FACTORS IN SELECTING SERUM SAMPLES FOR USE IN DETERMINING THE POSITIVE/NEGATIVE THRESHOLD (CUT-OFF) IN ELISA

R. H. JACOBSON
New York State College of Veterinary Medicine, Cornell University, Ithaca, New York, USA

Abstract

FACTORS IN SELECTING SERUM SAMPLES FOR USE IN DETERMINING THE POSITIVE/NEGATIVE THRESHOLD (CUT-OFF) IN ELISA.

The threshold (cut-off) that defines whether a test result is seropositive or seronegative is calculated by testing serum samples from a subpopulation of animals that is assumed to represent the target population in all aspects. For this proposition to be true, it is essential to consider the variables in the target population that must be represented in the subpopulation. Without representation of the variables in the subpopulation, it is likely that the cut-off selected for the test will be errant and will misclassify animals as to their infection status. The purpose of this paper is to identify a few of the principal variables that need to be taken into account when selecting a subpopulation of animals for test validation.

1. INTRODUCTION

The threshold (also known as the cut-off) is the unit of activity in a serodiagnostic test above which animals are classified as positive and below which they are considered negative. Serologic antibody activity is often used to infer whether an animal is infected or uninfected with a particular agent of disease. The consequences for an animal or herd may be quarantine or depopulation of the herd as occurs in disease eradication programs or the data may be used in disease control and herd management strategies. The accuracy of the cut-off, therefore, has considerable herd health and economic implications. Placement of the cut-off on the scale of antibody activity for a disease agent is the basis for calculation of diagnostic sensitivity and specificity [1,2]. Diagnostic sensitivity and specificity are usually viewed as evidence of the test's performance when it is being considered for diagnostic uses. The selection of the cut-off and the ability to determine the appropriateness of the data underlying the cut-off have highly consequential ramifications. Cut-offs must be selected with utmost care.

2. FACTORS IN SELECTING A CUT-OFF

The cut-off is determined by applying the test to a selected subpopulation of animals of known infection and vaccinal status. The underlying assumption in any assay is that the subpopulation of animals upon which the cut-off was based is representative in every respect of the targeted population of animals for which the assay is intended. Because the targeted population for the test may be all of the cattle in a given country, it is apparent that these animals may represent a wide variety of genetic, environmental, and nutritional variables which may impact on a test's sensitivity, specificity, accuracy and precision. The first consideration in selecting a subpopulation for determining the cut-off in an assay is to obtain a group of animals that is unequivocally infected and another group that has never experienced an infection of the disease agent in question.

2.1. Infected animal group

Historically, the infection status of animals used as subjects in assay validation has been either known or inferred. Isolation of the organism from the animals that make up the infected group is proof that they are experiencing the infection at the time the serum sample was obtained. A second indication of infection is histopathology, which may be pathognomonic resulting in a strong inference that the animal was infected with the agent in question; for other agents, however, histopathology renders only a presumptive diagnosis. A third and less compelling inference of infection is the positive result of another serological test. This method is among the least useful because all of the
error in determining the sensitivity and specificity of this so-called "gold standard" test is carried over to the test being validated. A fourth method is clinical diagnosis without any laboratory evidence of infection; this is the least useful in designating an animal as infected or not. Experimental infection of animals with the agent in question is another means of assuring exposure to the agent. Route of administration and dose of the inoculum are variables that may elicit antibody responses that are qualitatively and/or quantitatively much different that those elicited by natural infection. For instance, inoculation of massive numbers of organisms intramuscularly may result in a completely different antibody response than when a tick vector introduces a few organisms into the host percutaneously. Conversely, if prior studies indicate that experimental and natural infections elicit antibodies in ways that are indistinguishable from natural infections, experimental infections may be the method of choice to develop a subpopulation of animals for test validation. It would be important, however, that environmental and genetic considerations not be overlooked in the choice of such experimental animals because these factors can have a profound effect on antibody responses or, through non-specific factors, a test's performance.

Among infected animals, the antibody response can be considerably different between individuals depending upon the stage of infection, the species of organism that elicits that response, and the age of the host. If the serum sample is taken at two to three weeks after exposure of adult animals to an acute infection that is known to elicit a strong antibody response, a relatively high level of antibody would be expected. However, if the sample was taken from such animals just after exposure to the organism or at 6 months after recovery from the infection, the antibody level may be much lower. Cattle exposed as neonates to the causative agent of paratuberculosis, Mycobacterium paratuberculosis, may not develop detectable antibody responses for several months or even years despite isolation of the organism from the faeces. Indeed, even adult animals will be anergic for months to this organism after initial exposure. The dynamics of exposure to an agent and antibody levels in serum that the agent elicits require that animals selected for the subpopulation be representative of all stages of the targeted infection in order to reflect the realities of the population at large. This requires knowledge of the host-parasite relationship at both the individual animal and population levels.

2.2. The uninfected group

Selection of a group of organism-negative animals as determined by isolation procedures, is not sufficient to assure that antibody to the agent is not present in the animal. The animal may have eliminated the organism and still have varying amounts of specific antibody in its serum, or the method to detect the organism may have lacked sensitivity. Animals negative on other antibody tests are only presumptively uninfected if they do not have detectable antibody; it is possible that the reference test is analytically less sensitive than the test being validated. Correlative test results by themselves, therefore, are not the best means of determining the animal's prior infection history and may be misleading in classifying animals as uninfected.

Animals from geographical areas that are known to be free of the disease in question are usually good candidates for the "uninfected group" in assay validation assuming that: 1) there is no history of disease in these animals clinically, 2) laboratory tests have always been negative during the last several years, 3) the herds are "closed" so that no movement of infected animals into the herds has occurred, 4) adjacent farms are also known to be infection/disease free, and 5) the environment, genetics, and nutrition of the animals is similar to that of the target population. This combination of criteria has a high likelihood of providing an uninfected group of animals that can be used to establish the specificity of the assay.

An additional factor that must be considered is the possibility that closely related organisms, endemic in the area, may induce antibody that is cross-reactive in the test being validated. The uninfected animals may then give results that raise the cut-off artificially and make the test not as diagnostically sensitive as it could have been were cross-reactive antibodies considered in validation of the assay.
3. CUT-OFF SELECTION

Assuming that the subpopulations of infected and uninfected animals have been properly selected, samples from these animals have been run in the ELISA, and variation within and between runs of the assay were minimal, the normalized data for the test samples then can be evaluated for selection of the optimal cut-off that will minimize false responses. (Normalization in this case means that optical density readings of the test samples are converted to percent of the high positive serum control in each plate as indicated in the FAO/IAEA kit instructions).

Diagnostic sensitivity and specificity in any assay is a trade-off for a range of ELISA values such as percent of positive control (PP). As the location of the PP cut-off point increases on the continuum from 0 to 100%, the sensitivity decreases and the specificity increases. If one sums the sensitivity and specificity values for each interval of PP points on the continuum of cut-off points (such as for PPs of 10%, 20%, 30%, etc.), one of the cut-off points will give the greatest sum of the sensitivity and specificity values: this is considered the most accurate cut-off point for that assay [1]. A receiver operator characteristic (ROC) curve is a similar way of describing the accuracy of a test over a range of cut-off points [3, 4] It serves as a nomogram for reading off the specificity that corresponds to a given sensitivity for the test. From this curve, a point can be found that describes the greatest accuracy for the assay and allows the optimal sensitivity and specificity to be read from the x and y axes. These methods provide a statistically sound means of determining a cut-off for an assay.

A commonly used method for establishing the cut-off is to select the mean plus two or three standard deviations of PPs from all uninfected animals. This method assumes that the PPs for all uninfected sera will form a Gaussian (normal) curve when plotted as a frequency distribution of PP values. Because such data are virtually never normally distributed, this method lacks validity statistically. Alternatively, the non-parametric method of two times the mean of PPs of the sera from uninfected animals has been used as the cut-off which is a reasonable way to establish that approximately 95% to 99% of the uninfected animals in the population will fall in the test negative group. A limitation of methods that determine cut-off only on the basis of the sera from uninfected animals is that they do not take into consideration the degree of overlap of PPs between uninfected and infected animals.

The actual cut-off selected is usually mandated by practical realities and not statistical optimization. Using ROC curves may give the statistically optimal cut-off for an assay for FMD. In practice, however, the cut-off may need to be lowered to assure that the test sensitivity approaches 100% in order to not misclassify any infected animals as negative-false negative results for a disease with devastating economic and animal health consequences are not acceptable. With such a cut-off, test results for some uninfected animals will be interpreted as positive, but will be falsely positive because the specificity of the test has been lowered as a consequence of lowering the cut-off. Alternatively, for diseases of low morbidity and low mortality, the consequences of false positive results, and the slaughter of non-infected animals, is not acceptable. In this case, the cut-off may be raised to assure that uninfected animals are not miss-classified serologically. The result of raising the cut-off is that some infected animals in the population will be misclassified as uninfected.

If a test is unable to detect antibody in about 40% - 50% of the infected animals when the specificity of the test is set at 99%, another strategy for interpretation of results is needed. This is the case with paratuberculosis in cattle. It may be necessary to test entire herds and determine the percentage of animals that fall into low, moderate, and high risk categories based on multiple cut-offs. Based on faecal culture data, all animals that fall below a selected low PP cut-off would be considered at low risk of having an infection whereas animals with PP values above a high PP cut-off would be considered as having a high risk of being infected; between the two cut-offs would be the animals of moderate risk of infection. Using multiple cut-offs to define risk levels gives the producer an indication of the status of the herd: if, for instance, 98% of the animals fall into the "low risk" category, and only 2% in the moderate risk category, the probability that this herd has paratuberculosis is very low. Alternatively, if 60% of the animals are in the low risk but 40% are evenly distributed among the moderate and high risk categories, the chances of paratuberculosis in this herd is very high. So, even for a test with poor performance characteristics, useful data can be obtained in the quest to classify herd status.

4. SIZE OF SUBPOPULATION REQUIRED FOR DETERMINING THE CUT-OFF
Standard tables are published in many epidemiology and statistical texts that give the sample size required to detect various minimum levels of infection in different sized herds at various confidence levels. These tables give some guidance in selecting the size of the known uninfected and known infected groups of animals for establishing a cut-off that will be valid for the entire target population. The guidance is imperfect because sampling theory assumes that the number and effects of unknown variables that contribute to the continuum of PP values for infected and uninfected populations of animals is fully represented in the sample. Indeed, when one tries to list the variables and the variation within each variable that must be considered in selecting the reference animals, the task becomes daunting. For instance, an infection process usually results in antibody production that is a dynamic variable as a function of time. Similarly, the antigen used in the test system may appear to be specific for a group of animals but then another group may give consistently elevated values attributable to some external unknown factor. These two examples represent but two of many such variables, all of which may affect PP values. For this reason, it is not possible to establish the set of underlying assumptions for a test's probable performance when applied to a subpopulation for the purpose of establishing a valid cut-off. It is, therefore, desirable to make an error on the high side when determining the number of infected and uninfected animals by which the test will be evaluated. Generally, the experience of ELISA developers has shown that several thousand samples from uninfected animals and at least several hundred from known infected animals, that are good representatives of the population at large, will allow development of a cut-off that "will stand the test of time" [4].

REFERENCES