SAMPLING SCHEMES TO DETECT VERY LOW CONCENTRATIONS

Terry C. Nelsen

Recommended Citation

SAMPLING SCHEMES TO DETECT VERY LOW CONCENTRATIONS

Terry C. Nelsen, USDA-ARS-MWA, 1815 North University Street, Peoria, IL 61604

Abstract

Many mycotoxins and certain drug residues can be important at very low concentrations in feeds and foods. Government regulatory agencies establish maximum acceptance concentrations (Action Levels) to avoid proven effects of known toxins and provide Advisory Levels where effects are not yet well established. The Action Levels for the aflatoxins range from 5 to 300 ppb (ng/g) depending on the particular toxin and the intended use of the feed or food. Other naturally occurring mycotoxins, such as DON, zearalenones, and fumonisons, have advisory levels in the range of 1 to 5 ppm (μg/g). Chemotherapeutic agents in feeds must be measured in ppm or even ppb. Common acceptance sampling techniques can usually be used for liquids, flours, meals, etc. A problem arises, however, when the substance to be measured is concentrated in individual kernels of grain. In some cases, a single "hot" kernel contains enough toxin to contaminate thousands of clean kernels. The common sampling techniques for large lots (such as truck, barge or ship loads) can be unreliable at concentrations around the Action or Advisory Levels. The problem will continue to arise as modern testing instruments become more sensitive at lower concentrations. This paper will review some biases of current sampling procedures and discuss methods of changing the sizes of the types I and II errors.

Key Words: mycotoxins, aflatoxin, deoxynivalenol, sampling, OC curves, action levels

1. Introduction

Mycotoxins are naturally occurring toxic products of fungal metabolism. The toxins produced by different fungi have different chemical structures and different effects on humans or livestock. They can adversely affect human health and cause economic loss for livestock producers. Mycotoxins have been found on just about

all agricultural commodities, including grains, nuts, meats, dairy products, fruits and vegetables. Anything that will support mold has a chance to be contaminated. These toxins can cause problems at very low concentrations, parts-per-million(ppm), or even parts-per-billion (ppb), so sampling, identifying and measuring toxins in a truck, barge or shipload can be a problem. Common acceptance sampling can be adequate for liquids in equilibrium, or in flours or meals that have been ground or milled, but may not be as useful when the toxins are concentrated in individual "hot" kernels which are randomly scattered throughout the load. Recent technological developments in testing procedures allow us to examine some different sampling schemes.

2. Mycotoxins.

All molds do not produce mycotoxins. The fungi most responsible for mycotoxin are *Aspergillus, Fusarium,* and *Penicillium* (Betina, 1984) which can produce toxins under certain conditions. Toxins are most common in crops that have been subjected to stress from too much or too little moisture, high temperatures, or insect damage (Davis et.al.1980, and Shotwell, 1977). These same stresses plus mechanical damage can cause problems in stored grains. The presence of mold is not always an indicator of mycotoxin but is a sign that care should be taken. Also, molds can grow in the field and then die off after the stress is relieved but leave a deposit of a toxin behind them in the grain. Most mycotoxin contamination occurs in damaged grain and can be eliminated by screening damaged kernels.
 Why, then, are we concerned about mycotoxins? First, recent research (Smart, et.al. 1990) has found deposits of mycotoxin in intact kernels. Second, technological advances are allowing us to measure the amount of mycotoxins faster, cheaper, and more accurately. Finally, evidence is gathering that extremely low levels of some of the toxins can have adverse effects on human or animal health (CAST, 1989).
    Many mycotoxins have been described and are currently being studied (Maracas and Nelson, 1987). I will briefly describe only a few of the better known plant pathogens and storage molds and some of their effects (see NAS, 1978, and Richard and Thurston, 1986).
    The tricothecenes are common plant pathogens. Deoxynivalenol (DON or Vomitoxin), T-2 and nivalenol have been shown to cause digestive disorders in pigs ranging from internal hemorrhage to diarrhea to simple appetite suppression at concentrations of 1 to 100 ppm (ppm = parts per million = $\mu g/g$). Zearalenone chemically resembles estrogen and at 50-100 ppm has caused a variety of reproductive disorders in pigs. The fumonisins are newly

described and as yet little known but may have effects at 1-5 ppm (Ross et.al. , 1991).

Aflatoxins B_1, B_2, G_1, G_2, M_1, and M_2 are storage molds that have caused liver and kidney damage and are carcinogenic at levels of 10 to 300 ppb (ppb = parts per billion = ng/g. One part per billion has been described as one drop in a tank car full.)

3. Action levels

There are few legal standards ("Action Levels") for determining a dangerous contamination level. The pathways of toxic effects are not well understood and accurate levels of the particular toxins have been difficult to measure. The U.S. Food and Drug Administration has action levels for aflatoxins of .5 ppb in milk, 20 ppb in foods, 100 ppb in breeding cattle and swine or poultry, and 300 ppb in feedlot cattle (CAST, 1989). The Canadian government has set an action level of two ppm for DON contamination. Until we learn more about these toxins, we will have to rely on "advisory levels" established by government or industry groups for particular toxins in each commodity. At one time some regulators thought of setting the action levels at "detectable limits" but this is no longer practical because the analytical chemists have developed and refined their instruments and measurement processes to where they can detect some toxins at the ppt (parts per trillion) levels. At ppt levels all commodities can probably be shown to have a residual level of at least one toxin. At ppt levels the toxins should not really be harmful. Humans are protected somewhat by eating a varied diet. Livestock are most at risk because most or all of their diet can be from one source and the toxins can accumulate in their bodies.

4. Detection Problems

Liquids, and to a lesser extent, flours or meals, should have the toxins spread homogeneously and therefore can be tested by conventional means. In whole grains, however, the toxins are concentrated in "hot" kernels which are randomly scattered throughout the truck, car, barge, ship or bin.

The toxin tends to be produced in a wounded or injured kernel and then spreads to surrounding kernels. (Bennett et.al. 1988, Smart et.al. 1990, and Wicklow et.al. 1987).

The problem in detecting mycotoxins overlaps with a problem in the pharmaceutical industry. Certain drugs which are administered to livestock in their feeds are delivered in individual carrier kernels. Kernels of grain are soaked in a concentration of the desired drug and then are mixed with the feed. At very low levels of drug concentrations these carrier

kernels are spread at random in the feed. Finding and counting the number of carrier kernels is a problem for quality control for the buyers and sellers of the feed and can be a problem for government or industry regulatory agencies that want to monitor drug use.

Mycotoxin levels have traditionally been measured by TLC (Thin Layer Chromatography) which requires hours for sample preparation and purification. Modern developments in HPLC (High Performance Liquid Chromatography), ELISA (Enzyme Linked Immunoabsorbant Assay), and Affinity Columns have expedited analyses and have become more specific for each mycotoxin. Also, test kits from assorted manufacturers now allow us to screen for a specific toxin at a specified level in under an hour. Thus traditional sampling schemes which may have been limited by the time and effort required in the measurement process can now be reevaluated in light of modern technology.

5. Sampling procedures and assumptions.

We will make the following assumptions for this exercise. 1) the contaminated ('hot') kernels contain aflatoxin at 36,700 ppb. Wicklow et.al.(1987) found *A. flavus* in wounded kernels of corn at a range of 13,200-82,000 ppb with a mean of 36,700 ppb. 2)the hot kernels are distributed at random in the container (i.e., truck, car, barge, ship, or bin) being tested. We assume that the damage was done to the kernels before they were loaded. There is no reason to assume "hot spots" or gradients within the storage container. 3) Because of the random distribution, we can use binomial probabilities to estimate our chances of finding contaminated kernels. 4) We will use the FDA action level of 300 ppb of aflatoxin in feed for finishing beef cattle. The results from this study should be applicable in principle to other action levels in other toxins. 5) The weight of a kernel of corn is .3 g.

We will also assume that once we have gathered our sample of whole kernels, we can then grind or mill our sample and analyze the sample as a homogenous mixture.

This analysis was done for both DON measured in ppm and for aflatoxin measured in ppb. The overall results and conclusions were similar so we will discuss only the aflatoxin results.

The Federal Grain Inspection Service recommends a sampling rate of 500 g per 2000 bushels (FGIS, 1986). For this exercise we will consider a load to be 2000 bushels of corn. A bushel of corn weighs approximately 25.4 kg. Our load of corn therefore weighs 5.08×10^7 g and contains 1.693×10^8 kernels of corn.

A hot kernel weighing .3 g and contaminated at a level of

3.67 x 10^4 ng/g therefore contains 1.101 x 10^4 ng of aflatoxin. If the load is contaminated at a level of 300 ng/g then it contains a total of 300 ng/g x (5.08 x 10^7 g) = 1.524 x 10^{10} ng of aflatoxin contained in (1.524 x 10^{10} ng) ÷ (1.101 x 10^4 ng/kernel) = 1.384 x 10^6 hot kernels. The percent of kernels that are contaminated is therefore (1.384 x 10^6 kernels) ÷ (1.693 x 10^8 kernels) = .8 %.

For our 500 g sample to be contaminated it must contain at least (300 ng/g) x (500 g) = 1.50 x 10^5 ng of aflatoxin which will be in (1.50 x 10^5 ng) ÷ (1.101 x 10^4 ng/kernel) = 13.6 kernels. We are measuring whole kernels so at these levels a 500 g sample with 13 hot kernels is acceptable and a 500 g sample with 14 hot kernels is contaminated.

The probability of finding x or more hot kernels in a sample of size n where p is the proportion of hot kernels as compared to all kernels is:

$$\binom{n}{x} p^x (1 - p)^{(n - x)}$$

Calculations were made in 5 ppb increments from 100 to 500 ppb.

6. Results.

The plot of toxin concentration versus the probability of rejection (Figure 1) is called an OC (Operating Characteristic) Curve. The probability of rejecting a load with exactly 300 ppb of a toxin is about 50%. The section of the graph with the horizontal striping to the left of the vertical 300 ppb line is the area associated with false rejection of a clean load (Type I error). The vertically striped area to the right of the 300 ppb line is the area associated with accepting a contaminated load (Type II error). The horizontal dashed line is the 95% probability of rejection line. The 95% line crosses our OC curve at about 470 ppb. The objective of this exercise is to explore sampling strategies as they affect the sizes of the striped areas. This is a classic case of exploring type I versus type II errors complicated somewhat by integer restrictions imposed by the hot kernels.

The implications of a Type II error are obvious. A contaminated commodity may be used and can cause health problems to humans or livestock. A Type I error, a false positive, may seem to be less serious, especially where public health is concerned, but it can have serious economic effects. An entire

load which is condemned has not only lost all of its value as a
commodity but now has to be disposed as contaminated waste. When
both the Type I and Type II errors are substantial the industry
can suffer a loss of confidence and trust. A load can be tested
more than once with conflicting results. People will naturally
question each other and the entire testing system if the results
of tests cannot be considered reliable.

The first strategy to try is to simply take a larger sample.
Figure 2 shows the OC curves for samples of 500,1000, 1500, 2000,
2500, 3000, and 3500 g. The curves are not evenly spaced because
of the rounding to integer values. Obviously, some progress in
shrinking the areas of error is made by taking a larger sample.
Above 1500 g appears to be rapidly approaching the point of
diminishing returns. With a 1500 g sample the probability of
rejecting a 300 ppb load is still about 50% but the 95% line
crosses the OC for 1500 g at about 390 ppb.

Another strategy is to take multiple 500 g samples and
analyze them separately. The load can be rejected if any one
sample is positive. We can approach 95% probability of rejecting
a 300-ppb load if we take five separate 500 g samples and analyze
them independently (Figure 3). The price paid is obvious in
Figure 3, however. The area for the false positives is greatly
increased. If we use the five x 500 g scheme, we will have more
than 50% chance of rejecting loads that are in the safe area of
210-300 ppb.

Consider the strategy where the samples are not collected as
500 g units. In practice, smaller samples are gathered from
specified locations and then bulked together to form the 500 g
sample. During a loading process smaller samples are gathered
from the loading stream at designated intervals until the 500 g
size is reached. Therefore, the 500 g standard sample could
easily be gathered and analyzed as fractions of 500 g. For
example, five 100 g samples could be gathered and tested with a
test-kit procedure without major changes in procedures used to
gather the standard 500 g sample. Figure 4 shows the OC curves
for testing up to five 100 g samples and rejecting the load if
any one sample is positive. When Figure 4 is compared to Figure
3 we can see that we can achieve about the same ability to detect
a contaminated load with four 100 g samples as we could reach
with four 500 g samples. The difference is in the large increase
in the false positive area. With five 100 g samples we would
reject too many loads which would be below 300 ppb or even 200
ppb.

The logical next step appears to be examining combinations
of 100 g samples and plotting rejection strategies which can keep
the type II error at a minimum but somehow reduce the type I
error. One strategy is to analyze the five separate 100 g samples

individually and then reject the load if three or more of the samples are positive. This strategy is not really any improvement on the simple 500 g sample as we can see in Figure 5 when we compare s-1 to the 500 g standard. The s-2 strategy is more complicated but does improve the OC Curve. In the s-2 strategy we take the five 100 g samples and analyze them separately. If none or just one of the 100 g samples is positive we accept the load. If four or five of the samples are positive, we reject the load. If two or three of the samples are positive, we take five more 100 g samples. Of the total of ten 100 g samples we now reject the load if five or more are positive. We can see that the OC Curve for s-2 in Figure 5 is an improvement over the 500 g standard.

7. Discussion.

We have now shown that different sampling schemes can alter the probabilities associated with the type I and II errors. We could proceed with some more complicated sampling schemes with different sizes of samples and perhaps introduce schemes which consider sequences of positives and negatives. But before we proceed any farther we should stop and reconsider what we have done to this point in light of some other factors.

First, we can alter the OC Curves but what have we done to the confidence limits around these curves? Second, we have not considered the distribution of the toxins inside these hot kernels. Microbiological growth tends to follow a logistic curve. Wicklow et.al.(1987) did not report the distribution of toxins within kernels, but only the range and mean. What effect does the distribution of the toxins have on the OC Curves? Third, what is the distribution of toxin concentration in loads of grain? If we design a complicated sampling system, we should try to estimate how often we will have to resample. Finally, any final sampling recommendation will have to have an economic analysis associated with it. What will be the consequences of both type I and type II errors? These consequences should be calculated from the economic points of view of the sellers and buyers of the grain as well as from the health interests of the American public.

REFERENCES

Bennett, Glenn A., Donald T. Wicklow, Rodney W. Caldwell and Eugene B. Smalley. 1988. Distribution of Trichothecenes and Zearalenone in *Fusarium graminearum*: Rotted Corn Ears Grown in a Controlled Environment. J. Agric. Food Chem., Vol. 36, No.3,

1988, pp. 639-642.

Betina, Vladimír (Ed.), 1984. <u>Mycotoxins: Production, Isolation, Separation and Puification</u>. Developments in Food Science 8. Elsevier, New York.

CAST (Council for Agricultural Science and Technology), 1989. Mycotoxins: Economic and Health Risks. Task Force Report No, 116. November 1989, CAST, Ames, Iowa

Davis, Norman D., J. William Dickens, Robert L. Freie, Pat B. Hamilton, Odette L. Shotwell, Thomas D. Wyllie and J. F. Fulkerson. 1980, Mycotoxins: Protocols for Surveys, Sampling, Post-Collection Handling and Analysis of Grain Samples Involved in Mycotoxin Problems. J. Assoc. Off. Anal. Chem.,Vol. 63, No.1, 1980

FGIS (Federal Grain Inspection Service), 1986. Grain Sampling, Book I. USDA-FGIS. U.S. Govt Printing Office : 1989 - 241-784/80828

Marasas, W.F.O. and Paul E. Nelson. 1987. <u>Mycotoxicology</u>. Penn State Univ. Press, University Park, PA

NAS (National Acadamy of Sciences), 1978, Interactions of Mycotoxins in Animal Production, Proceedings of a symposium, July 13, 1978, Michigan State Univ. N.A.S. Washington, D.C.

Richard, J.L. and J.R.Thurston (Eds), 1986. <u>Diagnosis of Mycotoxicoses</u>, Martinus Nijhoff, Boston.

Ross, P.F., L.G.Rice, R.D.Plattner, G.D.Osweiler, T.M.Wilson, D.L.Owens, H.A.Nelson, and J.L.Richard. 1991. Concentration of fumonison B_1 in feeds associated with animal health problems. Mycopathologia 114:pp.129-135.

Shotwell, O.L., 1977. Aflatoxin in Corn. J. Am. Oil Chemists Soc., March 1977 (Vol. 54, pp. 216A-224A).

Smart, M.G., O.L.Shotwell, and R.W.Caldwell. 1990. Pathogenesis in Aspergillus Ear Rot of Maize: Aflatoxin B_1 Levels in Grains around Wound-Inoculation Sites. Phytopathology, Vol. 80, No. 12, 1990, pp. 1283-1286

Wicklow, D.T., B.W.Horn and O.L.Shotwell. 1987. Aflatoxin Formation in Preharvest Maize Ears Coinoculated with *Aspergillus Flavus* and *Aspergillus Niger*. Mycolgia. 79(5), 1987, pp. 679-682.
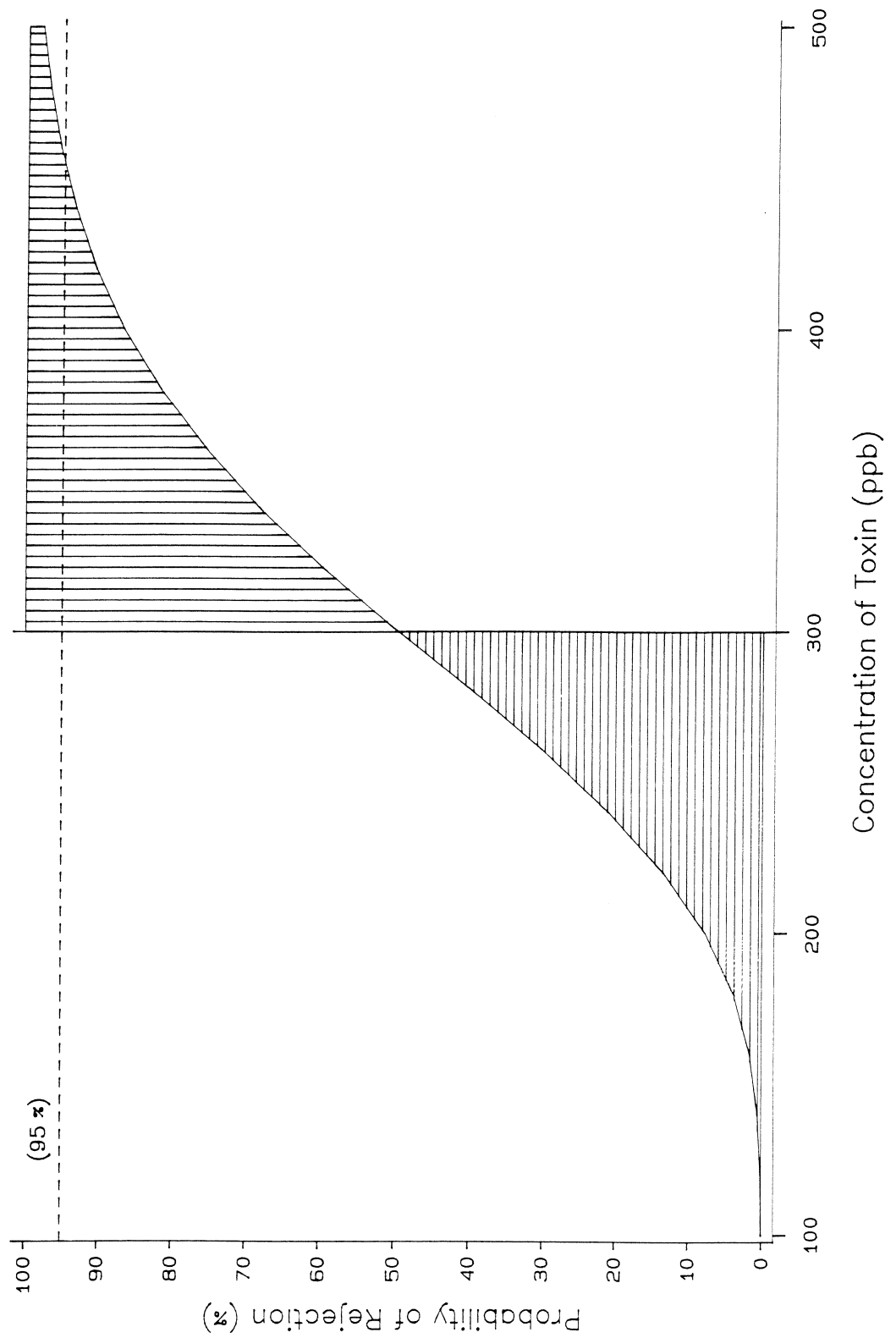
FIGURE 1: OC Curve for 500 g sample

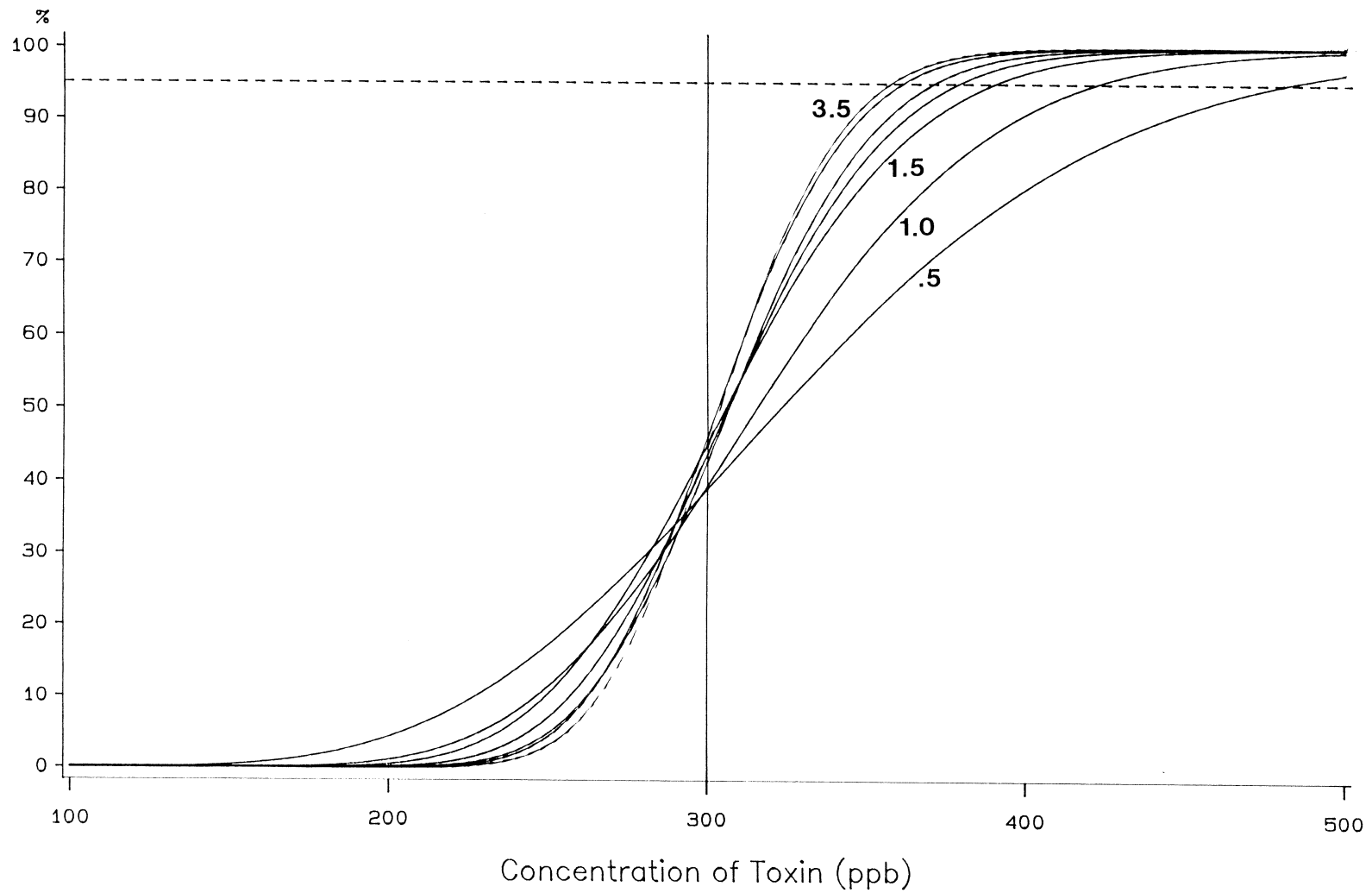# FIGURE 2: OC Curves for samples of size .5 to 3.5 kg

FIGURE 3: OC Curves for multiple 500 g samples
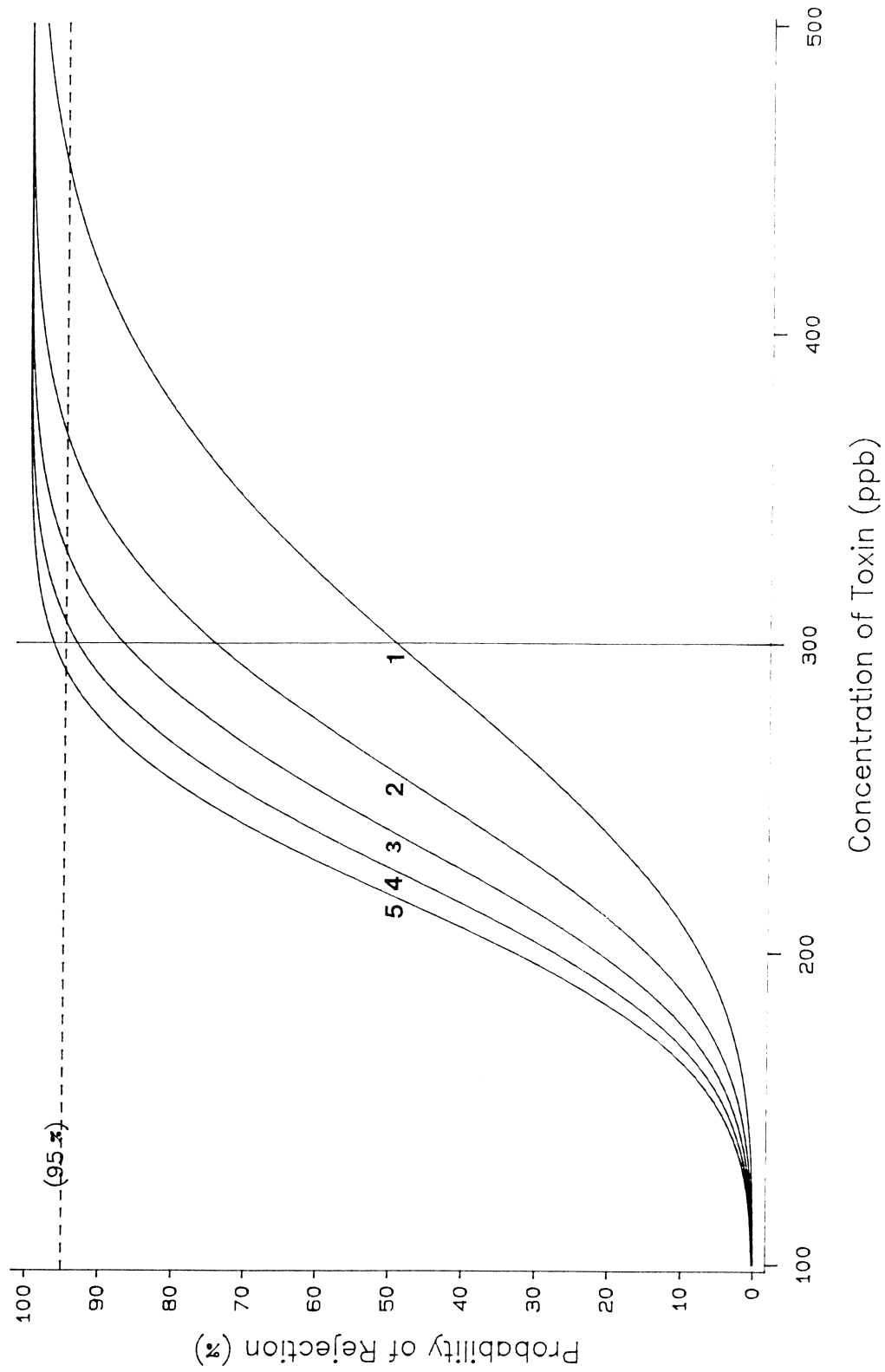Reject load if at least one is positive

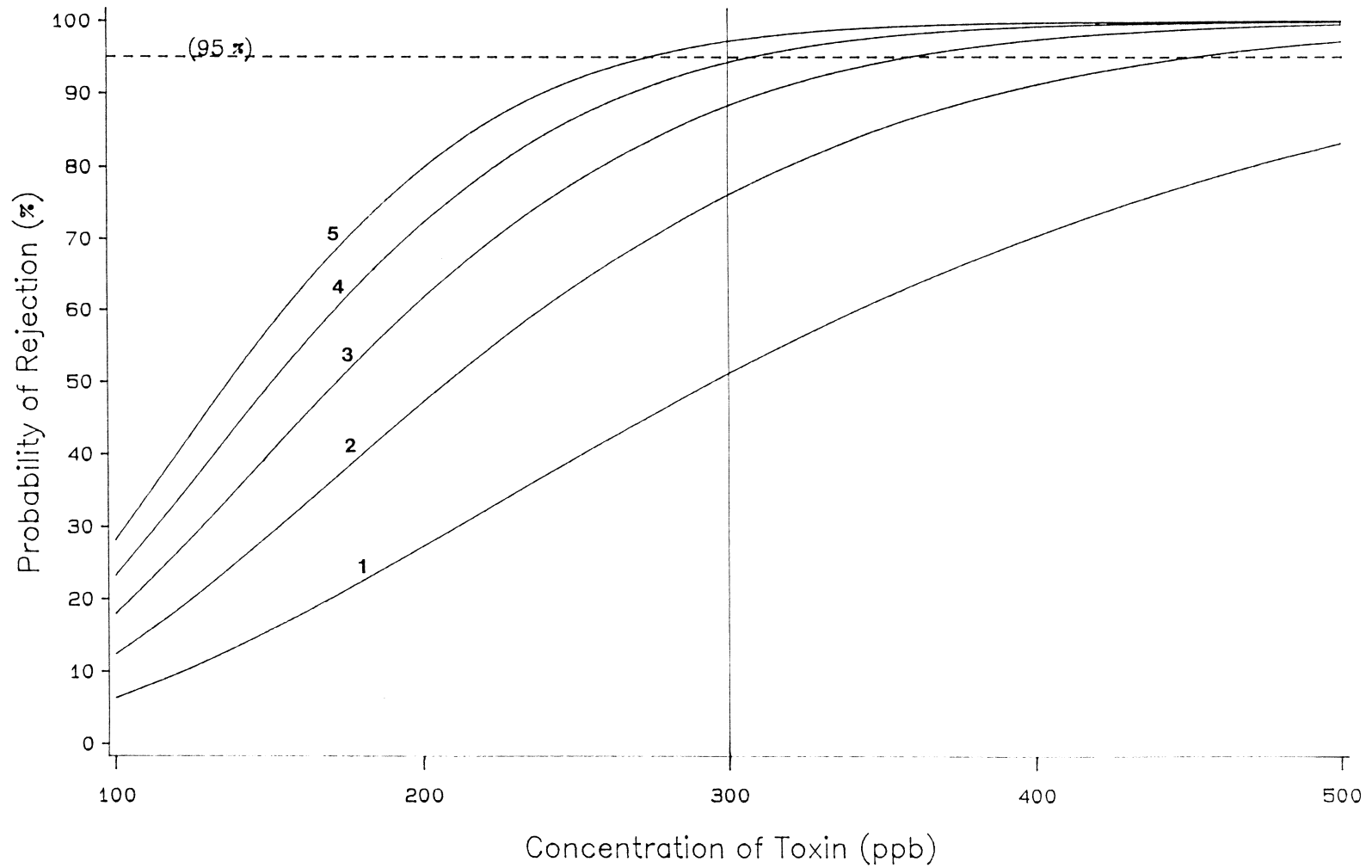FIGURE 4: OC Curves for multiple 100 g samples
Reject load if at least one is positive

FIGURE 5: Multiple 100 g samples:
Rejection strategy 100b