Size of Maize Sample Needed to Determine Percent Kernel Infection by *Aspergillus flavus*

Gene E. Scott and **Natale Zummo,** USDA-ARS and Mississippi Agricultural and Forestry Experiment Station, Mississippi State 39762

ABSTRACT

Scott, G. E., and Zummo, N. 1995. Size of maize sample needed to determine percent kernel infection by *Aspergillus flavus*. Plant Dis. 79:861-864.

Aspergillus flavus usually infects a low percentage of maize kernels, but it is a very important pathogen because it produces aflatoxin. The object of this study was to determine the needed kernel sample size per plot in replicated trials to ascertain the percentage of infection by *A. flavus* at a reasonable level of precision. We compared 65-, 130-, 195-, and 390-kernel sample sizes. Two experiments were simulated to determine which statistical attribute(s) was the most sensitive to changes in sample size. The changes in error mean squares and *F* values between 65- and 390-kernel samples were of a magnitude greater than six, which was larger than for the other statistical attributes. Based on averages of error mean squares, *F* values, and Spearman correlation coefficients over eight experiments with actual data, we concluded that a 390-kernel sample size is not too large, and a sample size of approximately 200 kernels should be considered a minimum. Based on other calculations, a minimum sample size of 100 kernels is needed just to have a reasonable probability of detecting the fungus when it is present.

Aspergillus flavus Link:Fr. produces aflatoxin in maize *(Zea mays* L.). Maize grain that contains more than 20 ppb aflatoxin cannot be sold in interstate commerce. Many countries will not accept maize grain with even less than this amount of aflatoxin.

Studies to evaluate for treatment or genotypic differences have been conducted utilizing number of kernels per sample of 40 to 50 (1), 50 (4), 100 (3), 200 (2), 260 (5), and 390 (6). These sample sizes apparently were chosen rather arbitrarily, or for convenience, because no study to determine the number of kernels needed per sample has been reported for *A. flavus* kernel infection.

The objectives of this study were (i) to determine which statistical parameter(s) was the most sensitive to changes in kernel sample size per plot in replicated trials, and (ii) to evaluate data from experiments when maize was inoculated with *A. flavus* to ascertain sample size requirement for separation of treatment or genotypic means for percent kernel infection by *A. flavus*.

MATERIALS AND METHODS

Data from eight experiments conducted from 1989 through 1992 were used in this study. The maize hybrids in these tests were grown in replicated 20-plant plots. Only the top ear on each plant was inoculated and harvested. Ears of most experiments were needle-inoculated (6) 6 days after midsilk, but in some experiments ears of some or all treatments and/or genotypes were pinbar-inoculated (1) 20 days after midsilk. Data from these experiments were considered to give a range of infection levels that would represent the levels of infection that might be obtained in studies on kernel infection with *A. flavus*. A listing of the experiments, with information on the number of entries, replications, and range of kernel infection, is given in Table 1.

All inoculated ears were harvested 60 days after midsilk, dried, and shelled. Ears that were needle-inoculated were mechanically shelled, and a sample of the grain was taken for assay. The two kernel rows adjacent to the inoculated row on pinbar-inoculated ears were removed by hand, and those kernels were used for determination of percent kernel infection.

A total of 390 undamaged kernels from each plot in each experiment was dipped in 70% ethanol, washed in 1.25% NaOCl for 3 min, and rinsed in sterile distilled water. Kernels were placed on Czapek solution agar amended with 7.5% NaCl (CSA-S) in 100-ml petri dishes. The 390 kernels from each plot were placed in 30 petri dishes, with 13 kernels per dish. Plates were incubated for 7 days at 28°C and then examined for fungal growth to determine the percent kernels infected by *A. flavus*.

The 390 kernels per plot were randomly divided into subsamples containing 65, 130, and 195 kernels. Petri plates containing 13 kernels per plate were randomly arranged in the incubator, and kernels from five plates (1–5, 6–10, 11–15, etc.) were combined to form six samples of 65 kernels each. To obtain three 130-kernel samples, kernels from plates 1–10 were combined into sample 1, plates 11–20 into sample 2, and plates 21–30 into sample 3. Finally, two samples containing 195 kernels each were obtained by combining plates 1–15 for sample 1 and plates 16–30 for sample 2.

Table 1. List of experiments (actual and simulated) with number of entries and replicates with mean and range of kernel infection by *Aspergillus flavus* used in this study

Exp[a]	Entries	Reps	Kernel infection	
			Mean (%)	Range (%)
1	40	6	10.6	7 to 21
2	30	6	8.2	4 to 16
3	25	5	7.5	3 to 14
4	32	4	7.9	2 to 13
5	12	6	4.0	1 to 9
6	12	6	3.4	1 to 7
7	14	6	4.0	1 to 8
8	30	5	6.9	1 to 17
Sim 1	14	6	8.4	2 to 16
Sim 2	24	6	14.3	3 to 25

[a] Sim = simulated.

Corresponding author: Gene E. Scott
E-mail: ges3@ra.msstate.edu

Approved for publication as Journal Article J-8420 of the Mississippi Agricultural and Forestry Experiment Station, Mississippi State University.

Accepted for publication 18 April 1995.

In addition to actual data, data for two "experiments" were simulated. That is, for each "plot," the number of "infected kernels" was determined in each of 30 dishes of 13 kernels based on a given expected percentage of infection for that entry. Assuming a given entry mean for percent kernel infection, each kernel was classified as infected or noninfected based on integers from 0 to 99 drawn at random. If the random number was 0, 1, 2, 3, 4, or 5, the kernel was designated as infected, and if the number was six or greater, the kernel was designated noninfected, when the assumed entry mean was 6% kernel infection. For example, one simulation for a plot with an assumed entry mean of 6% infection had petri plates 1 to 30 with 1, 1, 1, 1, 0, 1, 0, 2, 0, 0, 2, 0, 1, 2, 0, 1, 0, 2, 2, 0, 1, 0, 1, 2, 1, 0, 0, 2, and 1 infected kernels. This gave percent kernel infection levels of 7.7, 4.6, 7.7, 7.7, 6.2, and 6.2 for the six 65-kernel samples; 6.2, 7.7, and 6.2 for the three 130-kernel samples; 6.7 and 6.7 for the two 195-kernel samples; and 6.7 for the 390-kernel plot mean. Of course, this shows that even with simulated data, the plot mean is not necessarily identical to the entry mean.

All data were analyzed using SAS ANOVA, which provided error mean square, F value for entries, coefficient of determination (R^2 value), coefficient of variation (CV), and least significant difference (LSD). Spearman rank correlation coefficients for entry means were calculated by SAS CORR. Calculation of rank correlation coefficients (r) between a given sample size and a 390-kernel sample would represent comparison of a part with a whole; so the mean coefficients in Table 2 are among 65-kernel (15 values), 130-kernel (3 values), and 195-kernel (1 value) lots. For example, the first mean given in Table 2 (0.697) is the average of the 15 r values that were obtained when rankings of the six 65-kernel samples were compared in all possible paired combinations. The probabilities in Table 3 were calculated by the formula $1 - (1 - p)^n$ where p equals the percent kernel infection in the population and n equals the number of kernels in the sample.

Table 2. Mean and range of Spearman correlation coefficients obtained for entry rankings of percent kernels infected by *Aspergillus flavus* among sample sizes of 65, 130, and 195 kernels in eight field-grown and two simulated tests

	65 kernels		130 kernels		195 kernels
Exp[a]	Mean	Range	Mean	Range	Mean
1	0.697	0.611–0.798	0.766	0.726–0.828	0.869
2	0.759	0.670–0.878	0.847	0.791–0.893	0.871
3	0.756	0.677–0.851	0.851	0.782–0.886	0.908
4	0.794	0.679–0.895	0.876	0.838–0.903	0.915
5	0.866	0.761–0.961	0.902	0.846–0.993	0.930
6	0.680	0.197–0.865	0.835	0.769–0.881	0.874
7	0.753	0.571–0.896	0.835	0.779–0.893	0.943
8	0.929	0.885–0.958	0.956	0.950–0.968	0.983
Average	0.779		0.859		0.912
Sim 1	0.906	0.815–0.965	0.952	0.934–0.982	0.965
Sim 2	0.951	0.940–0.966	0.974	0.972–0.974	0.981

[a] Sim = simulated experiments.

Table 3. Probability[a] that at least some infection will be detected when different numbers of kernels are evaluated at kernel infection levels from 1 to 10%

Kernels	Actual kernel infection (%)							
per sample	1	2	3	4	5	6	8	10
20	18	33	46	56	64	71	81	88
40	33	55	70	80	87	92	96	99
60	45	70	84	91	95	98	99	100
80	55	80	91	96	98	99	100	100
100	63	87	95	98	99	100	100	100
120	70	91	97	99	100	100	100	100
140	76	94	99	100	100	100	100	100
160	80	96	99	100	100	100	100	100
180	84	97	100	100	100	100	100	100
200	87	98	100	100	100	100	100	100

[a] Probability = $1 - (1 - p)^n$ when p is the actual percent kernel infection of the population and n is the number of kernels per sample.

RESULTS AND DISCUSSION

All statistical parameters became closer to the values for 390 kernels as sample size increased in simulated experiments (Tables 4 and 5). That is, error mean square, CV, and LSD values decreased as sample size increased; and F values and R^2 values increased as sample size increased. The greatest changes over sample sizes were for error mean squares and F values.

Table 4. Statistical attributes obtained from the analyses of data from the first simulated test with different sample sizes of kernels infected by *Aspergillus flavus*

No. of kernels[a]	Error mean square	F value	R^2	CV	LSD
65-1	11.48	11.05	0.6957	40.12	3.91
65-2	12.81	8.17	0.6282	45.14	4.13
65-3	10.83	11.02	0.7057	37.80	3.79
65-4	12.62	8.40	0.6319	45.81	4.10
65-5	12.55	12.05	0.7125	41.15	4.08
65-6	8.17	13.26	0.7300	33.82	3.30
Mean	11.41	10.66	0.6840	40.64	3.89
130-1	6.66	16.14	0.7672	31.53	2.98
130-2	4.72	22.80	0.8254	25.28	2.51
130-3	4.57	27.12	0.8451	25.10	2.47
Mean	5.32	22.02	0.8126	27.30	2.65
195-1	4.41	24.63	0.8328	25.13	2.42
195-2	2.65	43.08	0.8963	19.10	1.88
Mean	3.53	33.86	0.8646	22.12	2.15
390-1	1.77	62.09	0.9257	15.74	1.53

[a] Number of kernels and designation of subsample.

Table 5. Statistical attributes obtained from the analyses of data from the second simulated test with different sample sizes of kernels infected by *Aspergillus flavus*

No. of kernels[a]	Error mean square	F value	R^2	CV	LSD
65-1	17.00	16.85	0.7734	28.90	4.72
65-2	19.67	18.83	0.7914	29.42	5.07
65-3	19.44	14.36	0.7475	31.48	5.04
65-4	18.59	16.76	0.7716	30.92	4.93
65-5	21.57	14.73	0.7491	31.47	5.31
65-6	18.74	15.19	0.7562	30.93	4.95
Mean	19.17	16.12	0.7649	30.52	5.00
130-1	9.02	35.80	0.8783	20.47	3.43
130-2	11.26	25.43	0.8365	24.01	3.84
130-3	9.18	32.19	0.8672	21.09	3.47
Mean	9.82	31.14	0.8607	21.86	3.58
195-1	5.48	55.46	0.9178	16.15	2.67
195-2	6.29	46.54	0.9037	17.62	2.87
Mean	5.89	51.00	0.9108	16.89	1.93
390-1	2.85	103.23	0.9539	11.77	1.53

[a] Number of kernels and designation of subsample.

Therefore, error mean squares and *F* values were the most sensitive to changes in sample size and thus should be the most effective parameters to use to make comparisons among sample sizes.

Based on results from these simulated data, a 65-kernel sample size was insufficient to obtain a desirable level of mean separation. That is, error mean squares were high and *F* values were low compared to those with larger sample sizes (Tables 4 and 5). The differences among error mean squares and *F* values narrowed as sample sizes increased, with the smallest difference between 195 and 390 kernels. This suggests that a sample size of approximately 200 kernels should be considered a minimum sample size.

The error mean squares for simulated data reflect the variability associated with data having a constant probability for infection for each kernel within a plot, and the true means of all replicates of a given entry are identical. However, for actual data from experiments, the error mean squares reflect differences of sample mean from the actual plot mean as well as differences of plot means from entry means. Therefore, it is not surprising that the error mean squares for simulated data are much lower than those for actual data (Table 6).

For actual data, the average error mean square for a sample size of 65 kernels was about 1.8 times greater than for samples with 390 kernels, compared to about 6.5 times greater for simulated data (Table 6). This illustrates the relative lack of precision in tests to determine percent kernel infection by *A. flavus* under field conditions. Thus, there is a greater need for larger sample sizes to more adequately estimate and separate entry means.

The error mean squares with a sample size of 390 kernels in the experiments with actual data were as large as or larger than mean squares for simulated data with 130 kernels. Thus, even the 390-kernel sample was probably not optimum. With somewhat larger samples from field plots, the precision obtained with simulated data would probably not be attained.

The average *F* value with a sample size of 390 kernels over all field experiments was 5.3, compared to 3.5 with a sample size of 65 kernels (Table 7). Although this is not a large numerical difference, relatively small differences in *F* values can markedly change our confidence in mean separation.

An adequate sample size provides data with low variability and good repeatability of ranking entries when tests are repeated. Spearman Rank Correlation coefficients (*r*) are used to measure similarity of rankings. The average correlation coefficients for actual experiments increased with sample size, indicating that a sample size of 65 kernels was poorer for consistency of ranking means than samples of 130 and 195 kernels. In fact, except for experiment 8, the coefficients among rankings of some 65-kernel samples were not significantly different from zero.

Tables are available to indicate when rank correlation coefficients differ significantly from zero, but that doesn't provide a mental picture with respect to what consistency of rankings is associated with a

Table 6. Error mean squares obtained for percent kernel infection by *Aspergillus flavus* with different kernel sample size in eight actual and two simulated experiments

No. of kernels[a]	Experiment designation									
	1	2	3	4	5	6	7	8	Sim 1	Sim 2
65-1	21.71	17.50	22.30	17.74	10.08	14.63	16.46	15.62	11.48	17.00
65-2	18.71	18.96	18.62	14.31	10.96	9.14	17.00	18.03	12.81	19.67
65-3	16.87	15.49	19.07	11.49	32.23	15.25	15.62	16.16	10.83	19.44
65-4	19.03	12.08	16.28	12.43	18.40	13.85	12.97	16.00	12.62	18.59
65-5	18.49	12.79	22.80	10.60	16.51	10.79	9.50	13.69	12.55	21.57
65-6	15.95	14.92	18.58	10.55	14.12	14.10	11.32	12.52	8.17	18.74
Mean	18.46	15.29	19.61	12.85	17.05	12.96	13.81	15.34	11.41	19.17
130-1	15.41	13.86	15.51	12.86	7.04	9.17	11.89	12.89	6.66	9.02
130-2	13.35	10.04	13.52	8.63	15.61	11.25	9.52	13.42	4.72	11.26
130-3	13.77	9.33	17.03	7.88	12.12	9.61	7.24	10.78	4.57	9.18
Mean	14.18	11.08	15.35	9.79	11.59	10.01	9.55	12.36	5.32	9.82
195-1	11.74	11.47	13.44	10.33	9.58	9.01	9.52	11.09	4.41	5.48
195-2	12.02	7.95	12.23	7.18	9.97	9.48	6.91	10.12	2.65	2.85
Mean	11.88	9.71	12.84	8.76	9.78	9.25	8.22	10.61	3.53	4.17
390-1	9.89	7.94	10.64	7.63	8.04	7.64	6.36	9.67	1.77	2.85

[a] Number of kernels and designation of subsample.

Table 7. *F* values obtained for percent kernel infection by *Aspergillus flavus* with different kernel sample sizes in eight actual and two simulated experiments

No. of kernels[a]	Experiment designation									
	1	2	3	4	5	6	7	8	Sim 1	Sim 2
65-1	3.52	3.65	2.14	2.63	6.01	2.15	1.99	8.19	11.05	16.85
65-2	2.74	2.52	2.36	2.92	2.99	2.37	1.89	6.07	8.17	18.83
65-3	3.29	3.58	2.11	3.27	2.15	3.04	2.62	7.29	11.02	14.36
65-4	3.35	4.30	3.17	3.29	2.38	2.49	2.57	6.77	8.40	16.76
65-5	2.43	4.43	1.78	3.81	3.82	2.16	3.23	8.26	12.05	14.73
65-6	4.04	2.96	2.70	2.78	2.73	2.47	3.50	6.98	13.26	15.19
Mean	3.23	3.57	2.38	3.12	3.35	2.45	2.63	7.26	10.66	16.12
130-1	3.70	3.66	2.59	3.09	6.29	2.37	2.31	8.98	16.14	35.80
130-2	4.17	5.09	3.16	4.27	3.25	3.28	3.40	8.22	22.80	25.43
130-3	3.64	4.53	2.44	4.05	3.95	2.68	4.41	9.02	27.12	32.19
Mean	3.84	4.43	2.73	3.80	4.50	2.78	3.37	8.74	22.02	31.14
195-1	4.49	4.30	2.80	3.67	5.29	2.89	2.84	10.24	24.63	55.46
195-2	4.15	5.64	3.38	4.50	4.27	2.93	4.36	9.60	43.08	46.54
Mean	4.32	4.97	3.09	4.09	4.78	2.91	3.60	9.92	33.86	51.00
390-1	4.96	5.80	3.60	4.48	5.61	3.16	4.30	10.76	62.09	103.23

[a] Number of kernels and designation of subsample.

given r value. A numerical example might be of benefit in forming this mental picture. Assume an experiment with 30 entries, and that these entries ranked from best to poorest in consecutive order from 1 to 30 in the first test. Further assume that when the experiment is repeated, the entries could differ in rank in any fashion except that the top, middle, and bottom 10 entries remained in their respective groups in both tests. If within each of the three groups in the second test, the 10 entries reversed (i.e., 10 best to 1 poorest), from the first test an r value of 0.780 would be obtained. This r value is equivalent to the average r value obtained for samples of 65 kernels (Table 2). Performance of entries from best to poorest of 1, 9, 2, 10, 3, 7, 4, 8, 5, and 6 within each group in the second test gives an r value of 0.920. This is equivalent to the average r value with 195 kernel samples (Table 2). Neither of these values reflects a particularly close agreement of rankings between the first and second tests, but certainly consistence of rankings is much better with 195-kernel samples than with 65-kernel samples.

Data on percentage of kernel infection by *A. flavus* show a high level of variability. Certainly, a sample size of 390 kernels is not an excessively large sample because error mean squares, F values, and correlation coefficients do not approach those obtained with simulated data. However, the differences for given statistical attributes among sample sizes of 65 to 390 kernels are progressively smaller. We conclude that the minimum sample size to evaluate for incidence of kernel infection by *A. flavus* would be around 200 kernels.

Samples of maize kernels are not only taken to determine differences among treatment or genotypic means in randomized trials. For instance, one might be interested only in whether or not a seedborne fungus is present. The probability of detection of the fungus with different sample sizes and infection levels can be calculated (Table 3). With a 20-kernel sample size and 1% infection level, there is an 18% probability of having at least one infected kernel in the sample. Even with a 200-kernel sample, there would be only an 87% probability of having at least one infected kernel in the sample. An 88% probability exists for having at least one infected kernel in a 20-kernel sample when the actual infection level is 10%. The data in Table 3 illustrate that a pathogen like *A. flavus*, which with natural inoculation often occurs at a low incidence, can go undetected if the sample size is not relatively large. It would appear that a sample size of 100 kernels would be required to have a reasonable probability of detecting the fungus when it is present, and in areas where the level of infection is expected to be very low, a larger sample would be required.

LITERATURE CITED

1. King, S. B., and Scott, G. E. 1982. Field inoculation techniques to evaluate maize for reaction to kernel infection by *Aspergillus flavus*. Phytopathology 72:782-785.
2. Payne, G. A., Hagler, W. M., Jr., and Adkins, C. R. 1988. Aflatoxin accumulation in inoculated ears of field-grown maize. Plant Dis. 72:422-424.
3. Rambo, G. W., Tuite, J., and Crane, P. 1973. Preharvest inoculation and infection of dent corn ears with *Aspergillus flavus* and *A. parasiticus*. Phytopathology 64:797-800.
4. Reddy, M. J., Banerjee, A., and Shetty, H. S. 1992. Seed resistance to *Aspergillus flavus* colonization in different cultivars of maize and sunflower and aflatoxin B_1 production in autoclaved kernels. Trop. Sci. 33:63-68.
5. Tucker, D. H., Jr., Trevathan, L. E., King, S. B., and Scott, G. E. 1986. Effect of four inoculation techniques on infection and aflatoxin concentration of resistant and susceptible corn hybrids inoculated with *Aspergillus flavus*. Phytopathology 76:290-293.
6. Zummo, N., and Scott, G. E. 1990. Cob and kernel infection by *Aspergillus flavus* and *Fusarium moniliforme* in inoculated, field-grown maize ears. Plant Dis. 74:627-631.