# Sample Preparation and Analytical Considerations for the US Aflatoxin Sampling Program for Shelled Peanuts

J. P. Davis*, M. D. Jackson, J.M. Leek, and M. Samadpour[1]

## ABSTRACT

The USDA aflatoxin sampling program for shelled peanuts is an important component of broader industry efforts to minimize aflatoxin occurrence in the edible market. In this program, official samples are milled with either a traditional hammer/automatic sub-sampling mill, commonly called the Dickens Mill (DM) or with a vertical cutter mill (VCM). Particle size reduction and sample homogenization are the primary objectives of sample preparation (milling) to generate subsamples which best represent the parent sample composition for downstream analysis. DM particle size reduction is limited by the 3.2 mm round hole screens internal to the mill which prevent pasting of the sample. VCM grinding converts the sample to a paste while simultaneously homogenizing the sample. Experiments demonstrate that when testing aflatoxin contaminated peanuts for equivalent sized subsamples prepared from the two mill types, made into water slurries per USDA specifications and subsequently extracted and tested for total aflatoxin per USDA specifications, VCM subsamples are more normally distributed around the sample aflatoxin mean, whereas DM subsamples are more positively skewed (median lower than mean) around the sample aflatoxin mean. Accordingly, milling official samples with a DM compared to VCM promotes more lot misclassifications. It is also demonstrated that for a given subsample after extraction and immunoaffinity column (IAC) purification, the total aflatoxin measured by either high performance liquid chromatography (HPLC) or fluorometry (both USDA approved) are practically equivalent from an accuracy perspective. There are costs (time and resources) associated with decreasing natural variation due to sampling, sample preparation and analytical testing in an aflatoxin sampling/ testing program. Sample preparation is a greater source of variation compared to that of the analytical testing. Resources would be better spent replacing DM with VCM mills than converting the final analytical step from IAC-fluorometry to IAC-HPLC in an effort to best classify peanut lots for the edible market.

---

---

Aflatoxins are toxic metabolites produced by a variety of *Aspergillus* species, notably *Aspergillus flavus* and *Aspergillus parasiticus* (Cole *et al*. 1995; Payne and Brown, 1998) These fungi are ubiquitous in tropical and temperate agricultural environments and under specific pre-harvest and/or post-harvest environmental conditions, these fungi can infect numerous crops, including peanut, maize, almonds and pistachios, among others, with aflatoxin. There are four primary aflatoxin types, designated $B_1$, $B_2$, $G_1$ and $G_2$, that can be commonly isolated from aflatoxin contaminated crops. Aflatoxin $B_1$ typically predominates and is also most toxic (Kensler et al. 2011). The World Health Organization, International Agency for Research on Cancer (IARC) classifies aflatoxins as established human liver carcinogens (2012). Aflatoxin exposure in developing countries is associated with increased liver cancer rates (further exacerbated by hepatitis viral infection) and increased incidences of childhood growth stunting (Kensler et al. 2011).

Aflatoxin was originally discovered in 1960 as the causative agent of "Turkey X" disease in which large numbers of turkey poults died in England after consuming contaminated peanut meal from Brazil (Lancaster *et al*. 1961; Cole *et al*. 1995). The toxicity was isolated to compounds derived from *Aspergillus flavus* mold which had been associated with the peanut meal (Sargeant *et al*. 1961). Since this initial discovery, the US Peanut industry has made extensive investments to develop and implement scientifically sound strategies for minimizing aflatoxin occurrence in edible peanuts. Examples include improved farming, harvesting, storage, and processing systems (Dorner, 2008). By the mid 1960's, in concert with newly established FDA regulations on finished product aflatoxin limits, the USDA began regulating that all lots of shelled peanuts in the United States, being considered for

[1]The authors are Jack P. Davis, Ph.D., Director of Technical Services, JLA International, Albany, GA, 31707 and Adjunct Faculty, Dept. of Food, Bioprocessing & Nutrition Sciences, North Carolina State University, Raleigh, NC 27695; James M. Leek, Chairman of the Board, JLA International; Mike Jackson, President, JLA International; and Mansour Samadpour, Ph.D., President and CEO of IEH Laboratories, Lake Forest Park, WA 98155.

*Corresponding Author Email: jackdavis@jlaglobal.com

the edible domestic market, be sampled and tested for aflatoxin. This program has evolved over the years and remains an important component of broader, industry efforts to minimize aflatoxin occurrence in edible peanuts.

When present, aflatoxin contamination in a shelled lot typically affects only a small frequency of kernels. Early work in select lots demonstrated an approximate 0.1% to 2.5% frequency contamination (Cucullu et al. 1966). Contaminated kernels commonly had visual damage, but damage was not a prerequisite for contamination (Cucullu et al. 1966). Furthermore, contaminated kernels demonstrated a range of aflatoxin concentrations, i.e. from just detectable up to approximately 1000 mg/kg (Cucullu et al. 1966). This highly skewed, non-uniform kernel to kernel contamination, makes determining the true (average) aflatoxin concentration of a contaminated, shelled lot of peanuts with 100% certainty impossible (Whitaker, 2006). Instead, the goal of a scientifically sound sampling plan is to estimate aflatoxin concentration of a lot with a reasonable level of statistical confidence to minimize bad lots (aflatoxin mean above the established threshold) accepted and good lots (aflatoxin mean below the established threshold) rejected per market expectations. With this goal in mind, there are three components of any aflatoxin sampling and testing program for shelled peanuts that must be recognized, carefully designed, and carefully controlled: sampling, sample preparation and analysis (Whitaker, 2006).

Shelled peanut lot sizes in US commerce are typically 20 metric tons (MT) or ~44,000 lbs but more rarely can be as large as 200,000 lbs. In the United States, the Federal State Inspection Service (FSIS) via USDA Agricultural Marketing Service (AMS) regulations, serves as the unbiased third party that collects official samples from positively identified shelled peanut lots. Sampling of shelled lots is best accomplished when the lot is being conveyed, typically just before final packaging, when multiple, regularly spaced incremental samples over the production of the lot can be automatically collected and subsequently aggregated (Whitaker, 2006). USDA AMS regulations require a 70-75 kg aggregate sample per lot (Whitaker and Slate, 2012). Approved random dividers then separate this aggregate sample into four samples, three of which are ~22 kg and used for aflatoxin testing per official USDA procedures. The remaining portion of the official aggregate sample is used by FSIS to grade the lot for damage, moisture, foreign material, and various kernel size parameters, among other grade factors.

The minimum sample considered for aflatoxin testing in the USDA AMS peanut program is 22 kg, which is commonly used to make a decision on a 20 MT lot. This represents about 0.11% of the lot in question. Ideally, the *entire* sample would be extracted for downstream analysis; however, it is currently not technologically practical from a time and resource perspective to entirely extract a 22 kg sample of peanuts. Sample preparation is the operation of converting the sample to a format that when subsampled, the smaller subsample, which will be used for downstream extraction and analysis, best represents the mean sample composition (Stoloff et al. 1969; Dorner and Cole, 1993; Whitaker and Slate, 2012). To accomplish this goal, the particle size of the sample should be reduced and the sample mixed. Two mills are approved by USDA for this purpose: One is commonly called a Dickens Mill (DM) which was originally developed by the USDA Agricultural Research Service for the peanut industry and USDA AMS to provide sample comminution and automated subsampling of comminuted samples prior to aflatoxin testing (Dickens and Satterwhite, 1969; Dickens et al. 1979; Whitaker and Slate, 2012). In brief, the DM has a series of rotating blades surrounded by a cylindrical screen with 3.2 mm round holes. Screen hole size is designed to minimize peanut pasting. As the sample is ground and forced through the screen, small openings in the outer wall of the mill are connected to spouts (typically two) which provide for automatic sub-sample collection. Vertical vanes are spaced about 20 degrees apart around the openings so that about 5% of the comminuted sample are captured by each spout. The second is a Vertical Cutter Mill (VCM) which has been previously described (Francis, 1979; Dorner and Cole, 1993). A VCM consists of a stainless steel scraped surface bowl and blade at the base of the bowl which when properly sized both mixes and converts the sample to a paste, i.e. the oil cells of the kernels are disrupted. Particle size reduction is improved about 10X for a VCM versus DM grind (Francis, 1979).

Regardless of mill type, subsamples are collected for downstream extraction and analysis. In the official USDA extraction method, subsamples are first mixed with water (1.455 water to peanut solids ratio) and made into a slurry using a high speed blender. USDA AMS regulations allow anywhere from 900-1300 grams to be subsampled and slurried. This water slurry procedure was designed to provide further particle size reduction/mixing and to save on downstream organic solvent usage (and subsequent disposal of solvents) necessary for extraction (Whitaker et al. 1980). Water slurry as a

component of sample preparation has been shown to decrease subsample variation and/or increase total aflatoxin measured in a number of commodities including peanut and pistachio (Whitaker *et al.* 1980; Schatzki and Toyofuku, 2003; Spanjer *et al.* 2006)

After slurrying, a portion of the slurry is removed for downstream aflatoxin extraction and subsequent analytical measurement. Traditionally, thin layer chromatography (TLC) was the final analytical technique for aflatoxin testing (Whitaker *et al.* 1974), and is still in limited use today, although TLC is falling out of favor due to the introduction of more accurate, precise, and user friendly analytical techniques (Holcomb *et al.* 1992; Reiter *et al.* 2009). Additionally, compared to updated analytical techniques, most common TLC protocols require the increased use of organic solvents (Holcomb *et al.* 1992), which pose human and environmental health hazards and require expensive waste disposal costs. Extraction protocols (commonly methanol/water based) coupled with immunoaffinity columns (IAC) are nowadays the most commonly applied clean up method for aflatoxin analysis in the US, and when coupled to HPLC or fluorescence detection, this combination is a reference method for measuring aflatoxin in food and feed (Reiter *et al.* 2009). The importance of IAC's in mycotoxin testing has been reviewed (Scott and Trucksess, 1997). HPLC or (benchtop) fluorometry coupled with upstream IAC clean up are approved by AOAC International (Trucksess *et al.* 1991) and by USDA AMS as an analytical technique in the shelled peanut aflatoxin sampling program. Regardless of the specific analytical technique, two measurement replications from a given subsample extract are required for all analytical measurements in the USDA peanut aflatoxin sampling program. These extract replicates are designated A and B, and the two results are averaged for the final test result. These measurements are designated as the 1AB, 2AB or 3AB where 1, 2, and 3 designate the three official 22 kg samples being prepared for aflatoxin analysis.

The US Peanut industry follows a modified sequential plan where one, two or three samples are prepared and analysed for aflatoxin to either accept or reject the lot of raw peanuts being considered for the edible market (Whitaker *et al.* 1995). The accept/reject limit selected by the US Peanut Industry for raw, shelled peanuts is an average of 15 ppb total aflatoxin for the 3 samples; however, specific accept/reject criteria are also in play after measuring the 1AB and 2AB. (The US Food and Drug Administration regulates finished products to 20 ppb.) Conceptually, this plan is designed to readily accept very good lots, i.e. those that should consistently test below 8 ppb on the 1AB and reject very bad lots, i.e. those that should consistently test above 45 ppb. In turn, lots with means near the accept/reject limit of 15 ppb will be statistically more likely to have multiple samples tested under this plan for final lot classification (Whitaker and Slate, 2012).

Considering the infrequent kernel to kernel nature of aflatoxin contamination, the negative binomial distribution has been used to mathematically model aflatoxin contamination in actual shelled lots (Whitaker *et al.* 1972). In this theoretical distribution, there is a high frequency of finding kernels with little to no aflatoxin contamination and a small frequency of finding kernels with extremely high aflatoxin contamination. The assumption about the aflatoxin distribution among individual kernels was based upon the experimental observations of the aflatoxin distribution among replicated sample test results taken from contaminated shelled peanut lots (Whitaker *et al.* 1972). The parameters of the negative binomial distribution were calculated from the experimentally measured variability (variance) associated with sampling, sample preparation, and analytical steps associated with the aflatoxin test procedure used to measure aflatoxin in shelled peanuts (Whitaker *et al.* 1974). By measuring the variability among sample test results and assuming the distribution among sample test results was negative binomial, the aflatoxin contamination rate among kernels in a lot at 50 ppb was calculated to 0.01% (Schuller *et al.* 1976). Over the years, the negative binomial distribution and the measured variances have been used to evaluate the risks of misclassifying lots for various aflatoxin sampling plans for the USDA (Whitaker and Dickens, 1979) and designing the current aflatoxin sampling plan used by USDA and the peanut industry (Whitaker and Slate, 2012).

Given the heterogeneous nature of aflatoxin contamination in shelled peanuts, sampling is the largest source of variation in the sampling, sample preparation and analytical chain to determine a test result (Whitaker and Slate, 2012). When using a DM mill to prepare a 22 kg sample, and HPLC for analytical measurements, the sampling, sample preparation, and analytical steps of the aflatoxin test procedure were estimated to account for 88.1, 8.1, and 3.7 percent of the total variance of the aflatoxin teste procedure in a lot at 15 ppb (Whitaker and Slate, 2012). Given this high opportunity for sampling variation, everything possible should be done to best prepare and analyse the official samples. It is recognized that

there is inherent natural variation among subsamples collected for a given sample that has been milled (Stoloff *et al.* 1969; Whitaker *et al.* 1974; Dorner and Cole, 1993). Subsample variation is a function of sample contamination, degree of particle size reduction, degree of homogenization and subsample size (Stoloff *et al.* 1969; Whitaker *et al.* 1974; Schatzki and Toyofuku, 2003). The improved particle size reduction and mixing provided by the VCM decreases subsampling variation compared to a DM (Stoloff *et al.* 1969; Dorner and Cole, 1993; Whitaker and Slate, 2012).

Disagreement exists within the US Peanut industry regarding the performance of the two mills types used to prepare samples for aflatoxin testing and their impact on sample results. Some in the peanut industry have seemingly observed that sample test results prepared from samples ground via a DM tend to skew lower than samples ground via a VCM (Whitaker and Slate, 2012). It was argued that the trends buyers and sellers were observing were less a function of mill type, but instead the reality of sampling theory for sequential samples being considered in relation to the 1AB USDA cut-off of 8 ppb (Whitaker and Slate, 2012). Authors further argued that the theoretical contribution of sampling preparation variation to the final test result is low enough (less than 10% of theoretical total variation derived from sampling, sample preparation and analysis) such that mill type minimally impacts official results. We agree that the sampling statistics on sequential samples can be a source of confusion and must be recognized and understood as demonstrated by Whitaker and Slate (2012); however, our experience has been that mill type does meaningfully influence final results with all other factors being equal, i.e. equivalent subsample size, equivalent water slurry procedure, equivalent extraction procedure and equivalent analytical measurements. Specifically, we have observed that the DM provides less comminution compared to a properly sized VCM, hence aflatoxin distributions among DM subsamples remain more positively skewed (median smaller than mean) than subsample distributions derived from VCM's which are more normally distributed around the sample mean. Because of the skewed aflatoxin distribution among subsample concentrations, DM subsample results will more frequently be lower than those derived from a VCM, but when aflatoxin is detected after a DM grind, it is often much higher than the sample mean because pockets of high aflatoxin peanuts are not as well comminuted and mixed.

While readily achievable, there are costs associated with reducing variation in the aflatoxin testing program that must be balanced against other factors. HPLC instrumentation is the preferred analytical measurement for aflatoxin detection (Holcomb *et al.* 1992; Reiter et al. 2009), but this instrumentation is more expensive than fluorometry or TLC instrumentation. Our experience suggested that with a high performance IAC in place, total aflatoxin measured via fluorometry and HPLC is quite comparable. Regardless, analytical variation is typically less important to the total variation of the final test result compared to subsampling variation and sampling variation. Accordingly, these experiments were designed to compare variability and distribution among subsample test results associated with the DM and VCM mill types while simultaneously comparing analytical performance (accuracy) of fluorometry to HPLC after IAC clean-up and concentration (Trucksess *et al.* 1991). (Trucksess *et al.* 1991). Such information can aid decisions on how to best spend valuable resources to deliver the safest peanuts possible.

## Materials and Methods

*Materials*. A lot of commercial medium runner peanuts from the 2015 US crop that did not pass the official USDA aflatoxin program was identified in cold storage. Considering the official samples for this lot, the 1AB and 2AB measured a 37 and 45 ppb total aflatoxin, respectively, with an overall average of 41 ppb. From this lot, a 1 MT tote was randomly selected for these research purposes. This tote was subsequently repackaged using a scoop into 46 individual samples, each weighing about 22 kg (48-49 lb).

*Sample Milling*. Samples were milled with one of two mill types: 1) DM or 2) VCM. Two DM's were used for these experiments, and the units were used alternately and equally throughout the experiments. Each DM had two spouts where subsamples are automatically discharged during milling. This is the standard design for a DM in the US peanut industry. As the DM blades rotate, they are surrounded by a cylindrical screen with 3.2 mm round hole openings and all ground material is forced through this screen. The two sup-sampling spouts correspond to two channels surrounding this screen, and these two channels each collect about 1100 g of the material passing through the screen at those locations. The remainder of the milled peanuts passes through the screen, but is not collected in the two subsample channels. Instead, this milled material falls out the bottom of the mill; however, this 'fall thru' is equivalent in particle size

reduction and comminution to that collected via the subsampling spouts. When milling a sample with a DM, samples were slowly metered into the DM's and subsamples collected from the two spouts built into the mill. Additionally, for some experiments, the 'fall thru' was also collected for subsequent testing.

Two Stephan (Hameln, Germany) VCM's were used in this research, a model 44 and model 60, with 44 L and 60 L stainless bowl capacities, respectively. Sharpened, dual serrated blades were used for both VCM's and each had 220V, 33A motors with 3600 RPM in addition to motorized, scrape surface baffles. The two VCM's were used alternately and equally throughout the experiments, and all samples were milled for 6 minutes.

After VCM milling, paste was manually subsampled by randomly collecting milled sample around the bowl with a spatula. For a given subsample, roughly equal portions of paste are taken from each of four quadrants within the bowl and aggregated to provide an 1100 g subsample.

***Subsample Water Slurry Preparation and Extraction.*** After milling (VCM or DM) all subsamples were slurried with water per USDA guidelines for the domestic, edible, shelled peanut aflatoxin sampling and testing program. While USDA guidelines for preparing water slurries allow milled subsamples weights to range from 900-1300 grams, all starting peanut material weights were 1100 grams and corresponding water weight was 1600 grams. Slurries were blended on high for 3 min. Immediately after slurry preparation, a 122.8 gram portion (50 gram peanut equivalent) was weighed into a tared blender jar, and 10.0 g of NaCl were added along with 177 ml of 85/15 methanol/water. This slurry portion with added methanol/water and NaCl was subsequently blended on high for an additional 2 min. Methanol and NaCl were both American Chemical Society grade. After blending, the material was gravity filtered with a P8 filter paper from Fisher Scientific (Fair Lawn, NJ). 20 ml of filtered extract was mixed 1:1 with deionized water and then gravity filtered through a G6 glass fiber filter from Fisher Scientific (Fair Lawn, NJ). Post filtration, 10 ml of the extract was then passed through disposable IAC's (Pi Biologigue; Seattle WA) at 1-2 drops/second using a vacuum manifold. Columns were then washed twice with 10 ml of deionized water before elution with 1 ml of methanol. If total aflatoxin measured above 60 ppb at the final analytical stage, the extract was diluted 1:10 and rerun thru a new IAC.

***Analytical Procedure.*** Post IAC, final eluates were analyzed for total aflatoxin via 1) fluorometry or 2) HPLC based on AOAC Method 991.31. For fluorometer measurements, the 1 ml eluates were diluted 1:1 with AflaTest® developer (VICAM; Watertown, MA) and total fluorescence measured via a VICAM (Watertown, MA) Series-4EX Fluorometer, and measured fluorescence was converted to ppb based on known calibration standards measured daily.

For HPLC measurements, the 1 ml eluates were diluted 1:1 with 1% acetic acid and loaded into HPLC vials. An Agilent HPLC 1100, equipped with a fluorescence detector, and a post-column photochemical reactor for enhanced detection (PHRED) (Aura Industries, NY, NY) was used for aflatoxin measurements. Fifteen μl of solution was injected onto a 4.6 × 150mm, Waters Nova-Pak C18 4μm analytical column held at 30°C. An isocratic method using 55:45 water: methanol [HPLC Grade] at a flow rate of 1.0 ml/min was used to separate individual aflatoxins at a pressure of approximately 200 bar. PHRED-enhanced peaks were detected by fluorescence with an excitation wavelength of 360 nm and emission wavelength of 440 nm. Aflatoxin $B_1$, $B_2$, $G_1$, and $G_2$ were quantified using a purchased 4-component aflatoxin mix from Supelco (Bellefonte, PA). Standard curves were generated daily prior to analyses. Quality assurance included matrix spikes, blanks and daily control.

## Results and Discussion

Sampling, sample preparation and analytical variation were systematically investigated by measuring aflatoxin distributions within a 1 MT tote of medium runners. As detailed in the Materials section, this tote was isolated from a shelled lot that failed the official USDA aflatoxin sampling program. In the first experiment, three sets (66 kg/set) of peanuts were prepared by selecting three samples (22 kg) randomly from the repackaged tote, thoroughly mixing these samples using a 3-way splitter and then dividing equally into three new samples (22 kg) per set. This generated three sets, three samples each, for a total of 9 samples. Three samples of this size are equivalent to the total sample available in the official USDA sampling and testing program for aflatoxin. All samples were first milled with a DM and the two spout subsamples collected. Additionally, the corresponding 'fall thru' from each sample was collected, milled in a VCM, and four subsamples collected after the VCM grind to determine the average aflatoxin in the 'fall thru'. This VCM

**Table 1. Comparison of aflatoxin measurements for subsamples after DM and VCM sample preparation. Three sets, 3 samples each, were prepared by thoroughly mixing ∼66 kg of medium runners which were then randomly split into 3 samples (22 kg). All samples were randomly selected from a 1 MT tote of medium runners that was in turn randomly selected from a 20 MT lot that had failed aflatoxin (1AB, 2AB average = 41 ppb). Each of the 9 samples was first milled via a DM and the two automatic subsamples (spouts) collected. Additionally, the 'fall thru' after DM milling was collected, milled in a VCM and then 4 subsamples collected. All subsamples, regardless of mill type, were extracted and analyzed equivalently, including water slurry preparation, extraction, IAC column cleanup and measurement of total aflatoxin via fluorometry.**

| Set | Sample | Dickens Mill Grind | | | | | VCM Grind - DM 'Fall Thru' | | |
| | | Spout 1 (ppb) | Spout 2 (ppb) | AVG (ppb) | SD (ppb) | CV (%) | AVG (ppb) | SD (ppb) | CV (%) |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 1 | 1 | 345.0 | 11.0 | 178.0 | 236.2 | 132.7 | 41.4 | 11.6 | 28.0 |
| 1 | 2 | 55.0 | 52.0 | 53.5 | 2.1 | 4.0 | 41.6 | 7.6 | 18.1 |
| 1 | 3 | 6.7 | 20.0 | 13.3 | 9.4 | 70.8 | 12.4 | 1.3 | 10.6 |
| 2 | 1 | 0.8 | 3.6 | 2.2 | 2.0 | 90.3 | 6.9 | 0.6 | 9.0 |
| 2 | 2 | 23.5 | 29.5 | 26.5 | 4.2 | 16.0 | 30.5 | 3.9 | 12.6 |
| 2 | 3 | 3.3 | 40 | 21.7 | 26.0 | 119.9 | 17.1 | 0.5 | 2.8 |
| 3 | 1 | 13.5 | 20 | 16.8 | 4.6 | 27.4 | 4.7 | 0.4 | 8.7 |
| 3 | 2 | 7.0 | 19 | 13.0 | 8.5 | 65.7 | 14.0 | 0.8 | 5.8 |
| 3 | 3 | 11.5 | 14 | 12.8 | 1.8 | 13.9 | 19.6 | 0.8 | 3.8 |
| | | | **AVG** | 37.5 | | 60.1 | 20.9 | | 11.1 |

milled 'fall thru' was assumed to have a particle size reduction similar to a typical VCM grind.

All subsamples, regardless of the mill used in preparation, were made into water slurries per USDA AMS specifications and extracted for aflatoxin per USDA AMS specifications using IAC's coupled with fluorometry or HPLC to measure total aflatoxin. Both fluorometry and HPLC coupled with IAC clean-up have been recognized as official AOAC International methods since 1991 for aflatoxin detection in peanuts (Trucksess et al. 1991). Total aflatoxin as measured by IAC-fluorometry across this experimental matrix is summarized in Table 1. Before discussing subsample and sampling variation observed in this experiment, several points on the analytical measurements will be discussed.

USDA AMS regulations require that all subsample extracts be split and analysed in duplicate, and this protocol was followed for data collected in Table 1. This practice derives from the origin of the US aflatoxin sampling program when TLC was the primary analytical option. The analytical variation for TLC is relatively high compared to other analytical methods such as HPLC (Whitaker et al. 1996) and averaging two measurements from a given subsample extract reduces this analytical variation by half (Whitaker and Dickens, 1981). Considering methanol/water extraction followed by IAC cleanup coupled with either fluorometry or HPLC measurements, earlier work internal to our lab had demonstrated excellent analytical repeatability for duplicate extracts. Accordingly, an excellent linear correlation ($R^2$ = 0.96) was observed across the experimental matrix after splitting each subsample extract, passing thru a pair of IAC's (designated A and B) and measuring total aflatoxin by fluorometry (Figure 1). The slope (0.965) near 1.0 and a y intercept less than 1.0 ppb, further demonstrate the excellent agreement for these measurements and corresponding performance of the immunoaffinity columns. Note that data from spout1, sample1, set1 was excluded from Figure 1, as these data were much higher than other paired points and would distort visualization for the majority of the data. These duplicate AB samples were 350 and 340 ppb for an average of 345 ppb.

HPLC is the most accurate and precise analytical measurement *commonly* used to measure aflatoxin globally in a variety of agricultural commodities (Reiter et al. 2009). Our experience suggested that with an IAC in place, total aflatoxin measured via fluorometry and HPLC is quite comparable (accuracy). The two analytical measurements were compared for equivalent subsample extracts (Figure 2). For this comparison, the filtered, diluted subsample extracts prepared from sets 2 and 3 (a total of 36 subsamples from Table 1) were split and passed thru three IAC's, two coupled with fluorometry and one coupled with HPLC. Comparison of total aflatoxin measured for fluorometry and that measured by HPLC showed an excellent correlation ($R^2$ = 0.97) (Figure 2). A primary advantage of HPLC compared to fluorometry, is that HPLC allows separation and quantitation of the individual aflatoxin types. Aflatoxin $B_1$ is the most carcinogenic aflatoxin
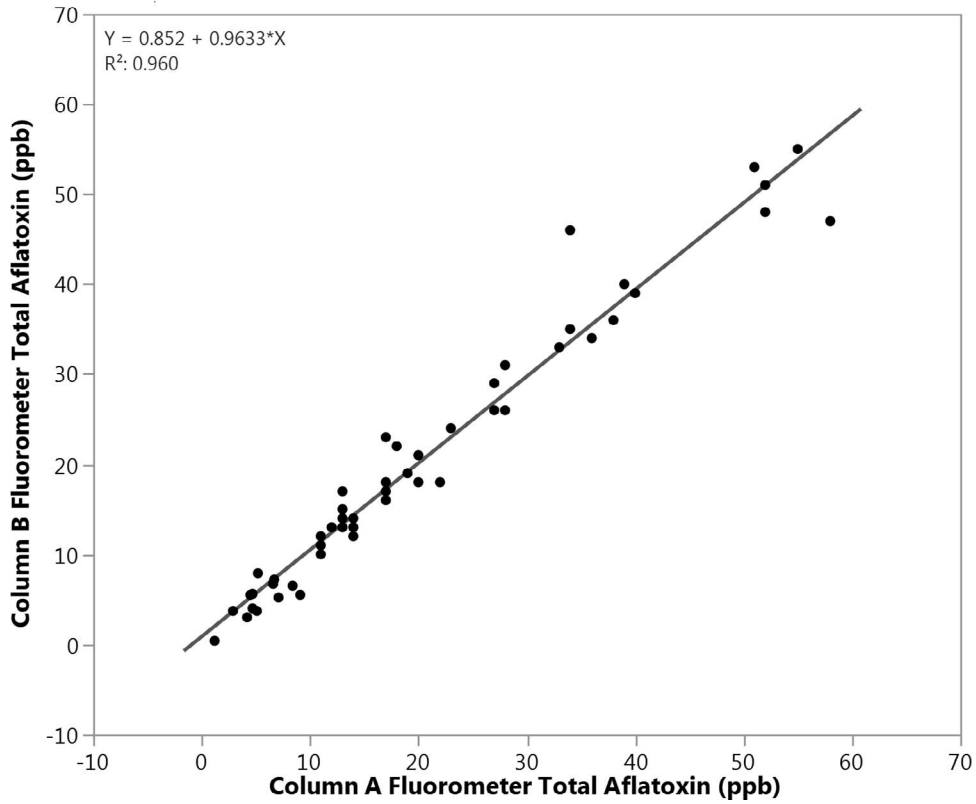
Fig. 1. Comparison of fluorometer total aflatoxin measurements prepared from equivalent subsample extracts and two immunoaffinity columns, A and B.
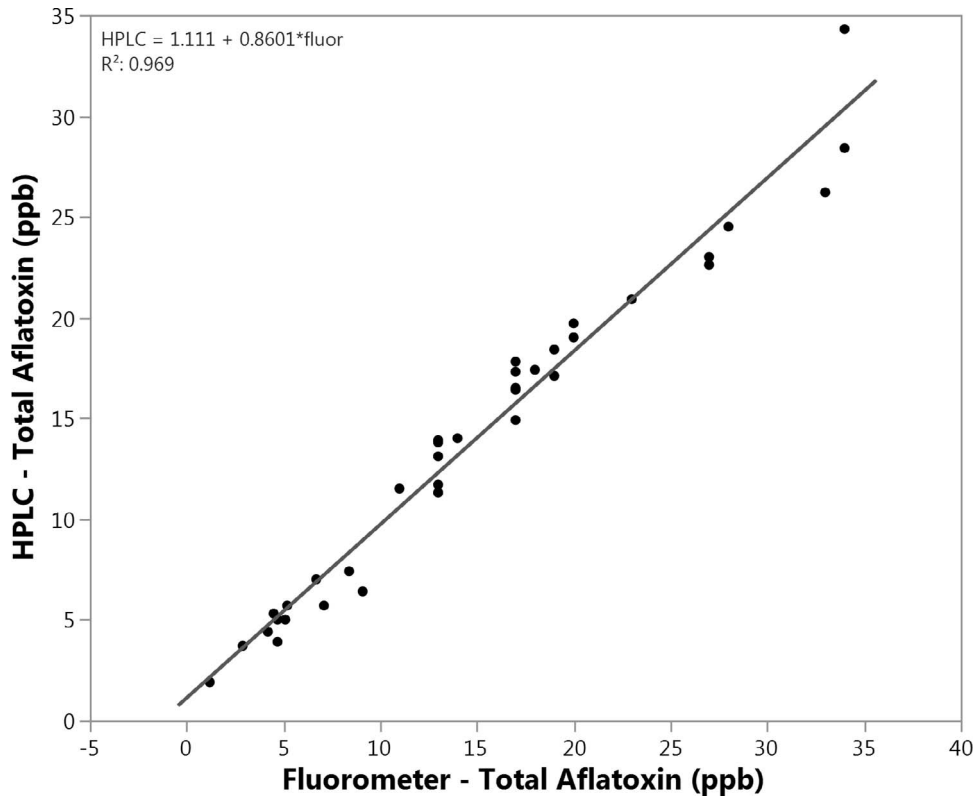


Fig. 2. Comparison of total aflatoxin measured by fluorometry vs total aflatoxin measured by HPLC for subsample extracts split prior to immunoaffinity cleanup. See manuscript for details.

type, and it typically predominates in nature (IARC 2012). Across all HPLC measurements completed for this study, aflatoxin $B_1$ averaged 81.4 $\pm$ 2.8% of the total aflatoxin detected. If measurements of the individual aflatoxin types are required, for example to meet EU export peanut aflatoxin regulations, than HPLC is an appropriate option. However, this data demonstrates that for a given subsample extract coupled with an IAC equivalent in performance to those used here, total aflatoxin measured by fluorometry and HPLC are practically equivalent in accuracy. IAC-fluorometry is more operationally efficient and cost effective than IAC-HPLC in a commercial setting. Given these benefits and the essentially equivalent accuracy compared to IAC-HPLC, IAC- fluorometry data was used when analyzing subsample and sampling variation. Future studies will be designed to measure the precision of the two methods.

Subsample variation for the paired DM subsamples is directly observed and CV's ranged from 4.0 to 132.7%, and average 60.1% across the nine samples (Table 1). The very high CV's observed at times in this experiment have important practical considerations. For example, in 5/9 samples prepared with a DM, one subsample was below the 15 ppb USDA regulatory limit for shelled peanuts, while the other subsample exceeded this limit, and at times this difference was excessive. For example, data observed in set1, sample1 (11 and 345 ppb) and set2, sample3 (3.3 and 40 ppb). In practice, lot classifications are based on one subsample, so in these cases the true value of the samples would be grossly underestimated or overestimated after DM preparation.

To further understand sample preparation performance, the 'fall thru' (see Methods) from the various DM preparations was collected, milled in a VCM and then 4 subsamples (1100 gram) removed, slurried, extracted, passed thru IAC's and total aflatoxin measured (Table 1). Four subsamples were collected from the VCM prepared 'fall thru' to better measure the average aflatoxin in the 'fall thru'. If the DM subsamples were representative of the overall sample mean, then the average of the 'fall thru' should correlate to the corresponding spout data, but large differences are readily observed. The majority of DM subsamples had values lower than the average measured in the corresponding VCM 'fall thru', but in some cases, aflatoxin measured in the DM subsamples was dramatically elevated, namely spout1 from set1, sample1 and spout2 from set2, sample3. This corresponds to 'hot spots' in the sample, i.e. contaminated kernels that the DM did not comminute sufficiently to disperse through the

prepared sample. It is established that VCM sample preparation decreases subsampling variation compared to that of a DM (Dorner and Cole, 1993; Whitaker and Slate, 2012), and this data provides some examples of how DM sample preparation can lead to good lots being rejected and bad lots being accepted.

When present in peanuts, aflatoxin contamination is highly positively skewed (median lower than mean), on a kernel to kernel basis, and sampling is the largest source of variation in the final test result determining lot acceptability (Whitaker, 2006; Whitaker and Slate 2012)). Sampling variation in the current experiment can be observed in Table 1, when considering VCM fall-thru averages for the 9 samples, all of which ultimately derived from a single 1 MT tote. Individual sample averages for the VCM fall thru ranged from 4.7 to 41.6 ppb. Set averages (three samples each) for 1, 2 and 3, were 31.8, 18.2 and 12.8 ppb, respectively, and the overall average was 20.9 ppb. As the sample size increases, sampling variation is decreased around the true lot mean (Whitaker, 2006), and that is observed in this example for the true tote mean when considering individual sample averages versus set averages (Table 1).

Given the large differences observed in aflatoxin measured for subsamples prepped from the DM versus a VCM, further experiments were pursued to understand implications of these differences. From the repackaged tote, 11 additional samples (representing approximately 24% of the tote) were pulled at random, prepared via VCM, and an additional 11 samples were pulled at random and prepared via DM. From each milled sample, a minimum of 12 individual 1100 gram subsamples were selected, slurried, extracted, and analysed post IAC for total aflatoxin via a fluorometer. Note that for DM subsamples, the two subsample spouts were collected and the 'fall thru' was riffle divided extensively before manually collecting additional subsamples. This riffle dividing, if anything, provided additional mixing for the 'fall thru' which is equivalent in particle size reduction and mixing for that collected in the spouts. The median, mean, standard deviation and CV among subsamples from each of the 22 samples are summarized in Table 2. The standard deviation and CV reflects the sum of the sample preparation and analytical variability. While the variability associated with the fluorometer and HPLC was not measured in this study, it was assumed to be minimal. Data were ordered basis sample means from low to high, for either the VCM or DM sample preparation. Subsample CV averaged 20.4 and 75.8 % for the VCM or DM, respectively (Table 2), values similar

**Table 2. Summary of total aflatoxin measured for multiple subsamples (1100 gram) prepared after milling samples (48-49 lb) with either a DM or VCM. 11 samples were milled using a VCM and 11 samples were milled using a DM. All samples were randomly selected from a 1 MT tote of medium runners that was in turn randomly selected from a 20 MT lot that had failed aflatoxin (1AB, 2AB average = 41 ppb). After milling each sample and collecting multiple 1100 gram subsamples (minimum of 12 per sample), all subsamples were extracted and analyzed equivalently, including water slurry preparation, extraction, IAC cleanup, and total aflatoxin measurement via fluorometry. Medians, means, standard deviations and CV among subsamples for all samples were calculated.**

| | Vertical Cutter Mill | | | | | Dickens Mill | | | |
|---|---|---|---|---|---|---|---|---|---|
| Sample | Median | Mean | Std Dev | CV | Sample | Median | Mean | Std Dev | CV |
| 1 | 2.8 | 3.2 | 1.5 | 46.4 | 12 | 5.8 | 6.3 | 3.4 | 53.1 |
| 2 | 3.8 | 4.0 | 1.1 | 26.4 | 13 | 6.2 | 7.1 | 3.9 | 55.2 |
| 3 | 13.0 | 14.7 | 6.0 | 41.1 | 14 | 5.6 | 7.1 | 4.3 | 60.8 |
| 4 | 18.0 | 18.1 | 1.3 | 7.2 | 15 | 7.3 | 8.9 | 4.4 | 49 |
| 5 | 18.3 | 18.6 | 2.0 | 10.6 | 16 | 10.5 | 11.5 | 5.9 | 51.5 |
| 6 | 20.5 | 21.5 | 4.0 | 18.9 | 17 | 16.0 | 18.3 | 9.6 | 52.3 |
| 7 | 24.0 | 25.1 | 3.2 | 12.9 | 18 | 18.0 | 19.4 | 12.1 | 62.5 |
| 8 | 25.0 | 25.2 | 4.7 | 18.6 | 19 | 18.0 | 31.8 | 45.8 | 144.1 |
| 9 | 26.0 | 26.5 | 4.8 | 18.1 | 20 | 21.0 | 35 | 37.2 | 106.2 |
| 10 | 27.5 | 28.3 | 4.1 | 14.4 | 21 | 29.0 | 43.6 | 42.2 | 96.7 |
| 11 | 45.0 | 45.4 | 4.6 | 10.2 | 22 | 40.8 | 73.5 | 75.2 | 102.2 |
| average: | 20.4 | 21.0 | 3.4 | 20.4 | | 16.2 | 23.9 | 22.2 | 75.8 |

to that measured in Table 1, and further demonstrating the improved performance of a VCM type mill. Earlier work with a DM, 280 gram subsamples and TLC analysis demonstrated a subsampling CV of about 50-30% depending on the average aflatoxin within a milled sample (Whitaker *et al.* 1974), values that are generally lower than that of the current study. In a another study, peanut subsample CV averaged 26.0% using a VCM and 47.0% for DM for 4 kg samples across a range of aflatoxin concentrations (Dorner and Cole, 1993). These CV values were similar to the current study; however, there were differences in study conditions, including VCM (smaller size previous versus larger size current), sample size (4 kg previous versus 22 kg current), subsample procedure (100 gram direct extraction previous versus 1100 gram water slurry current) and analytical techniques (HPLC previous versus IAC-Fluorometry current).

Subsample variation as measured by the standard deviation for both the VCM and DM increased as sample means increased (Table 2). This response during sample preparation is well documented for a variety of aflatoxin contaminated commodities including peanuts (Dickens and Satterwhite, 1969; Spanjer *et al.* 2006). The increase in subsample variation versus sample means was much greater after DM preparation compared to that of a VCM, especially for average values above approximately 20 ppb (Figure 3). Corresponding CV's for subsamples tended to decrease with increasing sample means after VCM preparation; whereas for the DM samples, subsample CV increased with increasing sample means across this range of aflatoxin contamination (Table 2). This latter observation for DM subsample CV increasing with sample means was not observed in earlier research (Whitaker *et al.* 1974). Using equations in Figure 3, subsample standard deviations at 15 ppb were calculated as 12.2 and 2.9 ppb for DM and VCM, respectively. Corresponding variances would be 150.0 and 8.4 ppb$^2$ for DM and VCM, respectively, values that are substantially different than subsample variances of 7.9 and 0.8 ppb$^2$ for DM and VCM calculated by Whitaker and Slate (2012). These differences warrant further consideration in future studies.

For the current study, in addition to improved CV values for VCM preparations versus DM, overall averages of the mean and median values among subsamples were more equivalent after VCM preparation: mean = 20.4 and median = 21.0, compared to DM preparation: mean = 23.9 and median = 16.2 (Table 2). This reflects the more positively skewed subsample distributions after a DM versus a VCM sample preparation. Aflatoxin histograms comprising all subsamples (N = 156) prepared from either of the two mill types for the 22 samples prepared reveal two distinct distributions (Figure 4). VCM derived subsamples have a more normal distribution around this tote. Alternatively, DM subsamples display more of a positively skewed distribution where the mean is greater than the median (Figure 4). Such a distribution reflects the typical kernel to kernel distribution of aflatoxin contamination in a shelled lot of peanuts, that is many kernels are aflatoxin free and some much smaller percentage are
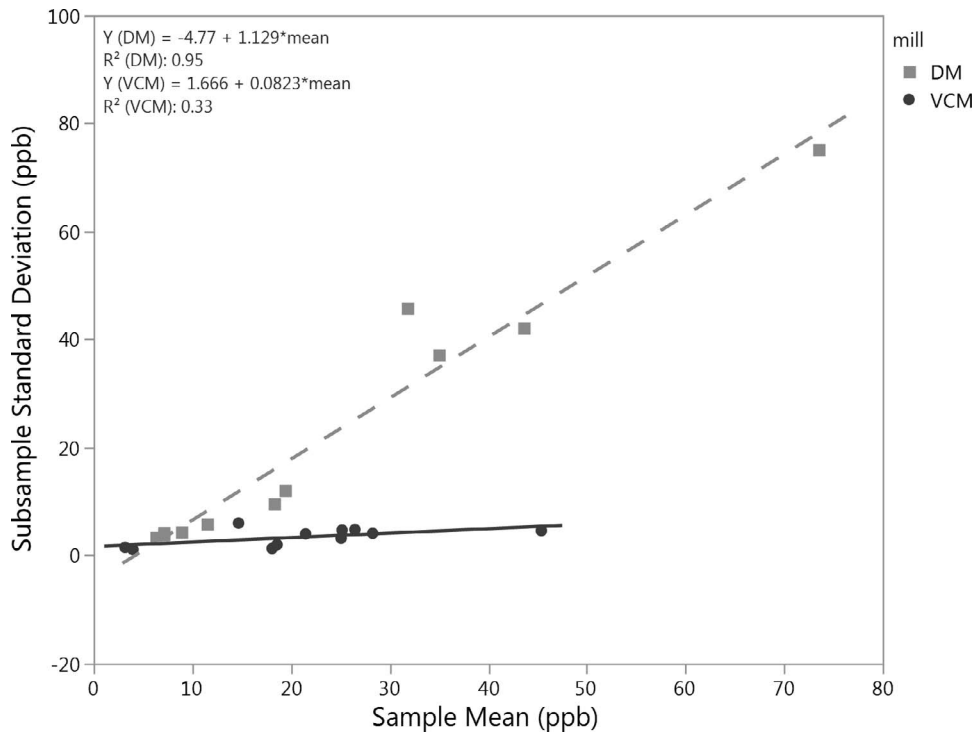
Fig. 3. Subsample standard deviation versus sample mean for VCM and DM preparations representing eleven ∼22 kg samples for each mill. Samples were randomly selected from a 1 MT of medium runners that was in turn randomly selected from a 20 MT lot that had failed aflatoxin (1AB, 2AB average = 41 ppb).

potentially highly contaminated. Due to sampling variation, it is not expected that the incoming aflatoxin loads for samples prepared by VCM versus DM were equivalent; however, given that 11 samples represents a large (∼24% by mass) portion of the parent tote, incoming aflatoxin should be comparable across the two sample sets, and this is reflected in the overall means: VCM mean = 20.9 and DM mean = 23.5. However, the subsample distributions around these means are clearly quite different as the DM subsamples skew lower with several high spikes. This has important practical consequences. As a reference line, 15 ppb which is the USDA accept/reject limit for edible lots is provided, and less than 28% of the subsamples for the VCM were below 15 ppb, whereas 56% of DM subsamples were below 15 ppb. The numerous spikes observed in the DM subsamples correspond to highly contaminated kernels that were not effectively comminuted within the sample after DM preparation, even with slurry preparation.

Subsample variation is inversely proportional to subsample size, and directly proportional to particle size reduction and degree of homogenization during milling (Stoloff et al. 1969; Schatzki and Toyofuku, 2003). As the official USDA peanut testing program evolved, the subsample size was increased from 275 to 1100 gram to improve subsample variation and skewness of the aflatoxin

distribution among subsamples (Whitaker and Slate, 2012). Related, a slurry step was incorporated into the official USDA aflatoxin testing program for shelled peanuts in the 1980's. Data demonstrated equivalent or improved subsample variation for the slurry preparation, and average aflatoxin measured increased about 15% after slurrying compared to the traditional extraction (Whitaker et al. 1980). Slurrying was demonstrated to provide further particle size reduction after the DM comminution while also saving on the excessive use of organic solvents necessary for TLC (Whitaker et al. 1980). This slurry requirement in the USDA peanut aflatoxin sampling/testing program remains today and was utilized for all data collected in this study. While a useful strategy for decreasing subsample variation, this data demonstrates that increasing the subsample size from 275 to 1100 g and incorporating the slurry step does not normalize the aflatoxin distribution among subsamples in the sample preparation of the sample with a DM to that of a VCM. It is also noted that while slurry incorporation does reduce use of organic solvents, the use of a slurry brings its own disposal challenges which is a cost carried by the testing lab (Spanjer et al. 2006).

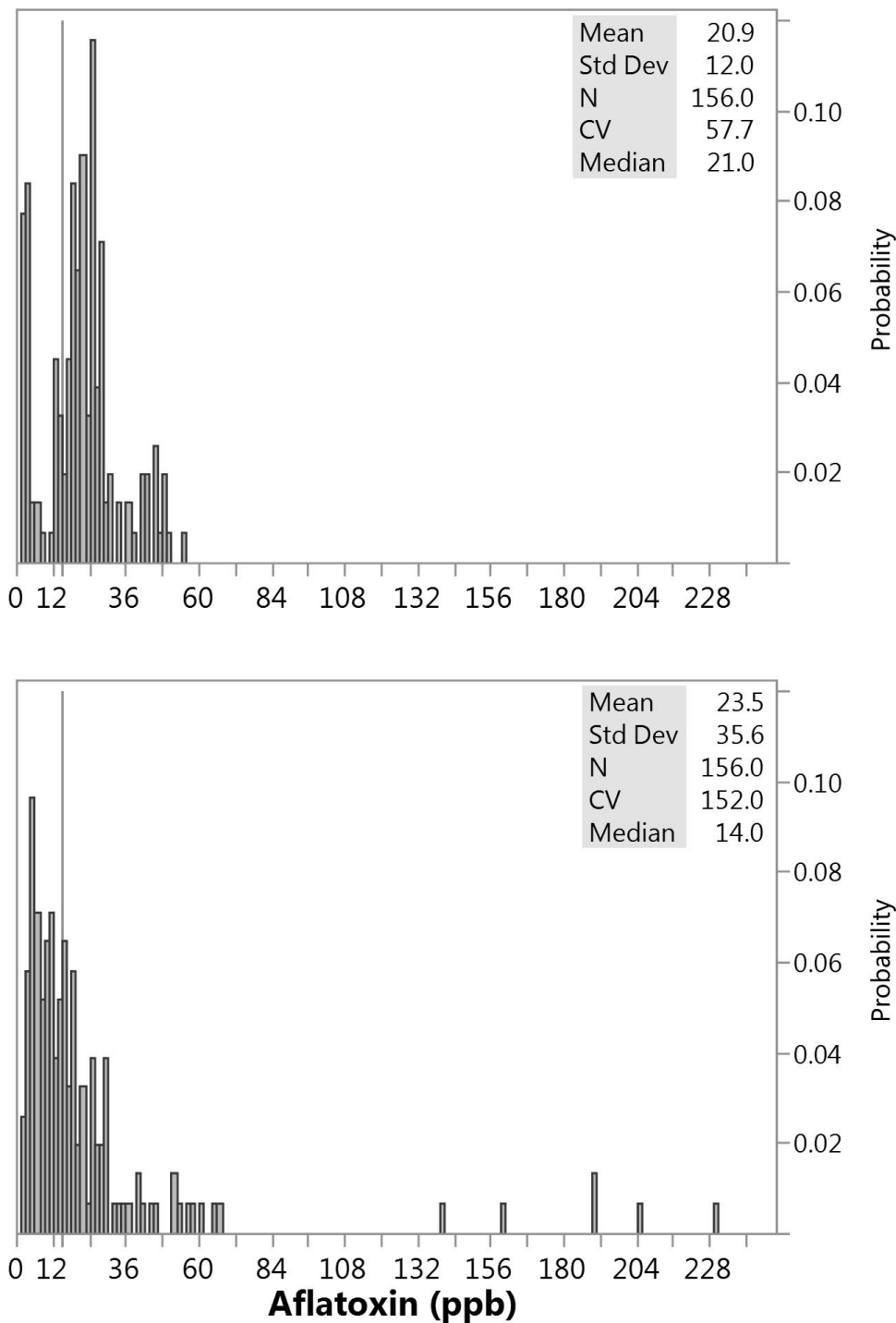There are opportunities to reduce variation in an aflatoxin sampling/testing program, but the

**Fig. 4.** Aflatoxin histograms for multiple subsamples (1100 gram) prepared after milling multiple samples (22 kg) with either a DM or VCM. Eleven samples were milled using a VCM (top panel) and 11 samples were milled using a DM (bottom panel). Samples were randomly selected from a 1 MT of medium runners that was in turn randomly selected from a 20 MT lot that had failed aflatoxin (1AB, 2AB average = 41 ppb). After milling each sample and collecting multiple 1100 gram subsamples (minimum of 12 per sample), all subsamples were extracted and analyzed equivalently, including water slurry preparation, extraction, IAC cleanup, and total aflatoxin measurement via fluorometry. A reference line is provided at 15 ppb, which is the USDA accept/reject limit.

benefits of implementing such opportunities must be balanced against their costs (Whitaker 2003). As previously discussed, when contamination is present, sampling is typically, by far, the largest source of variation in the sampling, sample preparation and analytical chain. While having optimized sample preparation and analytical steps in place is critical to best understand sampling variation, the sampling variation can only be reduced by considering larger sample sizes. This obviously

comes at a cost that must be balanced against various factors including risk exposure

Given sample preparation, this data provides some practical examples of how DM sample preparation can increase the chances of good lots being rejected and bad lots being accepted. The price of a DM is estimated to be $5000 whereas a VCM equivalent to those used in this research, which are sized to effectively process 22 kg of shelled peanuts, is estimated to be closer to $40,000. The VCM must be appropriately sized and built, as it has been reported that the conversion of oilseeds to a paste can clog mills and prevent mixing during sample preparation (Spanjer et al. 2006). When appropriately sized and utilizing a scrape surfaced design, our experience and this data demonstrates pasting to be desirable as it promotes mixing of the VCM contents. Clearly, there is a procurement cost to reduce sample preparation variation. Beyond procurement, the operational costs and turnaround time of the mill types should also be considered. DM are flow through type mills and grind time is proportional to the size of the sample being processed (Dorner and Cole, 1993). For 22 kg samples used in this research (and in the US commercial trade), we estimate the DM mill grind time to average 8-10 min. Milling time for a DM varies substantially as the technician, who may have to stand on an elevated platform, must transfer the sample contents into a funnel at the top of the mill and then meter the peanuts in slowly so as not to clog the mill. In addition to longer grind times, a DM preparation is more labor intensive versus a VCM grind, as for the latter the entire sample contents are poured into the mill at once. Cleanup for a DM is at a minimum 6-8 min, as air compressors and brushes must be used to carefully clean the many crevices in the instrument, which will retain the oily pieces of milled peanuts. During DM cleanup, respirators are advised for technicians given that aspiration is the primary cleaning mechanism (Spanjer et al. 2006). In contrast, a VCM cleanup is closer to 4-5 min: after subsampling, the technician tips the bowl on a built in swivel and removes the remaining paste to a container then follows a specific protocol of rinsing the unit with hot water, adding detergent, running the VCM briefly to clean, rinsing and drying. As the contents are pasted, no respirator is needed and the process is less labor intensive and more conducive for the rapid turnaround necessary to properly service the US peanut industry. These long term labor benefits are advantageous for operating VCM's versus DM's.

Given technological advances in analytical procedures over the past approximate 20 years, the analytical variation in the sampling, sample preparation and analytical chain to generate a test result is very low in modern aflatoxin sampling/testing programs. As previously mentioned, an exception is TLC, which while inexpensive to purchase instrumentation, has inherently high analytical variability, increased operational inefficiencies and requires the use of solvents which pose environmental and health risks. TLC is not advised for a modern aflatoxin sampling/testing program. However, while HPLC is the gold standard of common analytical techniques in the analysis of aflatoxin and other mycotoxins, our data demonstrates that with a high performance IAC in place during extraction, the more robust and cost effective fluorometer provide essentially equivalent measurements of total aflatoxin. Furthermore, the USDA requirement of duplicate analyses of extracts provides minimal improvement with IAC-fluorometry (or IAC-HPLC) in generating test results, and those resources could be better spent elsewhere in generating test results. An HPLC equivalent to that used in this research is estimated to cost $40,000 plus it has a very high operating cost, versus a fluorometer, which is estimated to cost $5000 and has very low operational costs. If available, resources would be better spent in improving sample preparation technologies then replacing IAC-fluorometer with IAC-HPLC.

## Summary and Conclusions

These experiments demonstrate the increased potential to misclassify shelled lots of peanuts in an aflatoxin testing program when preparing samples with a DM versus a VCM. A VCM is more expensive, but the clear benefits in better classifying aflatoxin contamination, operational efficiency and operator safety justify this expense. Furthermore, assuming an appropriate IAC is used upstream during extraction, a fluorometer can provide near equivalent performance in accuracy versus an HPLC for detecting total aflatoxin.

## Literature Cited

Cole, R.J., Dorner, J.W., Holbrook, C.C. 1995. Advances in mycotoxin elimination and resistance. Pages 456–474. In H.T. Stalker and H.E. Pattee (eds). Advances in Peanut Science, American Peanut Research and Education Society, Inc. Stillwater, OK.

Cucullu, A.F., Lee, L.S, Mayne, R.Y, Goldblatt, L.A. 1966. Determination of aflatoxins in individual peanuts and peanut sections. J. Amer. Oil Chem. Soc. 43: 89–92

Dickens, J.W., Satterwhite, J.B. Subsampling mill for peanut kernels. 1969. Food Technol. 23:90–92.

Dickens, J.W., Whitaker, T.B., Monroe, R.J., Weaver, J.N. 1979. Accuracy of subsampling mill for granular-materials. J. Amer. Oil Chem. Soc. 56: 842–844.

Dorner, J.W. Management and prevention of mycotoxins in peanuts. 2008. Food Addit. Contam. 25:203–208.

Dorner, J.W., Cole, R.J. 1993. Variability among peanut subsamples prepared for aflatoxin analysis with 4 mills. J. AOAC Int. 76:983–987.

Francis, O.J. 1979. Sample preparation of some shelled treenuts and peanuts in a vertical cutter-mixer for mycotoxin analysis. J. Assoc. Off. Anal. Chem. 62: 1182–1185.

Holcomb, M., Wilson, D.M., Trucksess, M.W., Thompson, H.C. 1992. Determination of aflatoxins in food-products by chromatography. J Chromatogr. 624:341–352.

International Agency for Research on Cancer. 2012. Aflatoxins. *IARC Monograph Chemical Agents and Related Occupations* 100F, 225–248

Kensler, T.W., Roebuck, B.D., Wogan, G.N., Groopman, J.D. 2011. Aflatoxin: a 50-year odyssey of mechanistic and translational toxicology. Toxicol. Sci. 120:S28–S48.

Lancaster, M.C., Jenkins, F.P., Philp, J.McL. 1961. Toxicity associated with certain samples of groundnuts. Nature. 192: 1095–1096.

Payne, G.A, Brown, M.P. 1998. Genetics and physiology of aflatoxin biosynthesis. Ann. Rev. Phytopathol. 36:329–362.

Reiter, E., Zentek, J., Razzazi, E. 2009. Review on sample preparation strategies and methods used for the analysis of aflatoxins in food and feed. Molec. Nutr. Food Res. 53:508–24.

Sargeant, K., Sheridan, A., O'Kelly, J., Carnaghan, R.B.A. 1961. Toxicity associated with certain samples of groundnuts - reply. Nature. 192:1096–1097

Schatzki, T.F., Toyofuku, N. 2003. Sample preparation and presampling of pistachios. J Agric. Food Chem. 51:6068–6072.

Scott, P.M., Trucksess, M.W. 1997. Application of immunoaffinity columns to mycotoxin analysis. J AOAC Int. 80:941–949.

Schuller, P.L, Horwitz, W., Stoloff, L. 1976. Review of sampling plans and collaboratively studied methods of analysis for aflatoxins. J. Assoc. Off. Anal. Chem. 59:1315–1343.

Spanjer, M.C., Scholten, J.M., Kastrup, S., Jorissen, U., Schatzki, T.F., Toyofuku, N. 2006. Sample comminution for mycotoxin analysis: dry milling or slurry mixing? Food Addit. Contam. 23:73–83.

Stoloff, L., Cambell A.C., Beckwith, A.C., Nesheim, S., Winbush, J.R., Fordham, O.M. 1969. Sample preparation for aflatoxin assay - nature of problem and approaches to a solution. J. Am. Oil Chem. Soc. 46:678–684.

Trucksess, M.W., Stack, M.E., Nesheim, S., Page, S.W., Albert, R.H., Hansen, T.J., Donahue, K.F. 1991. Immunoaffinity column coupled with solution fluorometry or liquid-chromatography postcolumn derivatization for determination of aflatoxins in corn, peanuts, and peanut butter - collaborative study. J. Assoc. Off. Anal. Chem. 74:81–88.

Whitaker, T.B., Wiser, E.H., Dickens, J.W., Monroe, R.J. 1972. Comparison of the observed distribution of aflatoxin in shelled peanuts to the negative binomial distribution. J. Am. Oil Chem. Soc. 49:590–593.

Whitaker, T.B., Dickens, J.B., Monroe, R.J. 1974. Variability of aflatoxin test results. J. Am. Oil Chem. Soc. 51:214–218.

Whitaker, T.B., Dickens, J.B. 1981. Errors in aflatoxin analyses of raw peanuts by thin layer chromatography. Peanut Sci. 8, 89–92.

Whitaker, T.B., Dickens, J.B., Monroe, R.J. 1980. A water slurry method of extracting aflatoxin from peanuts. J. Am. Oil Chem. Soc. 57:269–72.

Whitaker, T.B., Springer, J., Defize, P.R., DeKoe, W.J., Coker, R. 1995. Evaluation of sampling plans used in the United States, United Kingdom, and the Netherlands to test raw shelled peanuts for aflatoxin. J AOAC Int. 78:1010–1018.

Whitaker, T.B., Horwitz, W., Albert, R., Nesheim, S. 1996. Variability associated with analytical methods used to measure aflatoxin in agricultural commodities. J. AOAC Internat. 79: 476–485

Whitaker TB. Standardization of mycotoxin sampling procedures: an urgent necessity. Food Control. 2003;14(4):233–7.

Whitaker, T.B. 2006. Sampling foods for mycotoxins. Food Addit. Contam. 23:50–61.

Whitaker, T.B., Slate A.B. 2012. Comparing the USDA/AMS subsampling mill to a vertical cutter mixer type mill used to comminute shelled peanut samples for aflatoxin analysis. Peanut Sci. 39:69–81.