom 4 have disease. This gives a predictive value of 4/14 or .29. Same sensitivity - same specificity - yet a markedly different predictive value.

Two absurd but illustrative examples of the effect of prevalence in extreme situations are that the predictive value of a positive test applied to a totally disease-free population is 0.00 while the predictive value of a positive test applied to a totally diseased population is 1.00, regardless of the sensitivity and specificity (as long as they are not 0.00). A more practical example of extreme prevalence is in screening tests. Assume there is a very good test for a rare disease that has both a sensitivity and specificity of .99. The prevalence of the disease is .001 and we then apply this test to all the 1,000,000 people in Dallas. The predictive value calculates to .09. There are 10,980 patients with positive tests of whom only 990 have disease. Much of the worth of the effort would obviously depend on whether we had a logical next step.

A published misconception (Katz) which was later corrected (Galen-Gambino) is that a predictive value of .50 is equivalent to flipping a coin. This is incorrect. Flipping a coin as a test is equivalent to a sensitivity of .50 and a specificity of .50 which gives a predictive value equal to the prevalence. Imagine for the moment (if the misconception were indeed true) how great a diagnostician you would be if armed only with a coin you could predict with 50% accuracy those patients in whom the rare diseases of the world were present.

Although the discussion has centered around diagnostic tests, the same general principles hold for symptoms, physical signs, and even for risk factors (when used as diagnostic tests).

The predictive value of a test can always be calculated in the manner shown. However in some instances, the mathematical formulations of these calculations are easier to use (McNeil, Barnoon, and Wolfe). These are shown below.

$$P(D+|T+) = \frac{P(T+|D+)\,P(D+)}{P(T+|D+)\,P(D+) + [1-P(T-|D-)]\,[1-P(D+)]}$$

$$P(D-|T-) = \frac{P(T-|D-)\,[1-P(D+)]}{P(T-|D-)\,[1-P(D+)] + P(D+)\,[1-P(T+|D+)]}$$

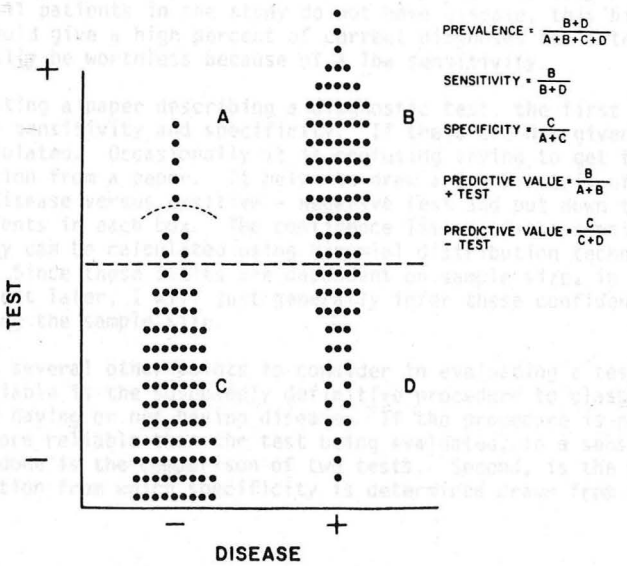
The symbols read as would be stated. P (D+|T+) is the probability of having disease when the test is positive (predictive value of a positive test). P (T+|D+) is the probability of having a positive test when the disease is present (sensitivity). P (D+) is the probability of disease before the test is performed (prevalence). P (T-|D-) is the probability of having a negative test when disease is absent (specificity). P (D-|T-) is the probability of no disease when the test is negative (predictive value of a negative test). Similar formulas for P (D-|T+) and P (D+|T-) can be written using the procedure outlined in the text.

The approach to diagnosis outlined here is based on Bayes Theorem, which was first published in 1763 but has only been applied to medical diagnosis since the early 1960's. The diagnostic approach based on Bayes Theorem differs somewhat from the diagnostic approach based on classical statistics (Cornfield, Edwards). While classical statistics emphasizes acceptance or rejection of an hypothesis (or diagnosis), Bayesian statistics emphasizes probabilities. After evaluating a patient with chest pain, the final diagnosis based on the classical approach would be that the patient does (or does not) have coronary artery disease. The final diagnosis based on the Bayesian approach would be that the patient has a certain probability of having coronary artery disease. A blood test result based on the classical approach would be either normal or abnormal. A blood test based on the Bayesian approach would (in conjunction with the prevalence of the disease in question) indicate a certain probability that the patient had disease. This emphasis on probability has utility in deciding courses of action. For instance, if after all tests are done, the diagnosis is not certain but is estimated at a certain probability, this probability in conjunction with the costs and benefits of the different available courses of action, can be used to determine the most appropriate course of action (Pauker, Pauker, Barnoon and Wolfe).

The Bayesian approach is not universally accepted as a scientific tool. The mathematical formulation is undisputed, but the need to *estimate probability* (in our case - prevalence) is the chief controversial issue. The critics of the approach argue that the estimation of probability is imprecise and therefore the approach lacks scientific value, while proponents of the Bayesian approach argue that the techniques used by the critics to avoid estimating probability are more involved and less valid than what they are trying to avoid (Cornfield). The fact of the matter is that the Bayesian approach is either directly used or its principles are involved in almost all tests we use today. Concerning the estimation of probability (or prevalence), the essential point to remember is that *the predictive value of the test you are using is dependent on the probability of disease that you personally assign to the patient before the test.* Fortunately small differences in probability are not critical and so a "ball park" estimate is usually sufficient (Lusted, 1968).

Another potential problem with the Bayesian approach is present when several tests are performed on the same patient. If the tests are independent of each other -- that is, the results of one test do not influence the sensitivity or specificity of the other tests -- they can simply be performed in series. The predictive value of the first test becomes the prevalence for the next test, etc. If on the other hand, the tests are dependent on one another, a more complex situation arises. For each different result of the previous test, a different set of criteria must be used for the following test. An example of this problem is the long standing question of whether EKG T wave changes are independent of, or dependent on, age. If these two tests for coronary disease, age and T wave change, are *independent* of each other than one sensitivity and specificity for a given T wave change is applicable for patients of all ages. If on the other hand, the T wave is *dependent* on age, then different sensitivities and specificities must be used for different age groups. In general tests are considered independent unless otherwise stated. However, dependence is a common problem and one should be alert to the possibility of its influence on a test (Barnoon and Wolfe, Lusted, Hall).

For perspective and a quick review, I will compare the Bayesian approach that we have been discussing with some other ways that tests are presented. See Figure 4 below.



One common way to present a test is just to give the "normal limits" of the test, usually calculated by doing the test in a series of normal people and calculating the mean ± 2 standard deviations. In Figure 4 this normal limit is shown by the parenthesis in the patients with no disease. It is apparent that this "normal limit" does not take into account either the distribution of the test in the diseased patients or the prevalence of the disease in the population under study. Consider how different the worth of the test would be if the patients with disease had test values distributed similarly to the patients without disease versus the patients with disease having test values much higher than the patients without disease. Yet this difference in worth of the test is not reflected in the "normal limits". Sometimes the predictive value of a test is given alone without giving sensitivity or specificity. At first glance this seems acceptable because after all predictive value is the end result that you calculate from knowing sensitivity, specificity, and prevalence. But the catch with just knowing predictive value is that it is only applicable to the unique prevalence in the situation in which the test was evaluated. Consider a test that was evaluated in a population where the great majority of the patients had disease. If sensitivity and specificity are at all reasonable, the predictive value of a positive test would be excellent, but not necessarily because of a good test, but because the "deck was stacked". Another way to rate a test is to determine the percent of the total diagnoses that are correct. Again this can be deceiving. For instance consider a test evaluated in a population where almost all of the patients do not have disease. The cutoff value for the test is set high so that almost all patients have a negative test. Sensitivity would probably be very low, but specificity would be very high. Since almost all patients in the study do not have disease, this high specificity would give a high percent of correct diagnoses for a test that may actually be worthless because of a low sensitivity.

In evaluating a paper describing a diagnostic test, the first step is to find the sensitivity and specificity. If these are not given they should be calculated. Occasionally it is confusing trying to get this basic information from a paper. It helps to draw a 2 x 2 matrix of disease - nondisease versus positive - negative test and put down the number of patients in each box. The confidence limits of the sensitivity and specificity can be calculated using binomial distribution technique (Zar, Galen). Since these limits are dependent on sample size, in the studies I present later, I will just generally infer these confidence limits by giving the sample size.

There are several other points to consider in evaluating a test. First, how reliable is the supposedly definitive procedure to classify the patient as having or not having disease. If the procedure is not considerably more reliable than the test being evaluated, in a sense all that is being done is the comparison of two tests. Second, is the non-disease population from which specificity is determined drawn from a

population similar to that which will be encountered in the clinical situation? An example is in the evaluation of an enzyme test for myocardial infarction. If the specificity is determined from a group of medical students, it will probably be far better than the specificity determined from hospital patients with diseases likely to be confused with myocardial infarction. In essence this is an example of the effect of dependence. Third, when the test was evaluated, was it related in any way to the procedure used in final diagnosis? An obvious example of this would be if the test being evaluated was used in making the definitive diagnosis. A more subtle violation of this principle probably occurs more commonly when the test being evaluated is used to select patients to undergo the definitive procedure. Consider the results obtained if we decided to do a retrospective study on the sensitivity and specificity of exercise testing in the diagnosis of coronary artery disease. But in the period of time in which the patients were hospitalized we already believed in the test to the point that we were only referring for the definitive test of catheterization those patients who had a positive exercise test. Obviously our results would show artefactually high sensitivity and low specificity.

For further reading, the July 31, 1975 issue of the New England Journal of Medicine is excellent, as is the article by Schwartz and associates. The books by Galen and Gambino, Barnoon and Wolfe, and Lusted are good, but more involved. The reviews of the mathematics of Bayes Theorem by Cornfield, and by Edwards and associates, are both readable and excellent.

## TESTS

With this background in the principles of testing I will now discuss some tests commonly used in cardiology. I won't discuss tests that are promising but not yet well evaluated, but rather confine the discussion to those tests where enough data is available to make some judgement about the value of the test. In general I picked the tests because they are commonly used, have potential for misinterpretation, or they illustrate a point about testing.

### CORONARY ARTERY DISEASE:

Exercise testing. Exercise testing on the treadmill or bicycle is now common in cardiology. There are several purposes of this test, one of the most common of which is to attempt to diagnose the presence or absence of coronary artery disease. Interest in this diagnostic application is now high because of the sheer number of patients in whom diagnosis is difficult and because of the hope that early therapeutic intervention may be beneficial.

The values of sensitivity and specificity for maximal or near maximal exercise testing in the diagnosis of coronary artery disease are shown below in Table 1.

Table 1

| Study | Patients with Disease | Patients without Disease | Sensitivity | Specificity |
|---|---|---|---|---|
| Roitman, 1970 | 30 | 16 | .80 | .88 |
| Martin, 1972 | 63 | 37 | .62 | .89 |
| Keleman, 1973 | 48 | 24 | .54 | .96 |
| Bartel, 1974 | 258 | 123 | .64 | .91 |
| Linhart, 1974 | 71 | 47 | .80 | .89 |
| Rios, 1974 | 29 | 21 | .83 | .90 |
| Piessens, 1974 | 40 | 30 | .65 | .83 |
| Borer, 1975 | 39 | 11 | .33 | .91 |
| Goldshlager, 1976 | 269 | 141 | .64 | .93 |
| Approximate Mean | | | .70 | .90 |

In these studies, the cutoff point for a positive test was 1.0 mm ST depression. The groups studied were generally patients referred for diagnostic or presurgical evaluation who were considered ill enough to undergo catheterization. The standard for absolute diagnosis was coronary angiography. Patients with coexisting cardiac disease, conduction disorders on the EKG, and digitalis were usually excluded from the study while patients with previous infarcts and/or ST changes were generally included. As is evident from the table sensitivity is about .70 and specificity about .90, the figures most quoted by cardiologists. One study which was published in the New England Journal of Medicine in 1975 (Borer) is notable for its deviation in sensitivity - only .33. This study has been discussed rather widely and the reason for the deviation is not immediately apparent. The authors used the term specificity in an unconventional manner which critics were quick to point out, but the raw data shows that the sensitivity is reported correctly. Among the possibilities for the deviation is the fact that the authors apparently were biased against the value of the test (Redwood) relative to the generally optimistic approach of the other authors. Since the inter-

pretation of EKG changes is subjective, this bias may have had an important influence.

Using the commonly accepted sensitivity of .70 and specificity of .90, Table 2 below shows the expected predictive value of a positive and negative test at different levels of prevalence.

Table 2

| Prevalence | Predictive Value Positive Test | Predictive Value Negative Test |
|---|---|---|
| .01 | .07 | .997 |
| .10 | .44 | .96 |
| .50 | .88 | .75 |

When the prevalence (or probability that the patient has the disease before the test) is .01, the predictive value of a positive test (that is, the chance that a patient with a positive test has the disease) is .07, and the predictive value of a negative test (that is, the chance that a patient with a negative test does not have the disease) is .997. When prevalence is .10, the predictive value of a positive test is .44 and the predictive value of a negative test is .96. When prevalence is .50, the predictive value of a positive test is .88 and the predictive value of a negative test is .75. The predictive value of a prevalence of .01 suggests that screening asymptomatic individuals by exercise testing may be of little value.

Two studies bear on this question of the value of screening asymptomatic individuals (Erickssen, Froelicher). In these two studies presumably healthy middle-aged men were screened by exercise testing. Approximate calculations from one of these studies show the prevalence of coronary disease in this presumably healthy population was about .04. At this prevalence of .04, and using the sensitivity of .70 and specificity of .90 derived from studying symptomatic patients the predictive value of a positive test calculated to .23. In these two studies patients with a positive exercise test were studied by angiography with the result that 66% and 53% of the individuals had significant disease. The reason for this better predictive value than anticipated is apparently a specificity higher than .90 in asymptomatic individuals. This points out the danger in assuming that a test is independent (that is, that the sensitivity and specificity hold for all conditions). In the example just cited the specificity is dependent on the population studied. Another point is evident from this example. This is that in screening low prevalence populations, a lot of individuals who have positive tests are going to be free of disease. This fact must be taken into account lest more harm than good be done.

In these studies just discussed the cutoff point for a positive test was 1.0 mm ST depression. In Table 3 below the effect of using lesser and greater amounts of ST depression as cutoff points is shown.

Table 3

| Study | Sensitivity at ST Δ of .5 | 1.0 | 2.0 | Specificity at ST Δ of .5 | 1.0 | 2.0 |
|---|---|---|---|---|---|---|
| Bartel, 1974 | | .64 | .32 | | .91 | .99 |
| Keleman, 1973 | | .54 | .27 | | .96 | 1.00 |
| Piessens, 1974 | .70 | .65 | | .67 | .83 | |
| Martin, 1972 | .84 | .62 | .38 | .57 | .89 | 1.00 |
| Approximate Mean | .75 | .60 | .33 | .60 | .90 | 1.00 |

As expected, at 0.5 mm ST depression sensitivity is highest at .75 and specificity is lowest at .60. At 2.0 mm depression, sensitivity is lowest at .33 but specificity is very high at virtually 1.00. The amount of ST depression chosen for the cutoff point varies with the purpose of the test. From the figures it is apparent that 1.0 mm is a compromise between sensitivity and specificity and therefore is the value most often used. However, it should be apparent that if a patient has 2.0 mm ST depression he is virtually certain to have disease, unless his probability of having the disease before the test is very low.

Since longevity in patients with left main coronary disease has now been shown to be increased by surgery (Takaro) and patients with left main coronary disease are known to show deep ST depression during exercise more commonly then other patients; exercise testing has been used to screen for this condition. In Table 4 below the sensitivity and specificity of 2.0 mm ST depression in detecting left main coronary disease is shown. The prevalence of left main disease in the population of coronary disease patients is also listed.

Table 4

| Study | Patients with Disease | Patients without Disease | Sensitivity | Specificity | Prevalence |
|---|---|---|---|---|---|
| Keleman, 1973 | 6 | 42 | .83 | .81 | .12 |
| Bartel, 1974 | 31 | 227 | .52 | .70 | .12 |
| Khaja, 1974 | 18 | | .78 | | .04 |
| Cheitlin, 1975 | 11 | 95 | 1.00 | .64 | .10 |
| Cohen, 1975 | 42 | | .81 | | .04 |
| Approximate Mean | | | .80 | .70 | .08 |

Sensitivity is about .80, specificity is about .70, and prevalence is about .08. Using these figures the predictive value of 2.0 mm ST depression in detecting left main disease is about .20 and the predictive value of a negative test is .98, as shown below in Table 5.

Table 5

| Prevalence | Sensitivity | Specificity | Predictive Value Positive Test | Predictive Value Negative Test |
|---|---|---|---|---|
| .08 | .80 | .70 | .19 | .98 |

For this approach to be valid it is of course necessary to allow the test to proceed long enough for the deep ST depression to be manifest. It is the policy in some exercise labs including those at Parkland and the VA to terminate the test after the ST segments have dropped 1.0 mm.

When patients with other cardiac diseases, patients taking digitalis, and/or patients with resting EKG abnormalities more severe than just ST change or infarct pattern comprise the studied population sensitivity stays the same but specificity drops to about .75 as shown in Table 6 below.

Table 6

| Study | Patients with Disease | Patients without Disease | Sensitivity | Specificity |
|---|---|---|---|---|
| Roitman, 1970 | 32 | 22 | .66 | .77 |
| Linhart, 1974 | 33 | 27 | .55 | .63 |
| Kansal, 1976 | 26 | 11 | .88 | .82 |
| Mean | | | .70 | .74 |

In these studies 1.0 mm ST depression is the cutoff point and any angiographically demonstrated coronary artery disease over 50% obstruction is the standard. It is generally felt that the sensitivity and specificity of exercise testing in women are lower than in men (Blomqvist, Bruce, Sketch) however at least one report shows no difference (Linhart 1974).

Ear lobe crease sign. In 1973, Frank suggested that a diagonal crease across the ear lobe was more common in patients with coronary artery disease and that this crease may be useful in diagnosis. Several studies of this relationship are shown in Table 7 below.

Table 7

| Study | Standard | Sensitivity | Specificity |
|---|---|---|---|
| Lichstein, 1974 | MI | 251/531 (.47) | 213/305 (.70) |
| Mehta, 1974 | Angiography | 89/159 (.56) | 26/52 (.50) |
| Christiansen, 1975 | All Methods of Diagnosis | 81/176 (.46) | 241/300 (.80) |
| Mean | | .50 | .67 |

Different methods were used to ascertain the presence of coronary disease, but the final sensitivity and specificity are reasonably close and average out to about .50 sensitivity and .67 specificity. The value of the test has been questioned (Mehta). The reply by the enthusiasts of the test center around the statistically significant difference of the frequency of the positive sign in coronary and non-coronary disease patients. But this difference in frequency does not tell the clinician how helpful the sign is. To tell how helpful the sign is, the predictive value should be calculated. This is done in Table 8 below.

Table 8

| Prevalence | Predictive Value Positive Test | Predictive Value Negative Test |
|---|---|---|
| .10 | .14 | .92 |
| .50 | .60 | .57 |

As can be seen from the table, the ear lobe crease sign probably has diagnostic value but this value is minimal.

Total Cholesterol. A risk factor is generally considered in terms of how much the factor increases a patient's chance of having the disease. However, these factors are also frequently used in the process of diagnosis. To understand how helpful the risk factor is in diagnosis, it can be analyzed like a test. A case in point is the total serum cholesterol. In Table 9 below 2 studies are evaluated in which total cholesterol in a series of patients being evaluated for symptoms of coronary artery disease is compared to the angiographic presence of obstructive coronary disease.

Table 9

| Study | Sensitivity at Cutoff 225 | 280 | 315 | Specificity at Cutoff 225 | 280 | 315 |
|---|---|---|---|---|---|---|
| Cramer, 1966 | .76 | .47 | .27 | .37 | .63 | .84 |
| Ascoop, 1971 | | .52 | | | .76 | |
| Approximate Mean | | .50 | | | .70 | |

There were 135 patients in the Cramer study and 96 patients in the Ascoop study. It is interesting that 7 studies of the correlation of lipids and coronary angiography were screened before this simple information could be extracted. Using the approximate sensitivity of .50 and the approximate specificity of .70 for the cutoff cholesterol value of 280, the predictive value of the total cholesterol is shown below in Table 10.

Table 10

| Prevalence | Predictive Value of Positive Test | Predictive Value of Negative Test |
|---|---|---|
| .10 | .16 | .93 |
| .50 | .62 | .58 |

It is obvious that the cholesterol value alone has some, but little, value.

MYOCARDIAL INFARCTION:

Technetium - 99 m Stannous Pyrophosphate Myocardial Scanning. Since Dr. Bonte and his colleagues demonstrated the potential of TC-PYP scanning for myocardial infarct diagnosis in 1973, (Bonte) the procedure has grown to be accepted as a valuable tool by most cardiologists. Much of this acceptance is due to the careful and extensive work of Dr. Willerson

and his colleagues in delineating the proper use of this test. About 12 hours after the infarct, the infarcted tissue takes up the TC-PYP such that a positive image is formed. The image becomes more positive for about 3 days and starts to fade at about 7 days. Probably only irreversibly damaged myocardium becomes positive, but there is still question as to whether severely ischemic yet still viable tissue also becomes positive. A significant difficulty in evaluating the scan is the lack of a good diagnostic standard. Because of this the standard is usually the EKG and enzymes, techniques well recognized to be imperfect and which in fact may well be less sensitive than the test they are evaluating -- the scan.

All studies indicate that the sensitivity of the scan, when done 1-7 days post infarct, is very good to perfect. The rare false negative scan seems to occur in subendocardial infarction (Willerson, McLaughlin, Prasquier, Ahmad). There is less unanimity however in the specificity, i.e. how often the scan is negative in the absence of infarct. This difference in specificity is listed below in Table 11.

Table 11

| Study | Population of Non-Infarct Patients | Positive Scans Number Patients | Specificity |
|---|---|---|---|
| Willerson, 1975a | R/O MI | 0/71 | 1.00 |
| Willerson, 1975b | R/O MI | 9/101 | .91 |
| McLaughlin, 1975 | R/O MI | 1/17 | .94 |
| Prasquier, 1977 | Bone Scans (non-cardiac indication) | 70/483 | .86 |
| | Unstable Angina | 8/18 | .56 |
| | Stable Angina | 4/41 | .90 |
| Ahmad, 1977 | Spectrum of Cardiac Patients | 31/67 | .54 |
| | Stable Angina | 7/30 | .77 |
| | Unstable Angina | 3/3 | .00 |
| | Chronic Aneurysm | 9/10 | .10 |
| | Post-op Coronary Surgery | 7/13 | .46 |
| | Cardiomyopathy | 4/4 | .00 |
| | Post-Cardioversion | 1/1 | .00 |
| | Acute Pericarditis | 0/2 | 1.00 |
| | Atypical Pain | 0/4 | 1.00 |

In the study by Prasquier 70 patients had positive scans out of 483 patients studied by whole body imaging for non-cardiac conditions, usually malignancy. In this study those patients with a positive scan usually had visualizable femoral arteries implying that the false positivity was due to visualization of the ventricular blood pool. In 120 patients that they then studied for cardiac indications in which they took multiple views of the heart their data is not clear, but calculation of their raw data implies that ventricular pool imaging was not a problem in these specific myocardial scans. This recalculated data is similar to that in the study by Ahmad.

In the study by Ahmad, the non-infarct patients were selected from patients admitted to the hospital with cardiac disease. In essence the breakdown of the patients in this study list the conditions where a false positive scan may be expected. The positive scans in the stable and unstable angina group have been noted by other investigators such as the study by Willerson. It is not clear whether these are truly false positives or just show that the scan is more sensitive for infarct than the objective standards used to evaluate it, i.e. infarction is occuring but the scan is the only test to become positive. Aneurysm, cardiomyopahty, and post cardioversion are all conditions where necrosis has been shown or could be expected. The post operative coronary surgery patients were studied an average of 30 months post-surgery. Scar and possible calcium deposition were postulated as reasons for the positive scan. Abnormal valvular calcium was shown to pick up TC-PYP in another study (O'Rourke). From this resume of the conditions causing false positive scans, it is obvious that the specificity chosen in figuring the predictive value of the scan depends on the population from which the infarct patients are to be picked. If the population is one of simple chest pain specificity is high and predictive value good. However, if one of the conditions listed with a high false positive rate is felt to be present in the population tested, specificity drops accordingly.

Serum creatine phosphokinase -MB. The diagnosis of acute myocardial infarction has characteristically relied heavily on the electrocardiogram and serum cardiac enzymes. The development of a characteristic EKG pattern of transmural infarction is a highly specific but not very sensitive test for infarction, while a rise in serum enzymes is a very sensitive yet not very specific test. Consequently little difficulty is encountered when the EKG shows a transmural infarction or conversely when the enzymes are normal. The problem comes when the EKG is not classic for infarction and enzymes are elevated. Does the enzyme elevation represent infarction or is it a false positive? Creatine phosphokinase (CPK) is an enzyme found in cardiac muscle which is released after infarction. However, CPK is also found in other tissues making an elevation in serum not very specific. CPK-MB is the isoenzyme found only in significant amounts in the myocardium. Its elevation in serum therefore, should be both sensitive and specific. Evaluation of the elevation of CPK-MB as a test for infarction is fraught with the same difficulty as evaluation of the TC-PYP scan, that is the lack of a good objective standard to judge the presence of infarction. All of the clinical studies evaluating CPK-MB rely on EKG and clinical judgement to decide whether infarction was indeed present. The enzyme begins to rise 2-12 hours after infarction, peaks in the first

day, and begins to fall by 24-72 hours. The test is very sensitive detecting almost all cases of infarction. In addition however it is much more specific than prior enzyme tests as shown in Table 12 below.

Table 12

| Study | Characteristics of Patients with No Infarct | Positive Test No. of Patients | Specificity |
|---|---|---|---|
| Wagner, 1973 | R/O MI | 1/182 | .99 |
| Varat, 1975 | R/O MI | 4/53 | .92 |
| Blomberg, 1975 | R/O MI | 0/143 | 1.00 |
| Galen, 1975 | R/O MI | 7/48 | .85 |
| Roberts, 1975 | R/O MI | 0/50 | 1.00 |
| | Non-cardiac surgery | 0/100 | 1.00 |
| | Cardiac cath | 0/50 | 1.00 |
| Ehsani, 1976 | Post-cardioversion | 2/30 | .93 |
| Coleman, 1976 | Cardiac surgery | 22/23 | .04 |
| Tonkin, 1975 | Selected patients with minor cardiac trauma | 4/4 | .00 |
| Konttiner, 1973 | Muscular dystrophy (Duchenne) | high | low |

Of pertinence is the fact that in the presence of cardioversion, cardiac surgery, minor cardiac trauma such as cardiac massage, and muscular dystrophy, it loses its specificity.

Electrocardiogram. In the diagnosis of myocardial infarction, probably the oldest but still one of the best diagnostic criteria is the Q wave on EKG. The Q wave appears at the time of infarction and usually remains as a marker from then on. The largest study to date on the significance of the Q wave is the study by Horan and co-workers in 1971 in which they compared .03 second Q waves on EKG to autopsy scars in 1184 patients. This study is summarized in Table 13 on the following page.

Table 13

| Location of Q Waves | Sensitivity | Specificity | Prevalence .50 Predictive Value + Test | Prevalence .50 Predictive Value - Test |
|---|---|---|---|---|
| Any one of following: Septal (V 1-2), Anterior (V 3-4) Anteroseptal (V 1-4) Inferior (II III AVF) | .61 | .89 | .85 | .70 |
| Any 2 of above and/or Lateral (I, V5-6) | .38 | .99 | .97 | .61 |

Note that this study again demonstrates that the Q wave is not very sensitive for infarction, but is quite specific. In the study Q waves in one area of the ventricle had a specificity of .89, but when the Q waves were present in two areas or on the lateral wall, specificity rose to .99. Assuming a prevalence of .50 in the population to be tested, this high specificity translates into a high predictive value for a positive test.

A popular EKG diagnosis is that of strictly posterior myocardial infarction by using the criteria of an R wave equal to or greater than .04 second duration in V1 and/or an R/S ratio in V1 of over 1. When this occurs in conjunction with the EKG finding of inferior infarct there is little problem, but what is the probability of the patient having had an infarct when this V1 configuration occurs as an isolated finding on an otherwise normal EKG? The study usually quoted as reference for this diagnostic criteria was done in 1964 (Perloff). Recall that sensitivity is determined by taking all patients with disease and seeing what percentage of these patients have a positive test. The presence or absence of disease should be determined by an independent standard. In the 1964 study the presence of disease was determined by picking out vector cardiograms which were felt diagnostic of posterior infarction. The vector cardiogram is certainly not independent of the electrocardiogram. In essence then sensitivity in this study was determined by diagnosing a group of patients as having disease by an abnormality on vector cardiography and then seeing what percentage of these patients had a highly related abnormality on EKG.

Recall that specificity is determined by taking all patients without disease and seeing what percentage of these patients have a negative test. Again the presence or absence of disease should be determined by independent criteria. In this posterior infarct study the patients who did not have disease were selected by having EKG's that were "within normal limits". Specificity was then calculated by seeing what percentage

of these "within normal limits" EKG's did not have the EKG abnormalities previously mentioned. It is obvious that in this study the criteria used to determine disease were very similar to the criteria being tested, which leaves the conclusions of the study -- that is that an R wave in V1 ≥ .04 seconds or an R/S ratio of V1 ≥ 1 is highly predictive of posterior infarction -- in serious doubt. Unfortunately to my knowledge there is no firm data on this EKG abnormality, although the military study on asymptomatic personel showed that the presence of an R/S ratio ≥ 1 in V1 varied from about .012 at age 20 to .002 at age 45 (Hiss). Hypothetically, using a specificity of .99 suggested by the military study, guessing at a sensitivity of .50, and using a prevalence of .01 such as may be reasonable in routine EKG reading without clinical information or other suggestive EKG abnormalities, the predictive value of a positive EKG is about .33.

AORTIC STENOSIS - by carotid pulse tracing and fluoroscopy:

Significant aortic stenosis should be treated surgically, and if left untreated has a poor prognosis. This means that accurate identification of patients with significant aortic stenosis is essential, but this identification is complicated by the large number of patients with murmurs simulating aortic stenosis who do not have significant disease. Experienced clinicians generally consider significant aortic stenosis one of the most difficult bedside diagnoses to make. Catheterization is the most definitive diagnostic test, however, the procedure is involved, consequently noninvasive tests are frequently used in an attempt to select those patients to be catheterized. The most commonly used tests for this purpose are the evaluation of the carotid pulse contour, either informally by palpation or formally by carotid pulse recordings, and the radiographic determination of the presence or absence of calcium in the aortic valve. The most important clinical decision is whether significant aortic stenosis can be safely excluded, hence the predictive value of a negative test is the value of most concern. In evaluating these tests only patients with minimal or no aortic insufficiency are included since significant aortic insufficiency is relatively easy to diagnose, its presence changes the sensitivity and specificity of the tests, and insufficiency in itself is indication for catheterization and surgery. The division between significant and insignificant stenosis is a 50 mmHg gradient and/or a .75 $cm^2$ aortic valve area. Although many studies use a control group of normal people, in this evaluation I am using only control groups of patients being studied for possible significant stenosis. I am omitting many of the earlier and frequently quoted studies because of the lack of documentation of the severity of stenosis. All studies described here had the severity of the stenosis documented by catheterization.

In Table 14 below I have listed the most commonly used measurements of the duration and shape of the carotid pulse contour.

Table 14

| Study | Measurement | Sensitivity | Specificity | Prevalence .50 Predictive Value + Test | Prevalence .50 Predictive Value - Test |
|---|---|---|---|---|---|
| Epstein, 1964 | Uncorrected ejection time | | | | |
| | >.34 secs | .78 | .47 | .60 | .68 |
| | >.36 secs | .56 | .80 | .74 | .65 |
| Bonner, 1973 | Corrected ejection time | | | | |
| | >.43 secs | .50 | .82 | .74 | .62 |
| Bonner, 1973 | Q-peak murmur | | | | |
| | ≥.20 secs | .61 | .82 | .77 | .68 |
| | ≥.24 secs | .19 | 1.00 | 1.00 | .55 |
| Bonner, 1973 | Max. rate of rise carotid | | | | |
| | <500 mmHg | .92 | .36 | .59 | .82 |
| | ≤400 mmHg | .81 | .64 | .69 | .77 |
| Epstein, 1964 | Upstroke (u) time | | | | |
| | >.12 secs | .91 | .27 | .55 | .75 |
| | >.17 secs | .62 | .47 | .54 | .55 |
| Epstein, 1964 | "t" time | | | | |
| | >.046 secs | .81 | .53 | .63 | .74 |
| | >.055 secs | .69 | .73 | .72 | .70 |
| Bonner, 1973 | Ejection time>.42 Max. rate rise<500 Q-peak M>.19 | .75 | .91 | .89 | .78 |

The ejection time is measured from the upstroke to the incisural notch on the carotid tracing. If the corrected value is used, it is calculated by the formula: LVET (corrected) = LVET (uncorrected) + 1.7 Heart Rate for males or LVET (corrected) = LVET (uncorrected) + 1.6 Heart Rate for females (Weissler). The Q-peak murmur is measured from the Q wave on EKG to the point of maximum intensity of the murmur on the phonocardiogram. The three measurements that are most commonly used to evaluate the delay in upstroke of the carotid pulse are the maximum rate of rise, the "u" time, and the "t" time. These measurements are the closest objective measurements to the subjective feel the clinician gets when palpating the carotid. The maximum rate of rise is measured by taking the most rapid point of upstroke on the pulse and by using an arm cuff blood pressure calculating the mmHg/second which the pressure rises (Bonner). The "u" time is measured from the upstroke to the peak of the pulse, and the "t" time is measured from the upstroke to half the height of the pulse (Epstein). In the study by Bonner, 36 patients had significant stenosis and 11 had insignificant stenosis. In the study by Epstein, 32 patients had significant stenosis, and 15 had insignificant stenosis.

In the table, the cutoff values are those suggested by the authors. It is apparent that none of these tests separate patients well. Sensitivity can not be raised by selecting a cutoff point without dropping specificity to intolerably low levels, and vice versa. When the predictive value of a test is calculated assuming .50 prevalence and the sensitivity and specificity given, it is apparent that the most important value, that is - the predictive value of a negative test -- is never very high meaning that significant stenosis cannot be safely excluded. When 3 criteria are combined such that the test is positive only if all all 3 criteria are positive, the specificity and predictive value of a positive test increase, but because of a low sensitivity, the predictive value of a negative test remains relatively low at .78.

In Table 15 I have summarized the value of using radiographic calcium in the aortic valve for predicting significant stenosis.

Table 15

| Study | $Ca^{++}$ in Aortic Valve | Sensitivity | Specificity | Prevalence .50 Predictive Value + Test | Prevalence .50 Predictive Value - Test |
|---|---|---|---|---|---|
| Glancy, 1969 | 1+ (fluoro only) | 1.00 | .46 | .65 | 1.00 |
| Eddleman, 1973 | 1+ | .95 | .40 | .62 | .89 |
| Glancy, 1969 | 2+ (Specks on CXR) | .85 | .54 | .65 | .78 |
| Glancy, 1969 | 3+ (heavy) | .63 | 1.00 | 1.00 | .73 |

1+ calcium is that seen only on fluoroscopy while 3+ is the heavy calcium seen in the lateral chest X-Ray. In the Glancy study 46 patients had significant stenosis and 13 patients had insignificant stenosis. In the Eddleman study, 40 patients had significant stenosis and 15 patients had insignificant stenosis. Several pitfalls should be pointed out. First, this data applies only to patients over 35 years old. Younger patients more commonly have significant stenosis without calcium. Second, when mitral stenosis is present the sensitivity probably drops since aortic stenosis without calcification probably occurs more commonly when of rheumatic origin. Third, proximal coronary calcification can be mistaken for aortic valve calcification. From the table it appears that calcification is the best noninvasive test, and fortunately is one of the most available. If calcium is not present on fluoroscopy in an older patient who does not have mitral stenosis the chance of him not having significant stenosis is very good. It is not perfect however. Conversely if heavy calcification is present, the chance of having significant stenosis is very high.

MITRAL STENOSIS - by echocardiography

The diagnosis of mitral stenosis was the first use of echocardiography, having been described 22 years ago. Since then echo has been used extensively for this purpose and most clinicians would consider the diagnosis of mitral stenosis one of the most significant uses of the tool. Yet even now the sensitivity and specificity are not completely clear. The relatively slow development of a complete understanding of the test is at least in part due to the subjective nature of echocardiography. Not only must the relationship of an echo to the disease process be studied, but in addition the interpretation by the clinician of a given echo must be studied. This interpretation obviously may vary between observers complicating evaluation of studies from different centers. The study of observer detection performance is complex, but some information is available, especially in regard to radiology (Metz, Swets, Lusted).

Originally a diminished E-F slope was used as the criteria for the presence of mitral stenosis (see Feigenbaum for explanation of echo terms). A diminished slope is very sensitive for the presence of anatomic mitral stenosis, but it soon became apparent that the diminished E-F slope is not specific for mitral stenosis for it is also found in other conditions which can be clinically confused with mitral stenosis, chiefly poor ventricular compliance and pulmonary hypertension. Therefore, the predictive value of a negative test (i.e. - normal E-F slope) is very good, but the predictive value of a positive test (i.e. a diminished E-F slope) is not very good. Consequently a new echo sign was studied and felt to be the best discriminator - that of diastolic anterior motion of the posterior mitral leaflet instead of the normal posterior motion (Duchak).

The procedure then for interpreting an echo in respect to anatomic mitral stenosis is to first evaluate the E-F slope. If the E-F slope is

normal, then mitral stenosis may be excluded since the predictive value of a negative test is very good. If however the E-F slope test is positive (E-F slope less then 40 mm/sec - Cope) then further evaluation must be done since the predictive value of this positive test is not very good. For this further evaluation the test is the movement of the posterior mitral leaflet. The population of patients to whom the posterior leaflet test is applied therefore are those patients with a positive E-F slope test. If the posterior leaflet moves in the abnormal anterior direction then the test is positive. We assume the specificity of the sign to be good, but to my knowledge this has not been systematically studied. On the other hand, the sensitivity of the posterior leaflet sign was initially thought to be perfect. Since this would eliminate any patients who had the disease from having a negative test, the predictive value of a negative test would be perfect (i.e. in the population of patients with a diminished E-F slope, the presence of normal posterior leaflet motion would rule out mitral stenosis). However, recent reports indicate that the sensitivity of the posterior leaflet test is not perfect, and in fact is only about .90 (Berman, Ticzan, Levisman, Cope), meaning that the predictive value of a negative posterior leaflet test is not perfect. In summary, if the E-F slope is normal, mitral stenosis is not present. If the E-F slope is abnormal and the posterior leaflet moves in the abnormal direction mitral stenosis is highly probable, but has not been shown to be certain. If the E-F slope is abnormal, and the posterior leaflet moves in the normal direction, mitral stenosis is usually but not always absent. It is obvious that echocardiography is chiefly qualitative at this stage of its development.

LEFT VENTRICULAR FUNCTION AND/OR HYPERTROPHY - by chest X-Ray and EKG

The chest X-Ray and electrocardiogram are used so commonly to judge cardiac function that we tend to take their use for granted. However, relatively recent studies using angiocardiography, echocardiography, or large autopsy studies as standards have better defined the limits of the X-Ray and EKG. In this section I will review the ability of the X-Ray and EKG to predict left ventricular function, left ventricular hypertrophy, and left atrial enlargement.

Cardiac enlargement on the PA chest film in the absence of volume overload has long been recognized as an indicator of poor ventricular function. Recognition of patients with poor ventricular function has achieved additional concrete importance lately with the realization that patients with poor ventricular function are poor candidates for coronary artery bypass surgery. If this were recognized prior to cardiac catheterization, some patients could conceivably be spared the procedure. In Table 16 on the following page, the ability of the chest X-Ray to predict poor ventricular function in patients with coronary artery disease is shown.

Table 16

| Study | Sensitivity | Specificity |
|---|---|---|
| Aintablian, 1976 | .41 | .90 |
| Stein, 1974 | .29 | .92 |
| Approximate Mean | .33 | .90 |

There were 207 patients in the Aintablian study and 64 patients in the Stein study. The standard was cardiac catheterization where either an ejection fraction of 50% or a mean rate of circumferential fiber shortening of .59 circumferences/second was used as the dividing point between normal and poor function. Cardiac enlargement on the X-Ray is said to be present if the cardiothoracic ratio is over 50% (Stein) or 1.5 cm greater than 50% (Aintablian). It is apparent that specificity is fairly good at .90, but that sensitivity is low at .33.

Using these figures of sensitivity .33 and specificity .90, Table 17 shows the predictive value at prevalences of .10 and .50.

Table 17

| Prevalence | Predictive Value Positive Test | Predictive Value Negative Test |
|---|---|---|
| .10 | .27 | .92 |
| .50 | .77 | .57 |

At a prevalence of .10 such as would be likely in routinely reviewing films in a hospital admitting room, little could be said on the basis of chest X-Ray. At a prevalence of .50 such as would be likely in using the X-Ray in a patient suspected of having poor ventricular function, the X-Ray is not very good, especially in ruling out poor ventricular function - i.e. a normal chest X-Ray increases a patient's chances of not having poor function from .50 to .57.

An increase in left ventricular mass is found in pressure and volume overload, various hypertropic states, and in heart failure. Its presence is frequently of aid in diagnosis and clarification of heart disease. In Table 18 on the following page, the sensitivity, specificity, and predictive value of a positive test are listed for several popular EKG criteria (Baxley, Romhilt).

Table 18

| Criteria | Sensitivity | Specificity | Predictive Value of Positive Test at Prevalence of .05 | .50 |
|---|---|---|---|---|
| Grant | .67 | .73 | .12 | .71 |
| Sokolow | .68 | .73 | .12 | .72 |
| SV1-2 + RV6 40mm | .60 | .89 | .22 | .85 |
| Left Atrial Enlargement | .48 | .95 | .34 | .91 |
| Estes 4 points | .62 | .97 | .52 | .95 |
| Estes 5 points | .58 | .97 | .50 | .95 |

The standard for increase in mass is either angiocardiography (LV mass > 125 gm/$m^2$) or autopsy (Heart mass > 1.9 body length (cm) + 38 gms. in males or > 1.78 body length + 8 gms. in females). There were 112 patients in the Baxley study and 150 patients in the Romhilt study. The criteria are listed in order of increasing usefulness. The predictive value of a positive test is given at a prevalence level of .05 simulating routine hospital EKG reading (without information) and .50 simulating patients being evaluated for heart disease.

The criteria of Grant and Sokolow are long and will not be listed here, but can be found in the article by Baxley. Two rather simple criteria -- that of the sum of the S wave in V1 or V2 plus the R wave in V6 exceeding 40 mm, and left atrial enlargement (by product of duration and depth of terminal deflection of P wave in V1 equaling or exceeding .04 mm-secs) -- rank well compared to other criteria. The best criteria seem to be the point system of Romhilt and Estes. In this system points are given as follows:

| | Points |
|---|---|
| 1) Any one of below | 3 |
| a) Largest R or S wave in limb leads ≥ 20 mm | |
| b) S wave in V1 or V2 ≥ 30 mm | |
| c) R wave in V5 or V6 ≥ 30 mm | |

| | |
|---|---|
| 2) Typical ST-T strain pattern | |
| a) without digitalis | 3 |
| b) with digitalis | 1 |
| 3) Left atrial involvement (terminal negativity of P wave in V1 1 mm or more in depth and .04 seconds or more in duration) | 3 |
| 4) Left axis deviation of - $30^{o}$ or more | 2 |
| 5) QRS duration ≥ .09 secs | 1 |
| 6) Intrinsicoid deflection in V5 or V6 ≥ .05 secs | 1 |

It is apparent that even in this the best system on routine reading, when LVH is not clinically suspected LVH can only be diagnosed with an approximate .50 probability. These criteria are generally applicable to patients over about 35 years of age. It is generally felt that EKG criteria for LVH, especially hypervoltage, is less specific in younger people. The EKG findings in normal military personel (Hiss, Averill) generally agree with this lower specificity in people under 30 - 35, but also support the high specificity in older people.

Left atrial enlargement is found in chronic pressure or volume overload of the left atrium. As such it is a valuable finding, especially of mitral valve disease or chronically elevated left ventricular end diastolic pressure. In Table 19 below, the sensitivity and specificity of the sign of any indentation of the barium filled esophagus on lateral chest X-Ray is given.

Table 19

| Study | Sensitivity | Specificity |
|---|---|---|
| Hirata, 1969 | .88 | .75 |
| Glover, 1973 | .84 | .57 |
| Approximate Mean | .85 | .65 |

The standard is left atrial angiography. The two studies have 24 and 99 patients with enlarged left atria, but only 4 and 7 patients with normal sized atria, hence the reliability of the specificity is questionable. However, taking the data at face value, sensitivity is high but specificity is low. Predictive value is shown later.

In Table 20 below the sensitivity and specificity of EKG criteria for left atrial enlargement are shown.

Table 20

| Study | Criteria | No. of Patients: Enlarged Atrium | No. of Patients: Normal Atrium | Sensitivity | Specificity |
|---|---|---|---|---|---|
| Chirife, 1975 | PII ≥ 105 ms | 12 | 36 | 1.00 | .89 |
| Termini, 1975 | PII ≥ 110 ms | 31 | 69 | .95 | .67 |
| Kasser, 1969 | PII ≥ 110 ms | 36 | 91 | .74 | .94 |
| Approximate Mean | PII ≥ 110 ms | | | .90 | .80 |
| Kasser, 1969 | V1 ≥ .03 mm-sec | 36 | 91 | .72 | .94 |
| Chirife, 1975 | V1 ≥ .04 mm-sec | 12 | 36 | .75 | .83 |
| Termini, 1975 | V1 ≥ .04 mm-sec | 31 | 69 | .89 | .88 |
| Approximate Mean | V1 ≥ .04 mm-sec | | | .80 | .90 |

The standard is left atrial size determined by angiocardiography or echo. Echo volumes correlate well with angiocardiography and so both methods are good standards. The two criteria used are a P wave in limb lead II of greater than 110 msec duration or a terminal negativity of the P wave in V1 equal to or greater than .04 mm-sec (determined by multiplying the duration of the negative deflection in secs. times the depth of the negativity in mms). Using these criteria both sensitivity and specificity are high at approximately .80 - .90.

Table 21 shows the predictive value of positive X-Ray and EKG signs just discussed in diagnosing left atrial enlargement. It appears that none of the tests are very definitive, but that the sign of terminal negativity in V1 is best.

Table 21

| Criteria | at Prevalence .05 | at Prevalence .50 |
|---|---|---|
| Lateral CXR | .12 | .71 |
| PII ≥ 110 ms | .19 | .82 |
| V1 ≥ .04 mm-sec | .30 | .89 |

BIBLIOGRAPHY

Diagnostic Process:

Baroon S, Wolfe H: Measuring the Effectiveness of Medical Decisions. Charles C. Thomas, 1972. Springfield, Illinois.

Cornfield J: Bayes Theorem. Proc 6th IBM Med Symposium. Oct 5-9, 163-196, 1964.

Edwards W, Lindamn H, Savage JL: Bayesian statistical inference for psychological research. Pychol Rev 70:193-242, 1963.

Galen RS, Gambino SR: Beyond Normality: The Predictive Value and Efficiency of Medical Diagnoses. John Wiley and Sons, 1975. New York.

Hall GH: The clinical application of Bayes' theorem. Lancet 9-9-67: 555-557, 1967.

Katz MA: A probability graph describing the predictive value of a highly sensitive diagnostic test. N Engl J Med 291:1115-1116, 1974.

Ledley RS, Lusted LB: Reasoning foundations of medical diagnosis. Science 130:9-21, 1959.

Lusted LB: Decision-making studies in patient management. N Engl J Med 284:416-424, 1971.

Lusted LB: Introduction to Medical Decision Making. Charles C. Thomas, 1968. Springfield, Illinois.

McNeil BJ, Keeler E, Adelstein SJ: Primer on certain elements of medical decision making. N Engl J Med 293:211-215, 1975.

McNeil BJ, Adelstein SJ: Determining the value of diagnostic and screening tests. J Nuclear Med 17:439-448, 1976.

Metz CE, Goodenough DJ, Rossmann K: Evaluation of receiver operating characteristic curve data in terms of information theory, with applications in radiography. Diagnostic Radiology 109:297-304, 1973.

Pauker SG, Kassirer JP: Therapeutic decision making: A cost-benefit analysis. N Engl J Med 293:229-234, 1975.

Pauker SG: Coronary artery surgery: The use of decision analysis. Ann Intern Med 85:8-18, 1976.

Rautaharju PM, Blackburn HW, Warren JW: The concept of sensitivity, specificity and accuracy in evaluation of electrocardiographic, vector-cardiographic and polarcardiographic criteria. J Electrocardiology 9:275-281, 1976.

Schwartz WB, Gorry GA, Kassirer JP, Essig A: Decision analysis and clinical judgment. Am J Med 55:459-472, 1973.

Sisson JC, Schoomaker EB, Ross JC: Clinical decision analysis: The hazard of using additional data. JAMA 236:1259-1263, 1976.

Swets JA: Signal Detection and Recognition by Human Observers. John Wiley and Sons, 1964. New York.

Zar JH: Biostatistical Analysis. Prentice-Hall, Inc., 1974. Englewood Cliffs, N.J.

Tests:

Ahmad M, Dubiel JP, Logan KW, Verdon TA, Martin RH: Limited clinical diagnostic specificity of technetium-99m stannous pyrophosphate myocardial imaging in acute myocardial infarction. Am J Cardiol 39:50-54, 1977.

Aintablian A, Hamby RI, Garsman J: Correlation of heart size with clinical and hemodynamic findings in patients with coronary artery disease. Am Heart J 91:21-27, 1976.

Ascoop CA, Simoons ML, Egmond WG, Bruschke AVG: Exercise test, history, and serum lipid levels in patients with chest pain and normal electrocardiogram at rest: Comparison to findings at coronary arteriography. Am Heart J 82:609-617, 1971.

Averill KH, Lamb LE: Electrocardiographic findings in 67,375 asymptomatic subjects. I. Incidence of abnormalities. Am J Cardiol 6:76-83, 1960.

Bartel AG, Behar VS, Peter RH, Orgain ES, Kong Y: Graded exercise stress tests in angiographically documented cononary artery disease. Circulation 49:348-355, 1974.

Baxley WA, Dodge HT, Sandler H: A quantitative angiocardiographic study of left ventricular hypertrophy and the electrocardiogram. Circulation 37:509-517, 1968.

Berman ND, Gilbert BW, McLaughlin PR, Morch JE: Mitral stenosis with posterior diastolic movement of posterior leaflet. Canadian Med Assoc J 112:976-979, 1975.

Blomberg DJ, Kimber WD, Burke MD: Creatine kinase isoenzymes. Am J Med 59:464-469, 1975.

Blomqvist CG: Use of exercise testing for diagnostic and functional evaluation of patients with arteriosclerotic heart disease. Circulation 44:1120-1136, 1971.

Bonner AJ, Sacks HN, Tavel ME: Assessing the severity of aortic stenosis by phonocardiography and external carotid pulse recordings. Circulation 48:247-252, 1973.

Bonte FJ, Parkey RW, Graham KD, Moore J, Stokely EM: A new method for radionuclide imaging of myocardial infarcts. Radiology 110:473-474, 1974.

Borer JS, Brensike JF, Redwood DR, Itscoitz SB, Passamani ER, Stone NJ, Richardson JM, Levy RI, Epstein SE: Limitations of the electrocardiographic response to exercise in predicting coronary-artery disease. N Engl J Med 293:367-371, 1975.

Bruce RA: Values and limitations of exercise electrocardiography. Editorial. Circulation 50:1-3, 1974.

Cheitlin MD, Davia JE, de Castro CM, Barrow EA, Anderson WT: Correlation of "critical" left coronary artery lesions with positive submaximal exercise tests in patients with chest pain. Am Heart J 89:305-310, 1975.

Chirife R, Feitosa GS, Frankl WS: Electrocardiographic detection of left atrial enlargement. Correlation of P wave with left atrial dimension by echocardiography. Br Heart J 37:1281-1285, 1975.

Christiansen JS, Mathiesen B, Andersen AR, Calberg H: Diagonal earlobe crease in coronary heart disease. (Letter) N Engl J Med 293:308-309, 1975.

Cohen MV, Gorlin R: Main left coronary artery disease. Circulation 52:275-285, 1975.

Cope GD, Kisslo JA, Johnson ML, Behar VS: A reassessment of the echocardiogram in mitral stenosis. Circulation 52:664-670, 1975.

Cramer K, Paulin S, Werkö L: Coronary angiographic findings in correlation with age, body weight, blood pressure, serum lipids, and smoking habits. Circulation 33:888-900, 1966.

Duchak JM, Chang S, Feigenbaum H: The posterior mitral valve echo and the echocardiographic diagnosis of mitral stenosis. Am J Cardiol 29: 628-632, 1972.

Eddleman EE, Frommeyer WB, Lyle DP, Bancroft WH, Turner ME: Critical analysis of clinical factors in estimating severity of aortic valve disease. Am J Cardiol 31:687-695, 1973.

Ehsani A, Ewy GA, Sobel BE: Effects of electrical countershock on serum creatine phosphokinase (CPK) isoenzyme activity. Am J Cardiol 37:12-18, 1976.

Epstein EJ, Coulshed N: Assessment of aortic stenosis from the external carotid pulse wave. Br Heart J 26:84-96, 1964.

Erikssen J, Enge I, Forfang K, Storstein O: False positive diagnostic tests and coronary angiographic findings in 105 presumably healthy males. Circulation 54:371-376, 1976.

Feigenbaum H: Echocardiography. Lea and Febiger, 1976. Philadelphia.

Froelicher VF, Thomas MM, Pillow C, Lancaster MC: Epidemiologic study of asymptomatic men screened by maximal treadmill testing for latent coronary artery disease. Am J Cardiol 34:770-776, 1974.

Froelicher VF, Yanowitz FG, Thompson AJ, Lancaster MC: The correlation of coronary angiography and the electrocardiographic response to maximal treadmill testing in 76 asymptomatic men. Circulation 48: 597-604, 1973.

Galen RS, Reiffel JA, Gambino SR: Diagnosis of acute myocardial infarction. JAMA 232:145-147, 1975.

Glancy DL, Freed TA, O'Brien P, Epstein SE: Calcium in the aortic valve. Ann Intern Med 71:245-250, 1969.

Glover L, Baxley WA, Dodge HT: A quantitative evaluation of heart size measurements from chest roetgenograms. Circulation 47:1289-1296, 1973.

Goldschlager N, Selzer A, Cohn K: Treadmill stress tests as indicators of presence and severity of coronary artery disease. Ann Intern Med 85:277-286, 1976.

Hiss RG, Lamb LE, Allen MF: Electrocardiographic findings in 67,375 asymptomatic subjects. X. Normal Values. Am J Cardiol 6:200-231, 1960.

Hirata T, Wolfe SB, Popp RL, Helmen CH, Feigenbaum H: Estimation of left atrial size using ultrasound. Am Heart J 78:43-52, 1969.

Horan LG, Flowers NC, Johnson JC: Significance of the diagnostic Q wave of myocardial infarction. Circulation 43:428-436, 1971.

Kansal S, Roitman D, Shiefield T: Stress testing with ST-segment depression at rest. Circulation 54:636-639, 1976.

Kasser I, Kennedy JW: The relationship of increased left atrial volume and pressure to abnormal P waves on the electrocardiogram. Circulation 39:339-343, 1969.

Kelemen MH, Gillilan RE, Bouchard RJ, Heppner RL, Warbasse JR: Diagnosis of obstructive coronary disease by maximal exercise and atrial pacing. Circulation 48:1227-1233, 1973.

Khaja F, Sharma SD, Easley RM, Heinle RA, Goldstein S: Left main coronary lesions. Circulation 49 & 50 (Supp.II):136-140, 1974.

Konttinen A, Somer H: Specificity of serum creatine kinase isoenzymes in diagnosis of acute myocardial infarction. Br Med J 1:386-389, 1973.

Levisman JA, Abbasi AS, Pearce ML: Posterior mitral leaflet motion in mitral stenosis. Circulation 51:511-514, 1975.

Lichstein E, Chadda KD, Naik D, Gupta PK: Diagonal ear-lobe crease: Prevalence and implications as a coronary risk factor. N Engl J Med 290:615-616, 1974.

Mehta J, Hamby RI: Diagonal ear-lobe crease as a coronary risk factor. (Letter) N Engl J Med 291:260, 1974.

Linhart JW, Turnoff HB: Maximum treadmill exercise test in patients with abnormal control electrocardiograms. Circulation 49:667-672, 1974.

Linhart JW, Laws JG, Satinsky JD: Maximum treadmill exercise electrocardiography in female patients. Circulation 50:1173-1178, 1974.

Martin CM, McConahay DR: Maximal treadmill exercise electrocardiography. Circulation 46:956-962, 1972.

McConahay DR, McCallister BD, Smith RE: Postexercise electrocardiography: Correlations with coronary arteriography and left ventricular hemodynamics. Am J Cardiol 28:1-9, 1971.

McLaughlin P, Coates G, Wood D, Cradduck T, Morch J: Detection of acute myocardial infarction by technetium-99m polyphosphate. Am J Cardiol 35:390-396, 1975.

Norris RM, Brandt PWT, Caughey DE, Lee AJ, Scott AF: A new coronary prognostic index. Lancet 2/8/69:274-278, 1969.

O'Rourke R, Righetti A, Schelbert H, Daily P, Ashburn W, Ross J: Usefulness of pre- and post-operative Tc-99m-pyrophosphate scans in cardiac surgical patients. (Abstr) Am J Cardiol 37:161, 1976.

Perloff J: The recognition of strictly posterior myocardial infarction by conventional scalar electrocardiography. Circulation 30:706-718, 1964.

Piessens J, Van Mieghem W, Kesteloot H, De Geest H: Diagnostic value of clinical history, exercise testing and atrial pacing in patients with chest pain. Am J Cardiol 33:351-355, 1974.

Prasquier R, Taradash MR, Botvinick EH, Shames DM, Parmley WW: The specificity of the diffuse pattern of cardiac uptake in myocardial infarction imaging with technetium-99m stannous pyrophosphate. Circulation 55:61-65, 1977.

Pugh BR, Buja LM, Parkey RW, Poliner LR, Stokely EM, Bonte FJ, Willerson JT: Cardioversion and "false positive" technetium-99m stannous pyrophosphate myocardial scintigrams. Circulation 54: 399-403, 1976.

Redwood DR, Borer JS, Epstein SE: Whither the ST segment during exercise? Editorial. Circulation 54:703-706, 1976.

Redwood DR, Epstein SE: Uses and limitations of stress testing in the evaluation of ischemic heart disease. Circulation 46:1115 1131, 1972.

Rios JC, Hurwitz LE: Electrocardiographic responses to atrial pacing and multistage treadmill exercise testing. Am J Cardiol 34:661-666, 1974.

Roberts R, Henry PD, Witteeveen SAGJ, Sobel BE: Quantification of serum creatine phosphokinase isoenzyme activity. Am J Cardiol 33:650-654, 1974.

Roberts R, Gowda KS, Ludbrook PA, Sobel BE: Specificity of elevated serum MB creatine phosphokinase activity in the diagnosis of acute myocardial infarction. Am J Cardiol 36:433-437, 1975.

Roitman D, Jones WB, Sheffield LT: Comparison of submaximal exercise ECG test with coronary cineangiocardiogram. Ann Intern Med 72:641-647, 1970.

Romhilt DW, Estes EH: A point-score system for the ECG diagnosis of left ventricular hypertrophy. Am Heart J 75:752-758, 1968.

Sketch MH, Mohiuddin SM, Lynch JD, Zencka AE, Runco V: Significant sex differences in the correlation of electrocardiographic exercise testing and coronary arteriograms. Am J Cardiol 36:169-173, 1975.

Stein PD, Lewinson H, Potts KH: Cardiac size and left ventricular performance. JAMA 229:1614-1620, 1974.

Takaro T, Hultgren HN, Lipton MJ, Detre KM: The VA cooperative randomized study of surgery for coronary arterial occlusive disease. Circulation 54 (Suppl. 3):107-116, 1976.

Termini BA, Lee Y: Echocardiographic and electrocardiographic criteria for diagnosing left atrial enlargement. Southern Med J 68:161-165, 1975.

Ticzon AR, Damato AN, Caracta AR, Lau SH, Gross L: Echocardiographic manifestation of "false" mitral stenosis that was. Ann Intern Med 83:503-505, 1975.

Tonkin AM, Lester RM, Guthrow CE, Roe CR, Hackel DB, Wagner GS: Persistence of MB isoenzyme of creatine phosphokinase in the serum after minor iatrogenic cardiac trauma. Circulation 51:627-631, 1975.

Varat MA, Mercer DW: Cardiac specific creatine phosphokinase isoenzyme in the diagnosis of acute myocardial infarction. Circulation 51:855-859, 1975.

Wagner GS, Roe CR, Limbird LE, Rosati RA, Wallace AG: The importance of identification of the myocardial-specific isoenzyme of creatine phosphokinase (MB form) in the diagnosis of acute myocardial infarction. Circulation 47:263-269, 1973.

Weissler AM, Garrard CL: Systolic time intervals in cardiac disease. Modern Concepts of Cardio Disease 40:1-4, 1971.

Willerson JT, Parkey RW, Bonte FJ, Meyer SL, Atkins JM, Stokely EM: Technetium stannous pyrophosphate myocardial scintigrams in patients with chest pain of varying etiology. Circulation 51:1046-1052, 1975.

Willerson FT, Parkey RW, Bonte FJ, Meyer SL, Stokely EM: Acute subendocardial myocardial infarction in patients. Circulation 51:436-441, 1975.