*Journal of Food Protection, Vol. 68, No. 6, 2005, Pages 1306–1313*

# Sampling Uncertainties for the Detection of Chemical Agents in Complex Food Matrices†‡

**THOMAS B. WHITAKER* AND ANDERS S. JOHANSSON**

*Agricultural Research Service, U.S. Department of Agriculture, Box 7625, North Carolina State University, Raleigh, North Carolina 27695-7625, USA*

MS 04-707: Received 27 May 2004/Accepted 17 October 2004

## ABSTRACT

Using uncertainty associated with detection of aflatoxin in shelled corn as a model, the uncertainty associated with detecting chemical agents intentionally added to food products was evaluated. Accuracy and precision are two types of uncertainties generally associated with sampling plans. Sources of variability that affect precision were the primary focus of this investigation. Test procedures used to detect chemical agents generally include sampling, sample preparation, and analytical steps. The uncertainty of each step contributes to the total uncertainty of the test procedure. Using variance as a statistical measure of uncertainty, the variance associated with each step of the test procedure used to detect aflatoxin in shelled corn was determined for both low and high levels of contamination. For example, when using a 1-kg sample, Romer mill, 50-g subsample, and high-performance liquid chromatography to test a lot of shelled corn contaminated with aflatoxin at 10 ng/g, the total variance associated with the test procedure was 149.2 (coefficient of variation of 122.1%). The sampling, sample preparation, and analytical steps accounted for 83.0, 15.6, and 1.4% of the total variance, respectively. A variance of 149.2 suggests that repeated test results will vary from 0 to 33.9 ng/g. Using the same test procedure to detect aflatoxin at 10,000 ng/g, the total variance was 264,719 (coefficient of variation of 5.1%). The sampling, sample preparation, and analytical steps accounted for 41, 57, and 2% of the total variance, respectively. A variance of 264,719 suggests that repeated test results will vary from 8,992 to 11,008 ng/g. Foods contaminated at low levels reflect a situation in which a small percentage of particles is contaminated and sampling becomes the largest source of uncertainty. Large samples are required to overcome the ''needle-in-the-haystack'' problem. Aflatoxin is easier to detect and identify in foods intentionally contaminated at high levels than in foods with low levels of contamination because the relative standard deviation (coefficient of variation) decreases and the percentage of contaminated kernels increases with an increase in concentration.

Chemical agents can occur in foods destined for human consumption, and assays for detection and quantification of these agents are needed. In research, quality assurance, regulatory, and biosecurity activities, correct decisions concerning the fate of bulk food lots can be made only if the chemical agent in the lot can be determined with a high degree of accuracy and precision. The concentration of a chemical agent in a bulk food lot is usually estimated by measuring the concentration in a small portion (a test sample) of the lot. The concentration in the bulk lot is assumed to be the same as that found in the test sample. Based on this measured sample concentration, some decision is made about the acceptability or edible quality of the bulk lot.

A sampling plan is defined by the test procedure used to quantify the chemical agent and an accept-reject limit. A test procedure is a multistage process and generally consists of three steps: sampling, sample preparation, and analysis. The sampling step includes how the sample will be selected or removed from the bulk lot and the size of the sample. For granular products, the sample preparation step is also a two-part process where the sample is ground in a mill to reduce particle size and then a subsample is removed from the comminuted sample. In the analytical step, the chemical agent in the subsample is quantified using approved analytical procedures.

The measured chemical agent concentration in the test sample will be used to either estimate the true concentration in the bulk lot or determine whether the lot is acceptable or unacceptable (below or above a defined legal limit). In a regulatory environment, the measured chemical agent concentration will be compared with a defined accept-reject limit that is usually equal to a legal limit. If biosecurity is an issue, the accept-reject limit may be zero or some nondetectable level. When the measured concentration is greater than the defined accept-reject limit, the lot is rejected and diverted from the food chain. Otherwise, it is accepted and processed into food or distributed to consumers. If the sample concentration does not accurately and precisely reflect the lot concentration, then the lot may be misclassified and there may be undesirable economic and/or health consequences. Sampling plans must be designed to minimize the misclassification of lots and reduce the undesirable consequences associated with regulatory decisions about the fate of bulk food lots. Therefore, it is important to understand the undesirable consequences of uncertainty associ-

* Author for correspondence. Tel: 919-515-6731; Fax: 919-515-7760; E-mail: whitaker@eos.ncsu.edu.

† The use of trade names in this publication does not imply endorsement by the U.S. Department of Agriculture or the North Carolina Agricultural Research Service of the products named nor criticism of similar ones not mentioned.

‡ Paper no. BAE 2004-04 of the journal series of the Department of Biological and Agricultural Engineering, North Carolina State University.

Downloaded from http://meridian.allenpress.com/jfp/article-pdf/68/6/1306/1676730/0362-028x-68_6_1306.pdf by guest on 19 January 2021

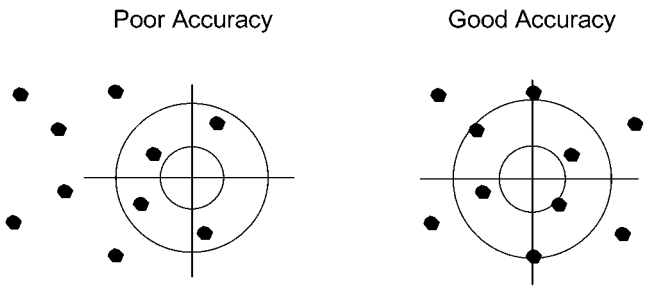


FIGURE 1. *High and low accuracy using a target as an example.*

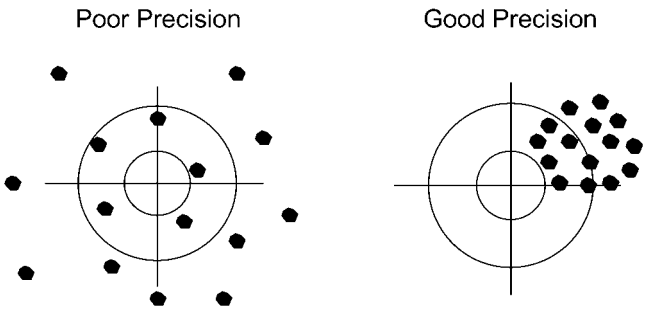


FIGURE 2. *High and low precision using a target as an example.*

ated with sampling plans and the sources of uncertainty and to determine ways to reduce uncertainty when estimating concentrations of chemical agents in food matrices. The uncertainties associated with detecting aflatoxin in shelled corn was used in this study as a model to describe the uncertainties that could be encountered when evaluating chemical agents in food matrices.

## MATERIALS AND METHODS

**Uncertainty.** Even when accepted sampling, sample preparation, and analytical procedures are used *(1, 6, 8, 12, 13, 18, 25),* there are uncertainties associated with each of the above steps of the test procedure *(7, 11, 16, 17, 23, 24, 26, 30, 31).* Because of these uncertainties, the true concentration of a chemical agent in the food lot cannot be determined with 100% certainty by measuring the chemical agent concentration in a test sample taken from the lot. Accuracy and precision are two types of uncertainties generally associated with sampling plans used to measure chemical agents in bulk lots. Accuracy is the closeness of the measured values to the true value *(2),* and precision is the closeness of the measured values to each other *(2).* Graphical examples of high and low accuracy and precision are shown in Figures 1 and 2 using a target as an example. Accuracy is usually associated with a bias in the test procedure, whereas precision is usually associated with variability in the test procedure. Biases are usually minimized by using equipment and procedures that give every member of the population an equal chance of being chosen, e.g., sampling equipment should not exclude certain members of the lot. Analytical methods used to quantify chemical agents should have 100% extraction efficiency and accurate reference standards. In this article, precision is the primary focus, including sources of variability and methods for reducing variability.

Even when biases of the test procedure are negligible, there can be random variation among test results. Ten replicated aflatoxin test results from each of six contaminated shelled corn lots are shown in Table 1 *(11).* For each test result in the table, the test procedure consisted of (i) comminuting a 1.1-kg sample of corn kernels in a Romer subsampling mill, (ii) removing a 50-g subsample from the comminuted test sample, (iii) extracting aflatoxins from a 50-g subsample using solvent as described by the Association of Official Analytical Chemists *(8),* and (iv) quantifying the aflatoxins using high-performance liquid chromatography (HPLC). The 10 aflatoxin test results from each lot are ranked from low to high to demonstrate several important characteristics about replicated aflatoxin test results taken from the same contaminated lot. Several conclusions can be drawn from examination of these data.

First, the wide range among replicated test results from the same lot reflects the high degree of variability associated with estimating the true chemical agent concentration in a bulk lot. In Table 1, the variability is described by the variance and the coefficient of variation (CV). The maximum test result can be several times larger than the mean of the 10 test results or the best estimate of the lot concentration.

Second, the amount of variation among the 10 test results appears to be a function of the lot concentration. As the lot concentration increases, the variance among test results increases, but the CV (square root of the variance divided by the lot mean) decreases.

Third, the distribution of the 10 test results for each lot in Table 1 is not always symmetrical about the lot concentration. The distributions are positively skewed, meaning that more than half of the sample test results are below the true lot concentration. However, the distribution of sample test results becomes more symmetrical as the lot concentration increases. This skewness can be observed by counting the number of aflatoxin test results above and below the lot concentration in Table 1 (average of the 10 sample test results). When a single sample is tested from a contaminated lot, there is more than a 50% chance that the sample test result will be lower than the true lot concentration. Although

TABLE 1. *Distribution of aflatoxin test results for 10 samples from each of six lots of shelled corn*

| Lot no. | Sample test results (ng/g)[a] | | | | | | | | | | Mean (ng/g) | Variance $(ng/g)^2$ | Coefficient of variation (%) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 0 | 0 | 0 | 0 | 0 | 3 | 10 | 26 | 27 | 33 | 6.8 | 179.1 | 196.3 |
| 2 | 0 | 0 | 0 | 5 | 8 | 10 | 15 | 38 | 40 | 49 | 11.8 | 347.2 | 158.1 |
| 3 | 0 | 0 | 11 | 14 | 21 | 28 | 39 | 55 | 62 | 90 | 26.1 | 866.8 | 113.0 |
| 4 | 0 | 9 | 22 | 26 | 31 | 33 | 59 | 65 | 69 | 92 | 35.1 | 856.5 | 83.4 |
| 5 | 48 | 55 | 73 | 89 | 98 | 106 | 117 | 133 | 147 | 170 | 95.5 | 1,558.2 | 41.3 |
| 6 | 165 | 211 | 223 | 256 | 268 | 295 | 326 | 378 | 379 | 444 | 276.8 | 7,629.4 | 31.6 |

[a] Each 1.1-kg sample was ground in a Romer mill, and aflatoxin was extracted from a 50-g subsample, of which one aliquot was assayed using HPLC.

Downloaded from http://meridian.allenpress.com/jfp/article-pdf/68/6/1306/1676730/0362-028x-68_6_1306.pdf by guest on 19 January 2021

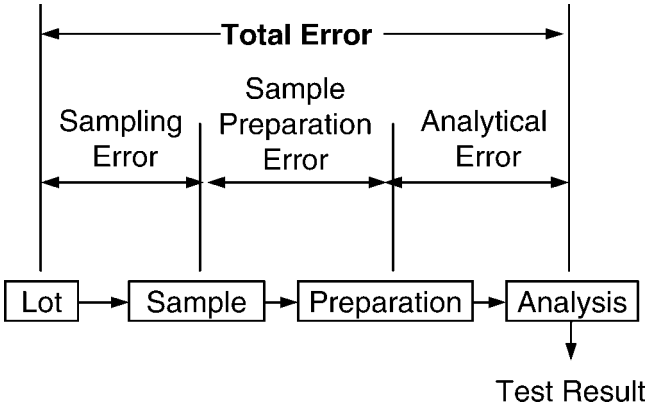


FIGURE 3. *Sources of variability associated with typical test procedures.*

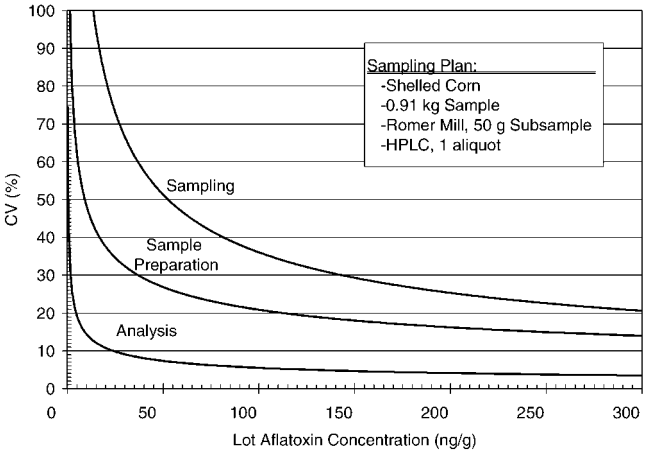


FIGURE 4. *Coefficients of variation (CV) for sampling, sample preparation, and analysis when testing shelled corn for aflatoxin.*

it is not apparent in Table 1, the skewness is greater for small sample sizes and the distribution becomes more symmetrical as sample size increases. These characteristics shown in Table 1 for aflatoxin estimation in corn are also generally found for other chemical agents and other commodities *(7, 11, 27, 30).*

**Sources of variability.** The variability among test results in Table 1 reflects the total variability associated with the test procedure. Each step of the test procedure contributes to the total variability observed in Table 1. As shown in Figure 3, the total variability (VT) (using variance as the statistical measure of variability) associated with a test procedure is equal to the sum of the variances associated with each step of the test procedure *(15).* The additive relationship of the variances for sampling (VS), sample preparation (VSS), and analysis (VA) is described by the following equation:

$$VT = VS + VSS + VA \quad (1)$$

**Sampling variability.** Results of studies on a wide variety of agricultural products (e.g., peanuts, cottonseed, shelled corn, and pistachio nuts) indicate that for small sample sizes the sampling step is usually the largest source of variability associated with the chemical agent test procedure *(7, 11, 16, 17, 24, 26, 30, 31).* Even when using accepted sample selection equipment and random sample selection procedures, sampling variability is large because of the distribution of contaminated particles within a lot. Results of aflatoxin studies on corn and peanuts indicate that about 0.1% of the kernels in a lot are contaminated, and the concentration on a single kernel may be extremely high *(9, 32).* Results of other studies on various products (e.g., peanuts, cottonseed, shelled corn, wheat, barley, and pistachio nuts) support the finding that a very small percentage of the particles in the lot are contaminated and the concentration on a single particle may be extremely high. Cucullu et al. *(3, 4)* reported aflatoxin concentrations greater than 1,000,000 ng/g (ppb) for individual peanut kernels and 5,000,000 ng/g for cottonseed. Shotwell et al. *(19)* reported finding more than 400,000 ng/g of aflatoxin in a corn kernel.

Because of this extreme range in aflatoxin concentrations among a few contaminated kernels in a lot, variation among replicated sample test results tends to be large. The sampling variance associated with testing shelled corn was estimated empirically *(11)* for any sample size, ns, as

$$VS = (12.95/ns)M^{0.98} \quad (2)$$

where $M$ is the aflatoxin concentration in the lot in nanograms of total aflatoxin per gram of corn (or ppb), and ns is the mass of shelled corn in the sample in kilograms (kernel count per gram was 3.0). From equation 2, the sampling variance is a function of the lot concentration $M$ and sample size ns. The sampling variance among 1.0-kg samples taken from a lot of shelled corn at 10 ppb is 123.8, and the CV is 111.3%. The sampling CV over a range of aflatoxin concentrations is shown in Figure 4.

Researchers have developed equations to describe the sampling variance for several commodities and chemical agents *(7, 11, 16, 17, 23, 24, 26, 30, 31).* The equations are specific for the type chemical agent and the type product studied but generally show that sampling variance is a function of concentration and increases with concentration.

**Sample preparation variability.** Once the test sample has been taken from the lot, the sample must be prepared so the chemical agent can be quantified. Often, it is not practical to quantify the chemical agent from a large test sample, and the chemical agent is usually extracted from a much smaller portion of product (subsample) taken from the test sample. If the commodity is a granular product such as shelled corn, it is essential that the entire test sample be comminuted in a suitable mill before a subsample is removed *(5, 25).* Removing a subsample of whole seed from the test sample before the comminution process is simply a sample size reduction process and eliminates the benefits associated with the larger sample of granular product. After the sample has been comminuted in a mill to reduce particle size, a subsample is removed. The assumption is that the distribution of contaminated particles in the comminuted sample is similar to the distribution among contaminated kernels found in the lot. As a result, there is also variability among replicated subsamples taken from the same comminuted test sample. However, the sample preparation variance is not as large as the sampling variance because of the large number of comminuted particles in the subsample. An example of sample preparation variance for aflatoxin and shelled corn is shown in equation 3 for any subsample size nss *(11)*:

$$VSS = (62.70/nss)M^{1.27} \quad (3)$$

where $M$ is the aflatoxin concentration in the test sample in ppb, and nss is the mass of shelled corn in the subsample in grams. The variance in equation 3 also reflects the use of a Romer mill that produces particles small enough so that most will pass through a no. 20 screen. From equation 3, the sample preparation variance is also a function of the aflatoxin concentration in the sample and the subsample size. The sample preparation variance associated with a 50-g subsample taken from a sample at 10 ppb

Downloaded from http://meridian.allenpress.com/jfp/article-pdf/68/6/1306/1676730/0362-028x-68_6_1306.pdf by guest on 19 January 2021

is 23.3, and the CV is 48.3%. The sample preparation CV over a range of aflatoxin concentrations is shown in Figure 4.

Researchers have developed equations to describe the sample preparation variance for several commodities, types of mills, and mycotoxins *(7, 11, 16, 17, 23, 24, 26, 30, 31).* The equations are specific for the type chemical agent, type of mill (particle size), and type of product used in the study. The type of mill affects the particle size distribution. When the average particle size decreases (number of particles per unit mass increases), then the subsampling variances for a given size subsample decreases.

**Analytical variability.** Once the subsample is removed from the comminuted test sample, the chemical agent in the subsample is quantified. Analytical methods usually involve several steps such as solvent extraction, filtration, centrifugation, drying, dilution, and quantification *(14).* As a result, there can be considerable variation among replicated analyses on the same subsample extract. The analytical variance associated with HPLC techniques (VAh) used to measure aflatoxin in corn is given by equation 4 *(11)* for any number of aliquots, na:

$$\mathrm{VAh} = (0.143/\mathrm{na})M^{1.16} \quad (4)$$

where $M$ is the aflatoxin concentration in the subsample in ppb, and na is the number of aliquots quantified by HPLC. The analytical variance and CV associated with using HPLC used to measure aflatoxin in a comminuted subsample of corn at 10 ng/g are 2.1 and 14.3%, respectively. The analytical CV over a range of aflatoxin concentrations is shown in Figure 4.

HPLC tends to have less variability than do other analytical techniques such as thin-layer chromatography (TLC) and enzyme-linked immunosorbent assay (ELISA) *(28).* Using precision estimates from collaborative studies, the analytical variances associated with TLC (VAt) and ELISA (VAe) methods to measure aflatoxin in corn are shown in equations 5 and 6, respectively:

$$\mathrm{VAt} = (0.316/\mathrm{na})M^{1.744} \quad (5)$$
$$\mathrm{VAe} = (0.631/\mathrm{na})M^{1.293} \quad (6)$$

The coefficients of variation associated with measuring aflatoxin in a corn subsample at 10 ppb with the TLC and ELISA methods are 41.9 and 35.2%, respectively. The variability associated with HPLC (10.7%, equation 4) is lower than those of TLC or ELISA.

All of this analytical variance information reflects results from single laboratories and does not reflect variances among laboratories. Some laboratories may have higher or lower variances than those described in equations 4 through 6. The variance among laboratories is about double the within-laboratory variance *(28).*

**Total variability.** As shown in Figure 3 and equation 1, the total variability is the sum of each variance component (represented by equations 2 through 4):

$$\mathrm{VT} = (12.95/\mathrm{ns})M^{0.98} + (62.70/\mathrm{nss})M^{1.27} + (0.143/\mathrm{na})M^{1.16} \quad (7)$$

Examples of the magnitude of the total variability associated with each step of a chemical agent test procedure (equation 7) are given below. The expected total variance associated with testing a shelled corn lot at 10 ppb when using a 1.0-kg sample, grinding the test sample in a Romer mill, taking a 50-g subsample from a comminuted sample, and quantifying aflatoxin in one aliquot by HPLC methods can be estimated from equation 7, where $M = 10$, ns = 1, nss = 50, and na = 1 (Table 2):

$$\mathrm{VT} = 123.8 + 23.3 + 2.1 = 149.2 \quad (8)$$

TABLE 2. *Variability associated with each step of the test procedure used to detect aflatoxin in shelled corn at 10 ng/g*[a]

| Test procedure | Sample size | Variance $(ng/g)^2$ | Coefficient of variation (%) | Ratio (%)[b] |
|---|---|---|---|---|
| Sample | 1 kg | 123.8 | 111.3 | 83.0 |
| Preparation | 50 g | 23.3 | 48.3 | 15.6 |
| Analysis (HPLC) | 1 aliquot | 2.1 | 14.3 | 1.4 |
| Total | | 149.2 | 122.1 | 100.0 |

[a] The contamination rate is 3 contaminated kernels per 10,000 kernels.

[b] The variance associated with each step divided by the total variance.

The variance, standard deviation (SD), and CV associated with the total test procedure when measuring a lot at 10 ppb are 149.2, 12.2, and 122.1%, respectively. The sampling, subsampling, and analytical variances account for 83.0, 15.6, and 1.4% of the total testing variance, respectively.

## RESULTS

**Reducing variability of test procedure.** The only way to achieve a more precise estimate of the true lot concentration is to reduce the total variability of the test procedure. Minimizing the variability associated with each step of the chemical agent test procedure reduces the total variability. Increasing the size of the sample can reduce sampling variability. Increasing the size of the subsample and/or increasing the degree of comminution (increasing the number of particles per unit mass in the subsample) reduces the sample preparation variability. Increasing the number of aliquots quantified by the analytical method and/or using a more precise quantification method (i.e., using HPLC instead of TLC) will reduce the analytical variability. If the variability associated with one or more of these steps can be reduced, then the total variability among test results from the same lot can be reduced.

In the above example, the sampling step accounted for 83.0% of the total variability. The best use of resources in this example may be to increase sample size to reduce the total variability. Increasing the sample size from 1.0 to 10.0 kg will reduce the sampling variance (equation 2) to 12.4. The total variance with the 10.0-kg sample, 50-g subsample, and quantification of one aliquot by HPLC now becomes

$$\mathrm{VT} = 12.4 + 23.3 + 2.1 = 37.8 \quad (9)$$

The variance, SD, and CV associated with the total testing procedure have been reduced to 37.8, 6.1, and 61.5%, respectively. Sampling now accounts for 32.8% of the total variability.

The range among test results associated with any size sample and subsample and the number of analyses about the lot concentration $M$ can be estimated from the SD (square root of the total variance) associated with the test procedure. Approximately 95% of all test results will fall between a low of $M - 1.96\mathrm{SD}$) and a high of $M + 1.96\mathrm{SD}$. The two expressions are only approximate because they are

Downloaded from http://meridian.allenpress.com/jfp/article-pdf/68/6/1306/1676730/0362-028x-68_6_1306.pdf by guest on 19 January 2021

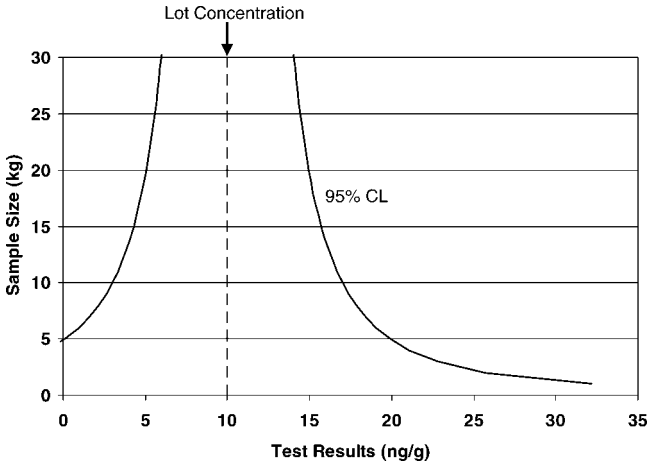


FIGURE 5. *Effect of increasing sample size on improving precision or reducing the range of aflatoxin test results when using a 1-kg sample, 50-g subsample, Romer mill, and 1 aliquot and HPLC to estimate aflatoxin in a lot at 10 ng/g.*

valid only for a normal distribution where test results are symmetrical about the mean. The distribution among aflatoxin test results is usually skewed but will approach a symmetrical distribution as sample size becomes larger. The effect of increasing sample size on the range of test results when testing a contaminated lot of shelled corn that has 10 ppb aflatoxin is shown in Figure 5. The range among the smallest and largest sample test results does not decrease at a constant rate as sample size increases. For example, at small sample sizes doubling the sample size has a greater effect on decreasing the range than it does at large sample sizes. This characteristic suggests that increasing sample size beyond a certain point may not be the best use of resources and that increasing subsample size or number of analyses may be a better use of resources for reducing the range of test results (improving precision) once sample size has become significantly large.

**Low contamination rates and sampling precision.** The aflatoxin model predicts that a contaminated lot at 10 ppb has about 3 contaminated kernels for every 10,000 kernels and the concentration on the contaminated kernels can vary from low to very high *(9)*. The contamination rate is 3/10,000 or 0.0003. Even using proper sample selection techniques, the variation among test sample concentrations is large because of small sample size and the sparse distribution of contaminated items in the lot. When sampling a lot with this distribution of contamination with small samples (1.0 kg), equation 8 (and Table 2) indicates that the sampling step accounts for most of the variability (uncertainty) associated with the total variability of the test procedure. Because of the sparse distribution among items in a contaminated lot, it is easy to miss the contaminated items in a lot with a small sample and thus to underestimate the true lot concentration or, worse, not detect the contamination at all. In the situation where the contamination rate is very low, it is important to use large sample sizes to minimize the chances of not getting a contaminated item in the sample. Increasing sample size helps ensure that a contaminated item will be detected, and a more precise estimate of the true level of contamination will be obtained (Figure 5).

TABLE 3. *Variability associated with each step of the test procedure used to detect aflatoxin in shelled corn at 10,000 ng/g*[a]

| Test procedure | Sample size | Variance $(ng/g)^2$ | Coefficient of variation (%) | Ratio (%)[b] |
|---|---|---|---|---|
| Sample | 1 kg | 107,713 | 3.3 | 40.7 |
| Preparation | 50 g | 150,764 | 3.9 | 57.0 |
| Analysis (HPLC) | 1 aliquot | 6,242 | 0.8 | 2.3 |
| Total | | 264,719 | 5.1 | 100.0 |

[a] The contamination rate is 2,300 contaminated kernels per 10,000 kernels.

[b] The variance associated with each step divided by the total variance.

**High contamination rates and sampling precision.** The aflatoxin model predicts that a contaminated lot at 10,000 ppb has about 2,300 contaminated kernels for every 10,000 kernels, and the concentration on the contaminated kernels can vary from low to very high *(9)*. The contamination rate is 23/100 or 0.23. The variance associated with a 1-kg sample, 50-g subsample, and HPLC analysis is shown in Table 3 when testing a lot at 10,000 ppb. Although the variances for each step of the test procedure are extremely large, the CVs for each step of the test procedure are small. A small CV indicates that the SD is a small percentage of the lot concentration (10,000 ppb). A 1-kg sample contributes about 41% of the total variability. The 41% is much less than the 83% that a 1-kg sample contributes when sampling a lot at 10 ppb. Increasing sample size reduces the variance associated with a 1-kg sample (Table 2). Figure 6 shows the effect of increasing sample size on improving precision. As sample size increases, the range among test results about the true lot concentration of 10,000 ppb decreases. In heavily contaminated lots such as

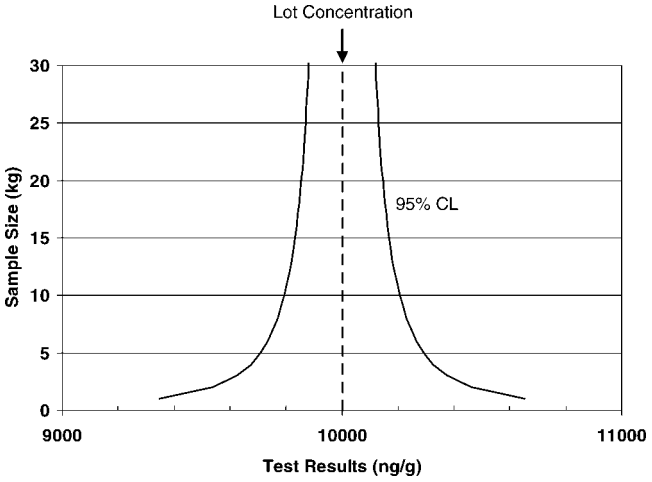


FIGURE 6. *Effect of increasing sample size on improving precision or reducing the range of aflatoxin test results when using a 1-kg sample, 50-g subsample, Romer mill, and 1 aliquot and HPLC to estimate aflatoxin in a lot at 10,000 ng/g.*

Downloaded from http://meridian.allenpress.com/jfp/article-pdf/68/6/1306/1676730/0362-028x-68_6_1306.pdf by guest on 19 January 2021

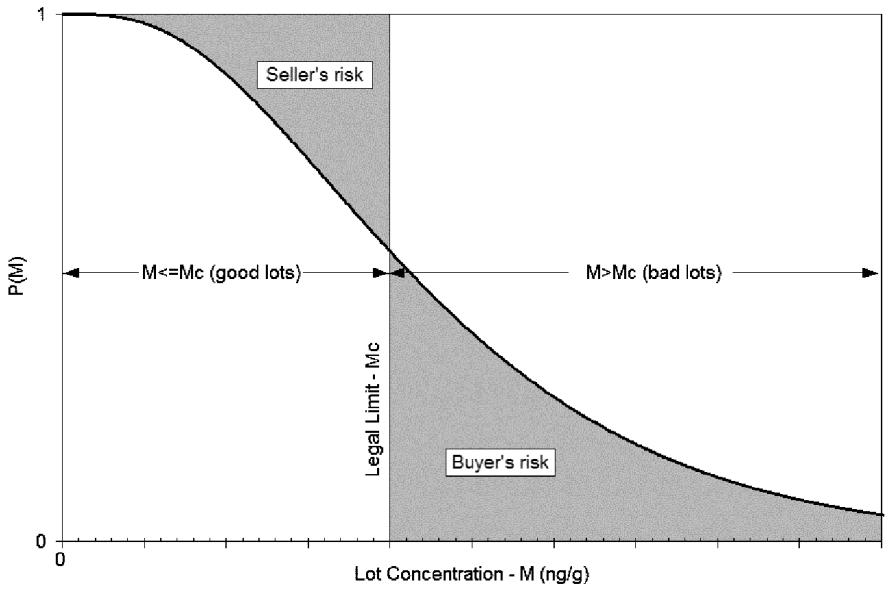


FIGURE 7. *Typical shape of an operating characteristic curve used to predict portion of lots accepted, P(M) and to evaluate the buyer's and seller's risks associated with a chemical agent sampling plan.*

the one in this example at 10,000 ppb, contamination will be detected with a high degree of certainty even with a small 1-kg sample. Because of the high percentage of contaminated items in the lot, it is highly likely that a contaminated item will be found even in a small sample taken from the lot.

**Misclassification risks.** Because of the large variability among test results (Table 1), two types of mistakes are associated with any sampling program. First, some good lots (lots with a concentration lower than or equal to the legal limit) will test bad and be rejected by the sampling program. This type of mistake is often called the seller's risk or a false-positive result because these lots will be rejected at an unnecessary cost to the seller of the product. Second, some bad lots (lots with a concentration higher than the legal limit) will test good and be accepted by the sampling program. This type of mistake is called the buyer's risk or a false-negative result because contaminated lots will be processed into food causing possible health problems and/or economic loss to the buyer of the product. To maintain an effective regulatory, quality control, and/or biosecurity program, the above risks associated with a sample design must be evaluated and reduced when considered too large. Based upon these evaluations, the costs and benefits (i.e., removal of contaminated lots) associated with a sampling program can be evaluated. For biosecurity issues, quantification of a chemical agent in a food lot may not be as critical as detecting the presence or absence of the agent in the lot. In the latter case, the only risk of concern is a false-negative result, i.e., not detecting a positive lot.

A lot is termed bad when the sample test result is above an accept-reject limit, and the lot is termed good when the sample test result is lower than or equal to an accept-reject limit. Although an accept-reject limit is usually equal to the legal limit, the accept-reject limit can be higher or lower than the legal limit. For a given sample design, lots with a chemical agent concentration $M$ will be accepted with a certain probability $P(M)$ by the sampling plan. A plot of the acceptance probability $P(M)$ versus lot concentration $M$ is called an operating characteristic (OC) curve. Figure 7 depicts the general shape of an OC curve. As $M$ approaches zero, $P(M)$ approaches 1, and as $M$ increases, $P(M)$ approaches zero. The shape of the OC curve is uniquely defined for a particular sampling plan design with designated values of sample size, degree of comminution, subsample size, type analytical method, number of analyses, and the accept-reject limit.

For a given sampling plan, the OC curve indicates the magnitude of the buyer's and seller's risks. When Mc is defined as the legal limit or the maximum lot concentration acceptable, lots with $M >$ Mc are bad and lots with $M \leq$ Mc are good. In Figure 7, the area under the OC curve for $M >$ Mc represents the buyer's risk (bad lots accepted), and the area above the OC curve for $M \leq$ Mc represents the seller's risk (good lots rejected) for a particular sampling plan. The lots rejected, $R(M)$, is $1 - P(M)$.

Because the shape of the OC curve is uniquely defined by the sample size, degree of comminution, subsample size, the number of analyses, and the accept-reject limit, these parameters can be used to reduce the buyer's and seller's risks associated with a sampling plan. Methods have been developed to predict the seller's and buyer's risks, the total number of lots accepted and rejected, the amount of chemical agent in the accepted and rejected lots, and the costs associated with a chemical agent inspection program for several commodities *(6, 10, 20, 21, 29)*. The basic information needed to evaluate the performance of a chemical agent sampling plan includes the variability associated with the chemical agent test procedure and the distribution among sample test results.

The effect of increasing sample size on the shape of the OC curve when testing shelled corn lots for aflatoxin is shown in Figure 8, where the accept-reject limit is 1.0 ppb. This accept-reject limit was chosen instead of 0.0 because the limits of detection of analytical methods are not truly zero but are some small finite positive value. As sample size increases from 1.0 to 3.0 to 10.0 kg, the percentage of false-negative results decreases, as indicated by the in-

Downloaded from http://meridian.allenpress.com/jfp/article-pdf/68/6/1306/1676730/0362-028x-68_6_1306.pdf by guest on 19 January 2021

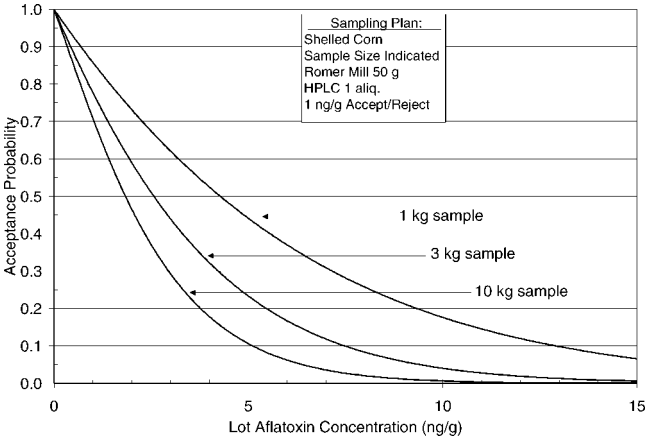


FIGURE 8. *Effect of sample size on the buyer's and seller's risks associated with sampling shelled corn for aflatoxin.*

crease in slope of the OC curve and a decrease in the area below the OC curve as sample size increases. The same effect can be obtained to a lesser extent by increasing the degree of sample comminution, subsample size, or number of analyses.

## DISCUSSION

Because of the variability associated with a chemical agent test procedure, it is difficult to determine with 100% certainty the true concentration of a bulk lot. Even when using acceptable sample selection procedures, there will be variability associated with the chemical agent test procedure. Each step of the test procedure contributes to the total testing variability. The total variance associated with a chemical agent test procedure is the sum of sampling, sample preparation, and analytical variances. For small sample sizes, sampling is usually the largest source of variability because of the sparse distribution of contaminated seed in the lot. Increasing sample size, the degree of sample comminution, subsample size, and the number of aliquots quantified can reduce the variability associated with a chemical agent test procedure. Reducing variability of the chemical agent test procedure will reduce the number of lots misclassified by the sampling plan.

The percentage of contaminated items in a lot also affects the number of lots misclassified by a sampling plan. When a high percentage of items in a lot are contaminated, it is highly likely that a contaminated item will be found even in a small sample taken from the lot. Conversely, when a small percentage of items in a lot are contaminated, it is easy to miss the contaminated items when a small sample is taken and thus to underestimate the true lot concentration or, worse, not detect the contamination at all. In the situation where the contamination rate is very low, it is important to use large sample sizes to minimize the chances of not getting a contaminated item in the sample. Increasing sample size helps ensure that a contaminated lot will be detected, and a more precise estimate of the true level of contamination will be obtained. Methods have been developed for several chemical agents and several commodities to evaluate and design sampling plans that minimize the misclassification of food shipments based on the presence or absence of chemical agents.

## REFERENCES

1. Campbell, A. D., T. B. Whitaker, A. E. Pohland, J. W. Dickens, and D. L. Park. 1986. Sampling, sample preparation, and sampling plans for foodstuffs for chemical agent analysis. *Pure Appl. Chem.* 58: 305–314.
2. Cochran, W. G., and G. M. Cox. 1957. Experimental design. John Wiley and Sons, London.
3. Cucullu, A. F., L. S. Lee, R. Y. Mayne, and L. A. Goldblatt. 1986. Determination of aflatoxin in individual peanuts and peanut sections. *J. Am. Oil Chem. Soc.* 43:89.
4. Cucullu, A. F., L. S. Lee, and W. A. Pons. 1977. Relationship of physical appearance of individual mold damaged cottonseed to aflatoxin content. *J. Am. Oil Chem. Soc.* 54:235A.
5. Dickens, J. W., and J. B. Satterwhite. 1969. Subsampling mill for peanut kernels. *Food Technol.* 23:90–92.
6. Food and Agriculture Organization. 1993. Sampling plans for aflatoxin analysis in peanuts and corn. FAO Food and Nutrition Paper no. 55. Food and Agriculture Organization, Rome.
7. Hart, L. P., and O. Schabenberger. 1998. Variability of vomitoxin in truckloads of wheat in a wheat scab epidemic year. *Plant Dis.* 82: 625–630.
8. Helrich, K. (ed.). 1990. Official methods of analysis of the Association of Official Analytical Chemists, 15th ed. Association of Official Analytical Chemists, Arlington, Va.
9. Johansson, A. S., T. B. Whitaker, W. M. Hagler, Jr., F. G. Giesbrecht, and J. H. Young. 2000. Testing shelled corn for aflatoxin, part II: modeling the distribution of aflatoxin test results. *J. AOAC Int.* 83: 1270–1278.
10. Johansson, A. S., T. B. Whitaker, W. M. Hagler, Jr., F. G. Giesbrecht, and J. H. Young. 2000. Testing shelled corn for aflatoxin, part III: evaluating the performance of aflatoxin sampling plans. *J. AOAC Int.* 83:1279–1284.
11. Johansson, A. S., T. B. Whitaker, W. M. Hagler, Jr., F. G. Giesbrecht, J. H. Young, and D. T. Bowman. 2000. Testing shelled corn for aflatoxin, part I: estimation of variance components. *J. AOAC Int.* 83:1264–1269.
12. Nesheim, S. 1979. Methods of aflatoxin analysis. National Bureau of Standards Special Publication no. 519. National Bureau of Standards, Washington, D.C.
13. Park, D. L., and A. E. Pohland. 1989. Sampling and sample preparation for detection of natural toxicants in food and feed. *J. AOAC Int.* 72:399–404.
14. Pohland, A. E., and M. W. Trucksess. 2000. Mycotoxin method evaluation, p. 3–10. *In* M. W. Trucksess and A. E. Pohland (ed.), Mycotoxin protocols, vol. 157. Humana Press, Totowa, N.J.
15. Remington, R. D., and M. A. Schrok. 1970. Statistics and applications to the biological and health sciences. Prentice-Hall, Englewood Cliffs, N.J.
16. Schatzki, T. F. 1995. Distribution of aflatoxin in pistachios. 1. Lot distributions. *J. Agric. Food. Chem.* 43:1561–1565.
17. Schatzki, T. F. 1995. Distribution of aflatoxin in pistachios. 2. Distribution in freshly harvested pistachios. *J. Agric. Food. Chem.* 43: 1566–1569.
18. Schuller, P. O., W. Horwitz, and L. Stoloff. 1976. A review of sampling and collaboratively studied methods of analysis for aflatoxin. *J. AOAC Int.* 59:1315.
19. Shotwell, O. L., M. L. Goulden, and C. C. Hessletine. 1974. Aflatoxin: distribution in contaminated corn. *Cereal Chem.* 51:492.
20. Whitaker, T. B., and J. W. Dickens. 1979. Evaluation of the Peanut Administrative Committee testing program for aflatoxin in shelled peanuts. *Peanut Sci.* 6:7–9.
21. Whitaker, T. B., and J. W. Dickens. 1989. Simulating aflatoxin testing plans for shelled peanuts in the export market. *J. AOAC Int.* 72:644–648.
22. Whitaker, T. B., J. W. Dickens, and R. J. Monroe. 1972. Comparison of the observed distribution of aflatoxin in shelled peanuts to the negative binomial distribution. *J. Am. Oil Chem. Soc.* 49:590–593.

Downloaded from http://meridian.allenpress.com/jfp/article-pdf/68/6/1306/1676730/0362-028x-68_6_1306.pdf by guest on 19 January 2021

23. Whitaker, T. B., J. W. Dickens, and R. J. Monroe. 1974. Variability of aflatoxin test results, *J. Am. Oil Chem. Soc.* 51:214–218.
24. Whitaker, T. B., J. W. Dickens, and R. J. Monroe. 1979. Variability associated with testing corn for aflatoxin. *J. Am. Oil Chem. Soc.* 56: 789–794.
25. Whitaker, T. B., and F. E. Dowell. 1995. Sampling methods to measure aflatoxin and grade factors of peanuts, p. 475–499. *In* H. E. Pattee and H. T. Stalker (ed.), Advances in peanut science. American Peanut Research and Education Society, Stillwater, Okla.
26. Whitaker, T. B., F. E. Dowell, W. M. Hagler, Jr., F. G. Giesbrecht, and J. Wu. 1993. Variability associated with sampling, sample preparation, and chemically testing farmers' stock peanuts for aflatoxin. *J. AOAC Int.* 77:107–116.
27. Whitaker, T. B., W. M. Hagler, Jr., F. G. Giesbrecht, and A. S. Johansson. 2000. Sampling, sample preparation, and analytical variability associated with testing wheat for deoxynivalenol. *J. AOAC Int.* 83:1285–1292.
28. Whitaker, T. B., W. Horwitz, R. Albert, and S. Nesheim. 1996. Variability associated with methods used to measure aflatoxin in agricultural commodities. *J. AOAC Int.* 79:476–485.
29. Whitaker, T. B., J. Springer, P. R. Defize, W. J. deKoe, and R. Coker. 1995. Evaluation of sampling plans used in the United States, UK, and The Netherlands to test raw shelled peanuts for aflatoxin. *J. AOAC Int.* 78:1010–1018.
30. Whitaker, T. B., M. W. Trucksess, A. S. Johansson, F. G. Giesbrecht, W. M. Hagler, Jr., and D. T. Bowman. 1998. Variability associated with testing shelled corn for fumonisin. *J. AOAC Int.* 81:1162–1168.
31. Whitaker, T. B., M. E. Whitten, and R. J. Monroe. 1976. Variability associated with testing cottonseed for aflatoxin. *J. Am. Oil Chem. Soc.* 53:502–505.
32. Whitaker, T. B., and E. H. Wiser. 1969. Theoretical investigations into the accuracy of sampling shelled peanuts for aflatoxin. *J. Am. Oil Chem. Soc.* 46:377–379.

Downloaded from http://meridian.allenpress.com/jfp/article-pdf/68/6/1306/1676730/0362-028x-68_6_1306.pdf by guest on 19 January 2021