General Session II

"Epidemiology for the Practitioner"

Moderator: **John Ferry**

Using and Interpreting Diagnostic Tests

Dr. Liz Spangler, *DVM, PhD*
Assistant Professor, Epidemiology
Department of Health Management
Atlantic Veterinary College
University of Prince Edward Island

Everything a veterinary clinician does ends in a decision: to treat, to do more tests, or to wait. The diagnostic process is, of necessity, a dualistic one. We have a need for a binary classification system because the decisions we must make are usually dualistic or binary in nature themselves. But the data we have to use is complex, not necessarily dualistic, subjective and not infrequently, contradictory. It must be reduced, consciously or unconsciously, to a single judgement, sick or well. The purpose of this exercise is to examine objectively how diagnostic decisions are made, with the objective of improving diagnostic capabilities in food animal medicine.

The diagnostic problem in either the clinical setting or the screening situation is similar: the correct classification of our patients' disease condition. In clinical medicine, the goal is to assign the correct diagnosis from a list of problems with similar presentation. In screening for disease, the goal becomes to classify herds or individuals correctly as either free of some particular disease of interest or not. The ringer in this process is that diagnostic decisions are never made in isolation. There are economic, social and psychological consequences of making a diagnosis that depend largely upon the environment of the individual or herd. Whatever happens next depends on factors such as the effect of labelling on the herd or owner, the likelihood of owner compliance with testing and/or recommendations which follow from the testing, human considerations, our best estimate of cost/benefit ratio of the procedures we are considering, and other issues.

Diagnostic Strategies

Traditionally, clinicians rely on one of several methods to arrive at a diagnosis, depending on the condition and the clinical setting (Sackett *et al.* 1985). In veterinary school we are taught the *exhaustive method* (complete history and physical examination). This technique may be the method of choice for the novice, but rapidly, experience and practicality lead us away from this time consuming approach.

Some diseases are easily recognized by their appearance or presentation, particularly those that are dramatic or frequently encountered. This is diagnosis by *pattern recognition,* and it is a reflexive process, not a reflective one. If asked why you treated a particular postparturient cow for milk fever, you would be able to give a list of the appropriate reasons, but you may not have gone through them all consciously when you made your decision, it was part of your repertoire of experience. Pattern recognition involves the use of all the senses: visualizing the S-curve in the neck of a cow with milk fever, the odor of the breath of an animal with ketosis, the characteristic "ping" of a displaced abomasum, the tone of the reproductive tract associated with early pregnancy, or the thought processes that require you to give a bottle of calcium to a post parturient downer cow after you draw a blood sample but before you have the lab result.

The usual diagnostic strategy, performed subconsciously by most veterinarians, is called the "*hypothetico deductive strategy*" by Sacket (Sacket *et al* 1985). When presented with a new case, any clinician will very quickly develop a short list of rule outs. Consider this example: a week following parturition a dairy animal develops profuse diarrhea. Many practitioners will respond immediately with possible diagnoses of indigestion, salmonellosis, BVD, paratuberculosis, ostertagiasis, renal amyloidosis, copper deficiency, winter dysentery, etc. or with the suggestion that more information (age of animal, number of others affected, history of various diseases in the herd, etc.) should be obtained. A visit to the farm may result in a brief physical exam, taking the animal's temperature, a qualitative sensory examination of the stool, some historical information about the animal and the herd, possibly the taking of a stool sample and a blood sample, and a decision to treat. By now the clinician would have consciously or subconsciously assigned priority to each item on her list.

Tests are then performed to classify the animal as "normal" or "abnormal" with regard to certain findings. Based on the information collected so far, and the results of any subsequent tests, the list is shortened to the point of taking action. Rarely, however, is certainty of diagnosis a privilege. Often at the point of taking action, the clinician may have decided that the action has a high probability of success, or at least, a low probability of causing further harm. Often in fact, response to therapy (or lack of it) is the definitive diagnostic test.

In the hypothetico-deductive strategy, the general approach is the formulation of a short list of possible diagnoses or actions, followed by the directed acquisition of information which will shorten the length of the list. Every attempt we make to gain information is a "test", whether it is questions asked in the history taking, visual examination, auscultation, palpation, serology, microbiology, hematology, etc. It differs from the exhaustive approach in that veterinarians are selective and directed in the acquisition of information. Tests are only selected if their results might possibly influence what they do next.

What kinds of data do we use in making diagnostic decisions?

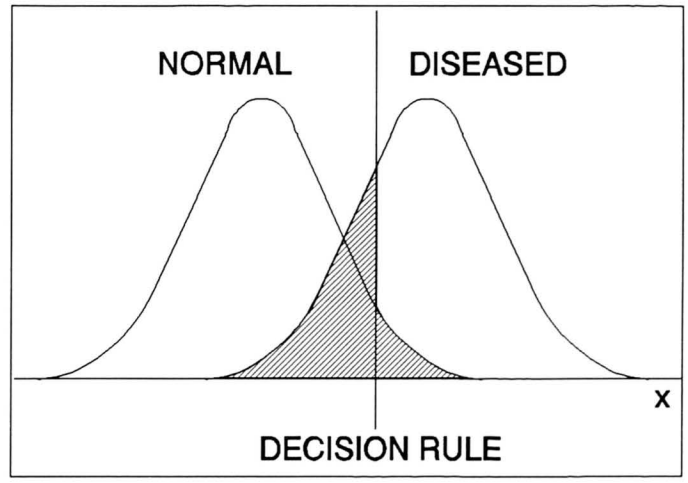
We use many kinds and sources of information when making a diagnosis. Traditionally, we think of laboratory data as being objective, dimensional and reliable, and so-called soft data as being subjective and less reliable, but, in reality, clinicians arrive at a correct diagnosis about 75% of the time after completing a history and physical exam. Clinical evaluation using this subjective data is far more powerful than laboratory examination, and more powerful than we give it credit for.

Clinical measurements are made on three different scales. Nominal scale data classify results into two categories, normal or abnormal, yes or no, pregnant or open, and so on. Questions asked in taking a history and information sought on physical examination are often nominal data. Many of the tests we use, like the California Mastitis test, or Body Condition Scoring, use a numerical scale to rank observations, relative to each other. This is ordinal data. Much of the so-called "hard" data (laboratory data like serum enzymes, antibody titer) clinicians deal with is measured on an interval scale (g/dl. serum optical density), although laboratory data can be measured on a nominal or ordinal scale as well.

What is normal?

"Normal" as it applies to diagnostic test interpretation can have several different meanings. Statistical normality refers to the distribution of interval measurements in a predictable bell shape about some average value. If we go out and measure a "reference population" of "normal" animals, we will get a range of values that may fall into a pattern like this. If a measurement, such as serum potassium, is normally distributed, then one possible way to differentiate normal from abnormal is to look at the distance an observation is from the mean value of the reference population of normal animals. With a normally distributed measurement, almost all of the observations will fall within 3 standard deviations of the mean, and 95% of observations will fall within two standard deviations of the mean. We could use this information to distinguish normal from abnormal, except for two problems. First, not very many clinical parameters have a normal distribution. Second, we are saying arbitrarily that 5% of the population will have abnormal clinical values, regardless of their underlying state of health. Another problem with the "normal limits" approach is that for most conditions and most tests, there is a gradual progression from the healthy to the diseased state. There is no magic line, no clear cut separation, between normal and diseased. Animals with test results within normal limits generally have a lower, but nonzero probability of disease (see Figure 1). Animals with results that are markedly deviated from normal ranges have a much greater likelihood of disease. Test results in the middle range give practitioners the most problems with interpretations.

Figure 1. Distribution of Normal and Diseased Populations



Often, data measured on an interval scale is reduced to a yes/no, positive/negative result before it is returned to the clinician for interpretation. There are two drawbacks to this. First, when a cutoff is chosen by a lab, and the result is reported only as positive or negative, information about the certainty of the result is lost to the clinician. When a test result is well beyond the cutoff between normal and abnormal, a clinician has greater confidence that the animal is truly abnormal than if a test result is fairly close to the cutoff (see Figure 1). If that information is not reported, it cannot be used. Second, the cutoff chosen by the lab may be the one that minimizes the total number of

errors, both false positives and false negatives, but it may not take into account the different costs of the two different errors. Depending upon the situation, a false positive diagnosis may be much more damaging than a false negative, and vice versa. For example, a veterinarian relying on the results of a test to make a decision about culling a cow probably wants to avoid a false positive diagnosis, particularly in the absence of other reasons for culling. On the other hand, when screening cattle for purchase into a herd, a false positive diagnosis would be much less harmful to your client than a false negative, which would allow diseased cattle into the herd. Ideally, a clinician needs the informaton to choose the cutoff that best meets his needs at the particular time.

For a thorough discussion of diagnostic data and normality, see **Clinical Epidemiology: the Essentials** by Fletcher, Fletcher and Wagner, pp. 19-41.

Critical evaluation of tests: how good a test is it?

Test selection is driven by the presenting signs and history of the animal, but the usefulness of the test is also dependent on certain characteristics of the test itself, as well as the setting in which the test is applied. When selecting a diagnostic test for use in you practice, it is important to realize that you can quantify the usefulness of the test to you. All tests, even some well accepted procedures currently in use, should be subjected to critical examination before you use them.

Some basic guidelines for the evaluation of tests follow. These guidelines should be followed by veterinarians before they make a decision whether or not to use a particular test. They should also be used in the critical examination of the literature on diagnostic tests, and therefore, this information MUST be provided by researchers developing and examining new and existing diagnostic tests.

The Gold Standard

The first question to ask is: "Has the test been compared to an appropriate "Gold Standard?" The appropriate gold standard is a test that is absolutely accurate, or at least the accepted standard for making a diagnosis at the present time. The researcher must have identified two groups, a group with the target disorder and a group without the target disorder, and the test must have been applied to both groups by someone who was BLIND to the true disease status of the animals being tested. From these two groups, sensitivity and specificity are calculated.

Keep in mind that the test populations should be similar to animals seen in your practice. The spectrum of disease in the diseased group should be similar to the spectrum of disease that you see, no better, no worse. In a clinical setting, the group without the target disorder should be free of the disease of interest, but it will have a variety of other conditions that could easily be confused clinically with the disease of interest. In a screening situation, the group without the target disorder are usually otherwise healthy, with few, if any, other disease conditions present.

It can be difficult to find an appropriate gold standard for comparison. Often, it is possible to fall back on the existing test, if it is well accepted as a diagnostic standard. A good example of such a test is fecal culture for paratuberculosis infection in cattle. The comparison of performance of a new test with fecal culture would properly be called *relative* sensitivity and specificity.

If no gold standard exists, and no appropriate comparisons can be made between an accepted test and a new test, then efforts should be made to compare the new test to some clinically relevant outcome. When two tests are being compared, and neither is the gold standard, then, at the very least, agreement and consistency should be determined.

The most important measures of a test describe the abilities of the test to discriminate between animals that truly have the target disorder and those which do not. They are measured againt the gold standard in a representative sample of affected and unaffected animals. The test's **sensitivity** is defined as the porportion of animals that actually have the disorder which test positive, or a/(a+c) in Table 1. The test's **specificity** is defined as the proportion of animals which do not have the target disorder which test negative, or d/(b+d) in Table 1. The results of repeated fecal culture (the gold standard) and Dot ELISA testing for paratuberculosis in three dairy herds (Woodruff *et al* 1991) are summarized in Table 2. Of 26 confirmed fecal culture positive cows in these herds, 18 of them had positive Dot ELISA tests. The sensitivity of the Dot ELISA relative to fecal culture, is 18/26 or 69%. Of 236 cows from which paratuberculosis was never isolated by fecal culture, only 5 were Dot ELISA positive, 231 were Dot ELISA negative. Therefore, the specificity of the Dot ELISA test is 231/236, or 98%, relative to fecal culture.

Table 1. Classification of Test Results and Disease Status in a Two x Two Table

	Disease Present (D+)	Disease Absent (D-)	
Test Positive (T+)	a	b	a+b
Test Negative (T-)	c	d	c+d
	a+c	b+d	n

Table 2. Determining the Sensitivity and Specificity of a Dot ELISA Test for Paratuberculosis (ref: Woodruff *et al* 1991).

	Fecal Culture Positive	Fecal Culture Negative	Totals
Dot ELISA Positive	18	5	23
Dot ELISA Negative	8	231	231
Totals	26	236	262

Sensitivity and Specificity are stable when they have been evaluated in an appropriate test population, and unless the spectrum of disease changes with the prevalence. If the test has a sensitivity of 100%, you can RULE OUT disease in test negative animals. If the test has a specificity of 100%, you can RULE IN disease in test positive animals. In the same study that reported the Dot ELISA results (Woodruff *et al* 1991), the AGID test was 100% specific, meaning that there were no AGID positive cattle among the fecal culture negative group. A positive AGID test can be used to rule in a diagnosis of paratuberculosis. All AGID positive cattle were also fecal culture positive. However, the sensitivity of the AGID test was only 34%, meaning that a large proportion of cattle with paratuberculosis confirmed by fecal culture were missed by the AGID test. A negative AGID test result is therefore not very useful.

If the test result is continuous, there is often an inverse relationship between sensitivity and specificity. Cutpoints can be chosen to maximize one or the other, or to minimize overall misclassification, depending on the relative costs of false positives and false negatives. Appropriate assessment of tests on a multilevel scale should include estimates of sensitivity and specificity at various levels of the test. This provides the most diagnostically useful informaton to the practitioner, as mentioned previously.

True Prevalence = (a+c)/n. This equation expresses the true prevalence of disease in the sampled population. This only represents the actual prevalence of disease in the target population if the sampled population is a true random sample of the target population. **Apparent Prevalence** = (a+b)/n. When using a diagnostic test all we know is the proportion of animals that test positive, called the apparent prevalence, (a+b)/n. We don't know the true disease status of the population. In Table 2, the true prevalence, determined by repeated fecal culture, in the three dairy herds, was 26/262 or 10%. The apparent prevalence, based on the Dot ELISA was 23/262 or 9%. The apparent prevalence was slightly less than the true prevalence estimate.

Predictive values describe the likelihood that a test result correctly identifies the animal's condition. **Predictive Value of a Positive Test** = the proportion of animals that test positive (a+b) that are truly affected = a/(a+b). **Predictive Value of a Negative Test** = the proportion of animals that test negative (c+d) that are truly unaffected = d/(c+d). In table 2, where the true prevalence of paratuberculosis is 10%, the positive predictive value of the Dot ELISA is 18/23, or 78%. This means that, when the Dot ELISA test is positive, the probability of paratuberculosis in a cow from one of these herds is 78%. There is a 22% chance that she doesn't have paratuberculosis. If the Dot ELISA is negative, the probability that the cow truly does not have paratuberculosis is 98%.

Predictive values are strongly influenced by prevalence. Low prevalence causes positive predictive value to be low. Disease control programs lower prevalence and positive predictive value, and increase the number of non-diseased reactors culled. Sometimes, diagnostic tests are evaluated in a high prevalence population. Be sure that the authors report the sensitivity of the test as well as the positive predictive value, which will be high simply because the prevalence of disease is high. There is more than one paper where it is the positive predictive value, and not the sensitivity or specificity, that has been reported as a demonstration of test efficacy. READER BEWARE!

Interpreting Results from Multiple Tests

Several ways to increse the predictive value of tests are possible. One way is to use tests in combination. If the test animals are screened first, and only strong suspects are tested, then the prevalence of disease is quite high, and hence, so is the positive predictive value. This is what happens in serial test interpretation. All animals are tested with one test. Animals with positive test results are tested with another test. Animals are classified as test positive only if both test results are positive. With serial testing, the specificity and positive predictive value are high. This strategy is applied in many disease control programs. In TB testing, the caudal fold test is applied first. It is relatively sensitive (good at picking out animals with disease) but not too specific (many animals without tuberculosis, some of which have been exposed to other mycobacterial antigens, may give a positive test). When a test positive animal is found, the test is followed by the comparative cervical test, which is more specific. Only animals which test positive to both tests are considered to have TB.

Parallel interpretation of tests requires that all animals be tested with two different tests. If an animal has a positive result on either, or both tests, then it is considered to be positive under parallel interpretation. The result is a test with high sensitivity but poor positive predictive value.

When two test are used in combination, the overall efficacy of the combination should be demonstrated in the literature. The combined sensitivity and specificity will likely be less than that calculated from the two tests used separately. This is because, within diseased individuals, both tests are likely measuring deviations from normal physiology that are not independent events, but related to the same underlying disease process.

Clinical Agreement

The measuring of agreement among test can be useful to determine how consistent or repeatable a test result is. For instance, you measure the degree of agreement between yourself and your partner on rectal palpation to determine estrus. Or you could measure your agreement with yourself. The results might surprise you!

Use "agreement" instead of sensitivity and specificity when: there is no gold standard, or you are comparing

similar tests, neither one of which can be considered a gold standard, or you are comparing agreement between two clinicians, or you are comparing one veterinarians's repeated examination of the same patients. Clinical disagreement arises when there is a variation in diagnostic criteria between examiners, when there is variation in skill, acuity, or bias between examiners, when there is biological and temporal variation in the patient from one evaluation to the next, or when there is variation in the conditions of the exam in the two situations being compared. Therefore, it is very unlikely that agreement between two tests will ever be 100%.

The Kappa Statistic is used to measure agreement (Kramer 1988). If you compare the results of two diagnostic tests on the same cows, you would expect them to agree some of the time by chance alone. For example, in one evaluation (Woodruff *et al* 1991), in which the prevalence of paratuberculosis (as determined by fecal culture) was 5.5%, results between Dot ELISA and fecal culture agree 97% of the time. In other words, in 97% of samples tested by both methods, the same result was obtained. However, the agreement expected by chance alone in that study was 90%. Kappa measures how well two tests agree after accounting for chance agreement. The Kappa value obtained between fecal culture and Dot ELISA in this study was 68%. This means that 68% of the potential agreement available beyond chance was actually achieved. From this it can be seen that simple observed agreement overestimates the agreement between two tests, especially when the disease condition is rare. A Kappa of 68% between fecal culture and Dot ELISA is fairly good. Generally, if two tests are measuring the same thing, you would expect the Kappa value to be at least 40% to 50%.

Is the Test Both Accurate and Consistent?

Every test should be described in terms of its accuracy and consistency. Accuracy is the ability of the test to measure the substance being evaluated, e.g., serum calcium, liver enzymes, or antibody concentration. Consistency refers to the ability of the test to give the same results on the same sample each time the test is performed. Consistency of test performance on the same sample is often ignored in the literature of test evaluation, but it is just as important as accuracy. A test is diagnostically useful when it is both accurate and consistent.

Most tests do not measure the target disease itself, but rather a physiologic response to the disease. While the test itself may be accurate and consistent, it may systematically miss certain cases of the target disorder. Many serologic tests measure serum antibodies, a by-product of infection, rather than the presence of an infectious agent itself. Many other factors, including the timing of the serum sample relative to infection, and factors specific for the disease itself, can cause results to be in error, even though the test is accurately measuring the antibody level present, and the result is consistent. An example is an antibody assay for BVD, which will not identify persistently infected animals. The test result may be accurate and repeatable, but the test is not appropriate to the situation. The best way to avoid this kind of error is to know the test well, and to understand the pathophysiology of the target disorder.

Potential Impact on Case Handling

The potential gain in certainty before and after testing must be great enough to justify the effort. If you are fairly sure that the target disease is present (or absent), only a very good test (high sensitivity and specificity) will have much influence on how you proceed. On the other hand, if disease prevalence in the group you are testing is about 50%, your uncertainty is greatest, and a diagnostic test will have the most impact on the handling of the case. The pre-test probability of disease is the same as the true prevalence. In order to make rational interpretations from diagnostic tests, you will need some estimate of disease prevalence. Prevalence estimated from your own experience is more relevant to your own situation than a textbook reference. When a test is chosen, if its sensitivity and specificity are known, and the pre-test probability of disease can be estimated, then a two by two table can be constructed. A hypothetical table, based on testing for paratuberculosis in subclinically affected dairy herds, is shown (Table 3). In this example, using an ELISA test, the true prevalence of disease is 10%, the sensitivity of the test is 73% and the specificity of the test is 85%. The goal is to calculate the post-test probability of disease for both positive and negative test results. If you know sensitivity, specificity and prevalence, then you can calculate post-test probability of disease:

Post-test probability of disease following a positive test = Positive Predictive Value

Post-test probability of disease following a negative test = the proportion of test negative animals that actually have the disorder = (1 – Negative Predictive Value)

Table 3. Utility of an ELISA test to diagnose subclinical Johne's Disease (unpublished data)

	Johne's Present	Johne's Absent	Totals
ELISA Positive	73	135	208
ELISA Negative	27	765	792
Totals	100	900	1,000

In this example, the post-test probability of disease following a positive test = positive predictive value = 73/208 = 35%. The post-test probability of disease following a negative test = 1 – negative predictive value = 1 –

(765/792) = 3%. In the situation of extreme prior probability (the clinician is at least 90% certain that the disease of interest is either present or absent before he runs the test), only a test with very high sensitivity and specificity will result in an appreciable gain in certainty before and after testing. In this case, raising the certainty of disease to 35% following a positive test result is not solid enough information for a veterinarian to recommend drastic action such as culling a cow. A better test should be sought, or a second test applied, keeping in mind the statement made earlier about the use of tests in combination. (Remember that a second serologic test using a similar antigen preparation is not independent, and a literature evaluation of the two in series should be sought before interpretation in that manner.)

A good discussion of the effect of pre-test probability on the utility of a diagnostic test can be found in a paper by Kelton (Kelton 1990), in which he discusses the effect of historical information and rectal palpation findings on the decision to treat open cows with prostaglandin to induce estrus. Kelton estimates the sensitivity and specificity of rectal palpation for a function corpus luteum to be 83% and 53%, respectively. The diagnostic objective is to discriminate effectively between cows with a functional corpus luteum (CL) that will respond to treatment, and those without. The effect of historical information on the utility of rectal palpation findings is summarized in Table 4. If it is known from the history that an open cow was in standing estrus 24 hours ago, then the probability of a functional CL is less than 10%. After palpation it rises to 16% in cows that the clinician palpates and finds a CL, and falls to about 3% in cows that he palpates and does not find a CL. In either case, the post-palpation probability of a responsive CL remains low, and palpation is not warranted, based on the lack of information gained. On the other hand, if the cow being presented for palpation was observed in standing estrus 10 days ago, then the post-palpation probability of a responsive CL is at worst 74%, suggesting that, given the small cost of a dose of prostaglandin, the cost of palpating a cow, and the small consequences of treating an open cow in the absence of a functional CL, the best option is to treat all cows observed in standing estrus 10 days ago, without palpating first. In the absence of historical information about previous heats, the information gained from palpation becomes most useful, tipping the balance in favor of, or away from, treatment, depending on palpation findings. Cows with a palpable Cl have about 3 chances in 4 of responding by coming into estrus, while cows without a palpable CL have only slightly better than 1 in 4 chance. The point being illustrated is that information used to select cows for prostaglandin treatment comes from two sources, history of estrus behavior and rectal palpation findings. Historical information changes the pre-test probability of response to prostaglandin treatment, and is valuable in selecting cows for treatment. The additional information provided by rectal palpation is most easily justifiable in the absence of historical information, when the uncertainty of response to therapy is greatest, at about 60%. However, neither test is 100% accurate. Each strategy must be weighed in terms of the benefits to the producer, and with due consideration to the costs involved with diagnosis and treatment.

Table 4. Utility of rectal palpation to select open cows for prostaglandin treatment (Kolton, 1990)

Pretest Probability of a Functional Corpus Luteum	Post-Test Probability of a Functional CL based on positive rectal findings	Post-Test Probability of a Functional CL based on negative rectal findings
10%	16%	3%
60%	73%	32%
90%	94%	74%

Summary

In summary, some key points deserve emphasis. It is important to answer some key questions when considering the adoption of a new or unfamiliar diagnostic test:

Does the test work in a population similar to the animals I see in my practice? Several factors, including disease prevalence, and host factors, including age, breed, sex and type (dairy or beef) may affect the performance of a diagnostic test. Be sure that the test is described under conditions that are similar to your own.

Can the test distinguish between healthy and diseased animals? Be sure the test sensitivity and specificity are described, and the results are consistent, as discussed above. If no gold standard exists, at the very least, agreement beyond chance should be described.

Can the test discriminate between animals with similar clinical signs but different target diseases? Generally in clinical practice, you will be using a test to diagnose a particular illness from other syndromes which present a similar clinical picture, or affect the same body system. Test specificity, in particular, will be affected by the composition of the group without the target condition.

Can the test diagnose animals in various stages of disease? Ideally, a test should be evaluated in a broad spectrum of cases, from acute to chronic, from mild to severe. If there are limitations, be sure they are described.

Will the results of the test change the way I handle this case? Finally, and most important, is the need to demonstrate the utility of the test. If the outcome cannot be improved by the use of a test, then the test is not needed. This is a social and an economic decision as much as it is an evaluation of the test. The outcome can only be improved if the test results in a significant gain in certainty after it is applied, if the cost of testing is affordable, relative to the cost of not testing, if the consequences of false positives and false negatives are not too great, and if the test offers significant advantages over other means diagnosis.

References

Fletcher, R.H., Fletcher, S.W., Wagner, E.H. 1988. Clinical epidemiology: the essentials. Second edition. Williams and Wilkins, Baltimore, 1988. Kelton, D.F. 1990. Using all available information to select cows for prostaglandin treatment. Soc. for Therio. Proc. Ann. Mtg. pp. 125-132. Kramer, M.S. 1988. Clinical epidemiology and biostatistics. Springer-Verlag, New York. Sackett, D.L., Haynes, R.B. and Tugwell, P. 1985. Clinical epidemiology: a basic science for clinical medicine. Little, Brown and Company, Boston and Toronto. Woodruff, T.S., Shulaw, W.P., Bech-Neilsen, S., Hoffsis, G.F., Spangler, E., Heider, L.E. 1991. Serodiagnosis of bovine paratuberculosis by use if a dit enzyme-linked immunosorbent assay. Am. J. Vet. Res. 52: 217-221.

Abstracts

Citrinin as a possible cause of the pruritis, pyrexia, haemorrhagic syndrome in cattle

I. B. Griffiths, S. H. Done

Veterinary Record (1991) **129**, 113-117

An outbreak of the pruritis, pyrexia, haemorrhagic syndrome affected eight of a herd of 175 cows which was divided into two groups of 115 and 60 according to yield. There was no difference in management between them but cirus pulp pellets were fed only to the larger group in which the eight cows were affected. Silage, which had been made without the use of additives, was also fed to both groups. The citrus pulp was visibly mouldy and contained 30 to 40 parts per billion of citrinin. Signs of the syndrome occurred within three days of the cows starting to ingest the citrus pulp, which was fed for 21 days, and the last case occurred six days after the feeding of citrus pulp ceased. Five calves whose dams had been fed citrus pulp were subsequently born with superior prognathism. In contrast to the eight cows that developed the syndrome only one out of 68 heifers which were fed larger quantities of citrus pulp for 10 days developed mild signs of the syndrome and then recovered, suggesting that older animals may be more susceptible. The clinical signs, gross pathology and histopathology are described and compared with those of previous outbreaks. Mycotoxins, particularly citrinin, were strongly implicated as the cause of this outbreak.

Treatment and control of an outbreak of fat cow syndrome in a large dairy herd

A. H. Andrews, R. Laven, I. Maisey

Veterinary Record (1991) **129**, 216-219

An outbreak of fat cow syndrome occurred in a herd of 300 Friesian and Friesian/Holstein dairy cows calving predominantly between January and May. The herd came in off grass in good condition despite a long and hot summer. The dry cows received a diet of grass silage, brewing waste and minerals until the end of December, but the grass silage was butyric and was partially replaced by maize silage. By January 23, 16 of 70 calving cows (23 per cent) had appeared to suffer milk fever. Subsequent blood tests revealed that the cows may have been ketotic, and clinical and post mortem examination showed that they were probably suffering from fat cow syndrome. The freshly calved sick cows were treated with glucose, and corticosteriods were injected every second day into those which remained ill. The cattle had received a high energy diet, but the cows still to calve were placed on a diet of low metabolisable energy (77 MJ/cow) but adequate levels of undegradable protein. The problem was associated with a possible clostridial infection in two cows and with reduced fertility.