

# Number of Experiments Needed to Determine Wheat Disease Phenotypes for Four Wheat Diseases

William W. Bockus, Department of Plant Pathology, Throckmorton Hall, Kansas State University, Manhattan 66506; Zhaohui Su, Statistical and Data Analysis Center, Harvard School of Public Health, Boston, MA 02115; Karen A. Garrett, Bikram S. Gill, and James P. Stack, Department of Plant Pathology, and Allan K. Fritz and Kraig L. Roozeboom, Department of Agronomy, Kansas State University; and T. Joe Martin, Kansas State University Agricultural Research Center-Hays, Hays 67601

## ABSTRACT

Bockus, W. W., Su, Z., Garrett, K. A., Gill, B. S., Stack, J. P., Fritz, A. K., Roozeboom, K. L., and Martin, T. J. 2007. Number of experiments needed to determine wheat disease phenotypes for four wheat diseases. *Plant Dis.* 91:103-108.

Disease phenotypes for winter wheat cultivars were determined in numerous inoculated greenhouse and field experiments over many years. For four diseases, *Fusarium* head blight, tan spot, *Septoria* leaf blotch, and *Stagonospora* leaf blotch, at least 20 cultivars each had been evaluated in a minimum of five experiments. Reference cultivars of known disease reaction were included in each experiment, which allowed transformation of the percent disease severity data to a 1-to-9 scale for comparisons between experiments. Variations in scale values obtained for individual cultivars among the different experiments were used to calculate standard deviations for disease phenotype data. Standard deviations were used to calculate the number of experiment repetitions needed within each disease to achieve different levels of accuracy (margins of error). A margin of error of  $\pm 1.5$  for the 1-to-9 scale was chosen as the best level of accuracy. Rounding values within this range would put the estimated disease phenotype within  $\pm 1$  unit of the actual phenotype. To achieve a margin of error of  $\pm 1.5$  for *Fusarium* head blight, tan spot, *Septoria* leaf blotch, and *Stagonospora* leaf blotch would require a mean that was calculated from a minimum of five, five, seven, and eight experiments, respectively. Personnel who report disease phenotype data to wheat producers or breeders should be aware of the number of experiments upon which they are basing their reports and adjust any disclaimers accordingly. Similarly, wheat breeders should be aware of the inherent variability in phenotyping these four wheat diseases and make appropriate adjustments to their selection protocols. With a minimum of five experimental repetitions, disease phenotype values obtained from inoculated greenhouse and field experiments had very high correlations ( $r = 0.81$  to  $0.92$ ,  $P < 0.0001$ ) with published Kansas State University Research and Extension ratings obtained from commercial fields.

Additional keywords: *Fusarium graminearum*, *Septoria tritici*

Host resistance to disease is often the best management procedure for limiting losses from plant diseases. Therefore, producers need to know the reaction of various cultivars to pathogens to decide which cultivars to plant. To help producers make those decisions, crop consultants and extension personnel must be able to collect and disseminate this type of information. Similarly, plant breeders need access to disease information to select which lines to

discard and which to advance. Finally, researchers also routinely need to know the disease phenotype of certain cultivars as they study various aspects of disease resistance. Determining the reaction of plant cultivars and breeding lines to pathogens (disease phenotyping) is one of the most important technical procedures in the science of plant pathology.

Wheat-breeding programs in Kansas recently have increased the number of diseases that are included in the selection process of new cultivars. Prior to 1977, wheat cultivars were resistant to only one or two diseases when they were released. However, by the late 1980s, cultivars had some level of resistance to at least eight diseases (4). Such a commitment to breeding for disease resistance has required increased effort directed toward disease phenotyping.

There are 20 diseases that routinely are observed in wheat plants in Kansas: leaf rust, wheat streak mosaic, stripe rust, barley yellow dwarf, tan spot, *Septoria* leaf blotch, *Stagonospora* leaf blotch, soilborne

mosaic, spindle streak mosaic, take-all, stem rust, powdery mildew, *Fusarium* head blight, common root and crown rot, strawbreaker foot rot, bacterial leaf blight, common bunt, loose smut, *Cephalosporium* stripe, and American wheat striate (1). Of these, three (strawbreaker foot rot, bacterial leaf blight, and American wheat striate mosaic) are of such minor importance in the state that no effort has been made to develop cultivars with improved levels of resistance. Two more (common bunt and loose smut) are controlled effectively by inexpensive chemical seed treatments, so that no recent effort has been made to breed for resistance to them. Two more diseases (powdery mildew and common root and crown rot) cause sufficient losses to warrant efforts to develop resistant cultivars and probably will be future emphases in the breeding programs. The remaining 13 diseases all have been the object of disease resistance breeding. This requires substantial effort to phenotype breeding lines for reaction to these diseases.

Cultivars with improved levels of resistance have been adopted readily by Kansas wheat producers, such that losses to four important diseases have been significantly reduced in recent years (4). These diseases are soilborne mosaic, spindle streak mosaic, *Cephalosporium* stripe, and tan spot. An important component of successful adoption of cultivars with improved disease resistance has been the dissemination of information about cultivars to wheat producers. Kansas State University (KSU) Research and Extension has two main publications that are used to disseminate this information, "Kansas Performance Tests with Winter Wheat Varieties" and "Wheat Variety Disease and Insect Ratings" (23,32). Both publications are available as hard copy or on the Internet. Production of such publications requires access to accurate disease phenotype data and constant updating of those data.

To produce resistant wheat cultivars and gather data on the reaction of cultivars to various diseases, an accurate way to determine disease phenotypes is needed. Three main methods have been used in the KSU wheat-breeding program to determine disease phenotypes: (i) breeding nurseries where diseases naturally occur,

Corresponding author: W. W. Bockus  
E-mail: bockus@ksu.edu

Contribution number 06-335-J from Kansas State Research and Extension.

Any opinions, findings, conclusions, or recommendations expressed in this publication are those of the authors and do not necessarily reflect the view of the United States Department of Agriculture.

Accepted for publication 8 September 2006.

DOI: 10.1094/PD-91-0103  
© 2007 The American Phytopathological Society

(ii) inoculated disease-testing nurseries in the field, and (iii) inoculated experiments in the greenhouse. To produce a new wheat cultivar in Kansas involves at least 40 location-years; therefore, breeders usually are able to evaluate their material several times for the more commonly occurring diseases such as leaf rust. Field disease-testing nurseries are established routinely for soil-borne mosaic, spindle streak mosaic, wheat streak mosaic, tan spot, barley yellow dwarf, and Fusarium head blight (5–7,13,14). Less often, nurseries are established for Stagonospora leaf blotch (3) and Septoria leaf blotch. Finally, greenhouse assays have been developed, or adapted from published methods, to test wheat lines for reaction to diseases. These include tan spot (30), Stagonospora leaf blotch (19,33), Septoria leaf blotch (37), Fusarium head blight (2), soilborne mosaic (8), wheat streak mosaic (27), and take-all (25).

When gathering disease phenotype data from these various experiments, there is variability in reaction for a given wheat cultivar or line from experiment to experiment. Although some of this variability may be due to new races of a pathogen, such as the leaf rust fungus, there is also variability among experiments due to the natural variation that occurs in biological systems. Because of this natural variability among experiments, an important question when evaluating a wheat cultivar or breeding line is: how many experiments are needed to achieve a certain level of accuracy?

Large amounts of data on disease phenotypes have been collected in Kansas during the past 25 years. Although the popularity of wheat cultivars changes from year to year, some cultivars have been evaluated many times over multiple years. These data sets make it possible to calculate the inherent variability in disease phenotyping experiments and, thus, the number of times a cultivar must be evaluated to reach a desired level of accuracy.

For four of the diseases (tan spot, Stagonospora leaf blotch, Septoria leaf blotch, and Fusarium head blight), at least 5 evaluations are available for at least 20 cultivars (average of 9.3 evaluations per cultivar per disease). These data for the four diseases were selected for analysis with the objective of determining the number of experiments necessary to achieve various levels of phenotyping accuracy. Accomplishing this objective was important to construct recommendations for the number of trials needed for successful and efficient comparison of lines in breeding programs and by wheat producers. A second objective was to correlate results from long-term greenhouse disease phenotyping experiments with published KSU Extension values obtained from field observations by Extension personnel. This objective was

important to determine the accuracy of the phenotyping experiments relative to the other ratings.

## MATERIALS AND METHODS

**Experimental design for leaf spot diseases.** Similar systems were used in the greenhouse for determining the reaction of wheat cultivars or lines to tan spot, Stagonospora leaf blotch, and Septoria leaf blotch. They were adapted from the procedures reported by Raymond et al. (30), Eyal and Scharen (19), and Yechilevich et al. (37). Briefly, they involved a randomized complete block design with 10 to 20 entries, four replications, and 10 plants per replication. Seedlings were grown in 2.5-by-13-cm cones (Stuewe and Sons, Corvallis, OR) in a soil:vermiculite mix (50:50) to the four-leaf stage (about 4 weeks). A spore suspension of one of the fungi was produced in the laboratory and applied to the leaves. Fungal isolates used in these experiments were obtained from infected wheat in commercial fields and represented the most virulent naturally occurring strains known to occur in Kansas. Inoculated plants were placed into a chamber where continual leaf wetness was maintained for 48, 72, or 96 h to produce tan spot, Stagonospora leaf blotch, or Septoria leaf blotch, respectively. After 7, 14, or 19 days, respectively, the bottom three leaves were assessed visually for percent leaf area affected by disease. Therefore, for a single experiment, the disease severity value for each cultivar is the mean of four replications of about 30 leaves each. Within an experiment, the data were subjected to analyses of variance to separate cultivar means, but such analyses are not reported here. Only the mean values of cultivars were used in the transformation described below.

**Experimental design for Fusarium head blight.** A modified protocol of Wang et al. (35) was used. Field experiments were established using a randomized complete block with 25 to 60 entries and four replications, and plots were single rows 2.3 m long and 50 cm apart. Seed was sown in the fall (about 1 October) at the standard rate using a cone-type planter. Air-dried corn kernels colonized by a single, aggressive isolate of *Fusarium graminearum* were spread throughout the test area on 1 April, 15 April, and 1 May (85 g/m<sup>2</sup> total). During anthesis, heads were kept wet by using small, overhead, impulse sprinklers for 3 min every hour from 9:00 p.m. until 6:00 a.m. Visual estimations of the percentage of spikelets affected by Fusarium head blight (FHB index; 34) for an entire plot were taken several times and averaged. Within an experiment, the data were subjected to analyses of variance to separate cultivar means, but such analyses are not reported here. Only the mean values of cultivars were used in the transformation described below.

**Comparison of data among experiments.** Because intensity of a given disease varied over experiments, data had to be transformed in order to compare results among many experiments. In Kansas, disease phenotypes for wheat cultivars are reported using a linear 1-to-9 scale, where 1 is highly resistant and 9 is highly susceptible. This scale is used in KSU Research and Extension publications (23,32) that disseminate disease phenotype information to wheat producers. To transform the data from disease severity percentages to this scale, four “reference” cultivars of known reaction were included in each experiment. The known reactions were averages from at least 20 experiments. Because the reaction of these cultivars was accurately known, a linear model could be produced that allowed data from all cultivars within an experiment to be transformed to 1-to-9-scale values. Using tan spot as an example, the reference cultivars were Karl 92 (known scale value = 3.16), Triumph 64 (4.98), TAM 107 (5.59), and TAM 105 (8.80). Therefore, given the data obtained in a single experiment, such as those shown in Table 1, and regressing the known scale reaction of the four reference cultivars against the percentage of leaf area affected in that particular experiment produced the equation  $Y = 8.6X + 1.05$ , where  $X$  is the scale value calculated from the percent leaf area affected ( $Y$ ). Once the equation was obtained from data of the reference cultivars, scale values for all 10 (or 20) cultivars in an experiment could be calculated, including the six (or 16) cultivars of unknown reaction (Table 1). Therefore, data from experiments over many years could be compared because they all were transformed to the same 1-to-9 scale. However, experiments were included only where moderate to severe disease occurred on the susceptible cultivars.

**Statistical analyses.** In the analyses, cultivars were included only if they had five or more observations (experiments) associated with a disease, although some cultivars had been evaluated over 20 times for reaction to a particular disease. This resulted in a total of 25, 25, 23, and 20 cultivars for Fusarium head blight, tan spot, Septoria leaf blotch, and Stagonospora leaf blotch, respectively. The average number of times an included cultivar had been evaluated was 8.0, 10.4, 8.4, and 10.5, respectively.

Tests of normality for the distribution of ratings for each cultivar to each disease were performed using the PROC UNIVARIATE procedure (SAS 8.2 software; SAS Institute, Cary, NC). The Shapiro-Wilk, Kolmogorov-Smirnov, Cramer-von Mises, and Anderson-Darling tests were used. For simplicity, only  $P$  values from the Shapiro-Wilk test are discussed in this article. The number of experiment repetitions needed was calculated based on the formula

$$t_{\alpha/2, n-1} \frac{s}{\sqrt{n}} = MOE \quad (\text{eq. 1})$$

where  $t_{\alpha/2, n-1}$  is the  $(1 - \alpha/2)$  quantile of the  $t$  distribution with  $n - 1$  degrees of freedom,  $n$  is the number of repetitions required,  $s$  is the standard deviation of the obtained ratings, and  $MOE$  (margin of error) is the prespecified deviation of the obtained rating from the true rating. For example, if a researcher wants the averaged rating to be within  $\pm 1.5$  units on the 1-to-9 scale, then  $MOE = 1.5$ . With a  $MOE$  of 1.5 or less, a value would be rounded to  $\pm 1$  unit of the correct value. For a specified value of  $MOE$ , the number of repetitions needed is determined by computing  $n$  iteratively based on equation 1 at a significance level  $\alpha$ . In this article, the significance level  $\alpha$  is 0.05 unless specified otherwise.

To obtain a plot summarizing the relationship between  $MOE$  and  $n$  for each disease when averaging all the cultivars within that disease, we took the mean (and the median) of  $n$  of all the cultivars within each disease at each specified  $MOE$ . In our study, the  $MOE$  was set at 0.5, 1, 1.5, 2, and 2.5 units.

A test of the null hypothesis of equal variance (or equivalently, standard deviation) of the estimated ratings for the four diseases was performed using the MIXED procedure in SAS 8.2. If the variances of the estimated rating for any two diseases are significantly different, then a significantly different number of repetitions of experiments are needed to reach the same specified  $MOE$ . The MIXED procedure was used to test whether the means of the standard deviations of the estimated ratings for all cultivars within each disease were equal for the four diseases. Here, the treatment is disease and the response variable is the standard deviation of ratings for each cultivar within each disease. We chose to model standard deviations rather than variances because the set of standard deviation estimates from all cultivars within each disease was more approximately normally distributed. Cultivar was modeled as the random effect.

A  $t$  test was used to determine whether there is a difference in estimating a cultivar's disease phenotype based on the inoculated experiments compared with that published by KSU Research and Extension (23,32). This test was done for each cultivar within each disease. The analysis was based on the data for Fusarium head blight, tan spot, and Septoria (note: a KSU Research and Extension rating for Stagonospora leaf blotch is not available). Standard linear regression was used to test the correlation between ratings obtained from inoculated greenhouse or field phenotyping experiments with those published by KSU Research and Extension for these three diseases.

## RESULTS

Numerous disease phenotyping experiments were conducted over a period of 7, 15, 13, and 11 years for Fusarium head blight, tan spot, Septoria leaf blotch, and Stagonospora leaf blotch, respectively (9,15,17). Several different people rated these experiments, including undergraduate students, graduate students, research assistants, and research faculty. All raters had been trained by the first author or his long-time research assistant and some raters were checked for accuracy by having an experienced rater re-rate an experiment. Some small discrepancies between raters were noted; however, this involved one rater having consistently lower ratings than another rater. No major discrepancies were noted between raters with regards to the ranking of cultivars.

When percent diseased leaf area data were transformed to the 1-to-9 scale (Table 1), there was variation in the values obtained for a cultivar among experiments. For example, the average value for the reaction of the cv. 2137 to tan spot from seven different experiments was 3.89; however, the values for individual experiments varied from 2.82 to 5.04. Variation of this type was noted for all cultivars and was used to calculate the standard deviations for the reactions of cultivars to diseases.

For testing the normality assumption for the estimated ratings of each cultivar within each disease, the SAS PROC UNIVARIATE procedure was used and the results show that only 1 of the 93 disease-cultivar combinations rejected the normality assumption at the 0.05 significance level. That combination was cv. Jagger with Stagonospora leaf blotch ( $P$  value = 0.012,  $n = 10$  experiments). This indicates that the normality assumption is reasonable for this study considering adjustment for multiple tests, for example, by the Bonferroni method.

For the four wheat diseases discussed in this article, the standard deviations of dis-

ease phenotype ratings for all cultivars were plotted against their averaged 1-to-9-scale ratings (Fig. 1). When linear and quadratic models were fit to these data, there was a significant quadratic effect for Fusarium head blight (Fig. 1A). Models for all other diseases were not significant (Fig. 1B, C, and D).

For each of the four diseases, the number of experiment repetitions needed at a specified margin of error was computed using equation 1 and is shown in Figure 2. A mixed model was used to test whether the four curves were significantly different from each other and the results are presented in the figure caption. The pair of curves for Fusarium head blight and tan spot are not significantly different from each other ( $P = 0.847$ ), and the pair of curves for Septoria leaf blotch and Stagonospora leaf blotch are not significantly different ( $P = 0.318$ ). The  $P$  values are 0.037, 0.023, 0.003, and 0.002 for comparing Septoria leaf blotch with tan spot, Septoria leaf blotch with Fusarium head blight, Stagonospora leaf blotch with tan spot, and Stagonospora leaf blotch with Fusarium head blight, respectively. After adjustment for multiple comparisons by the Bonferroni method, only comparisons with values lower than 0.05 divided by 6 = 0.008 are statistically different. Similar results were obtained when median values were used instead of the means (*data not shown*). The only difference was that the curves were slightly shifted to fewer repetitions ( $n$ ) when the median value was used.

Results from  $t$  tests did not detect any significant differences between the disease phenotype ratings for cultivars obtained from the greenhouse and field experiments and those published by KSU Research and Extension for Fusarium head blight, tan spot, and Septoria leaf blotch (23,32). KSU does not publish extension ratings for Stagonospora leaf blotch; therefore, that comparison could not be calculated. Correlations between average 1-to-9-scale values obtained from greenhouse or field

**Table 1.** Transformation of data from percent leaf area affected by tan spot to a 1-to-9 scale using four reference cultivars within an experiment involving 10 cultivars

Cultivar	Known reaction (1-to-9 scale) <sup>a</sup>	Leaf area affected (%)	Calculated scale value (1 to 9) <sup>b</sup>
Reference cultivars			
Karl 92	3.16	29.67	3.33
Triumph 64	4.98	40.59	4.60
TAM 107	5.59	50.51	5.75
TAM 105	8.80	77.20	8.85
Cultivars of unknown reaction			
Coker 9474	...	28.22	3.16
Prairie Red	...	52.95	6.03
TAM 302	...	61.03	6.97
Akron	...	65.39	7.48
Prowers 99	...	67.35	7.71
Scout 66	...	75.14	8.62

<sup>a</sup> Average values obtained from at least 20 observations (experiments).

<sup>b</sup> Transformation equation derived from data of the reference cultivars is  $Y = 8.6X + 1.05$ , where  $X$  is the calculated scale value and  $Y$  is the percent leaf area affected.

phenotyping experiments with Extension values were highly significant for all three winter wheat diseases. The correlation coefficient for Fusarium head blight was

0.922 ( $n = 25$ ,  $P < 0.0001$ ), for tan spot it was 0.882 ( $n = 25$ ,  $P < 0.0001$ ), and for Septoria leaf blotch it was 0.813 ( $n = 23$ ,  $P < 0.0001$ ).

## DISCUSSION

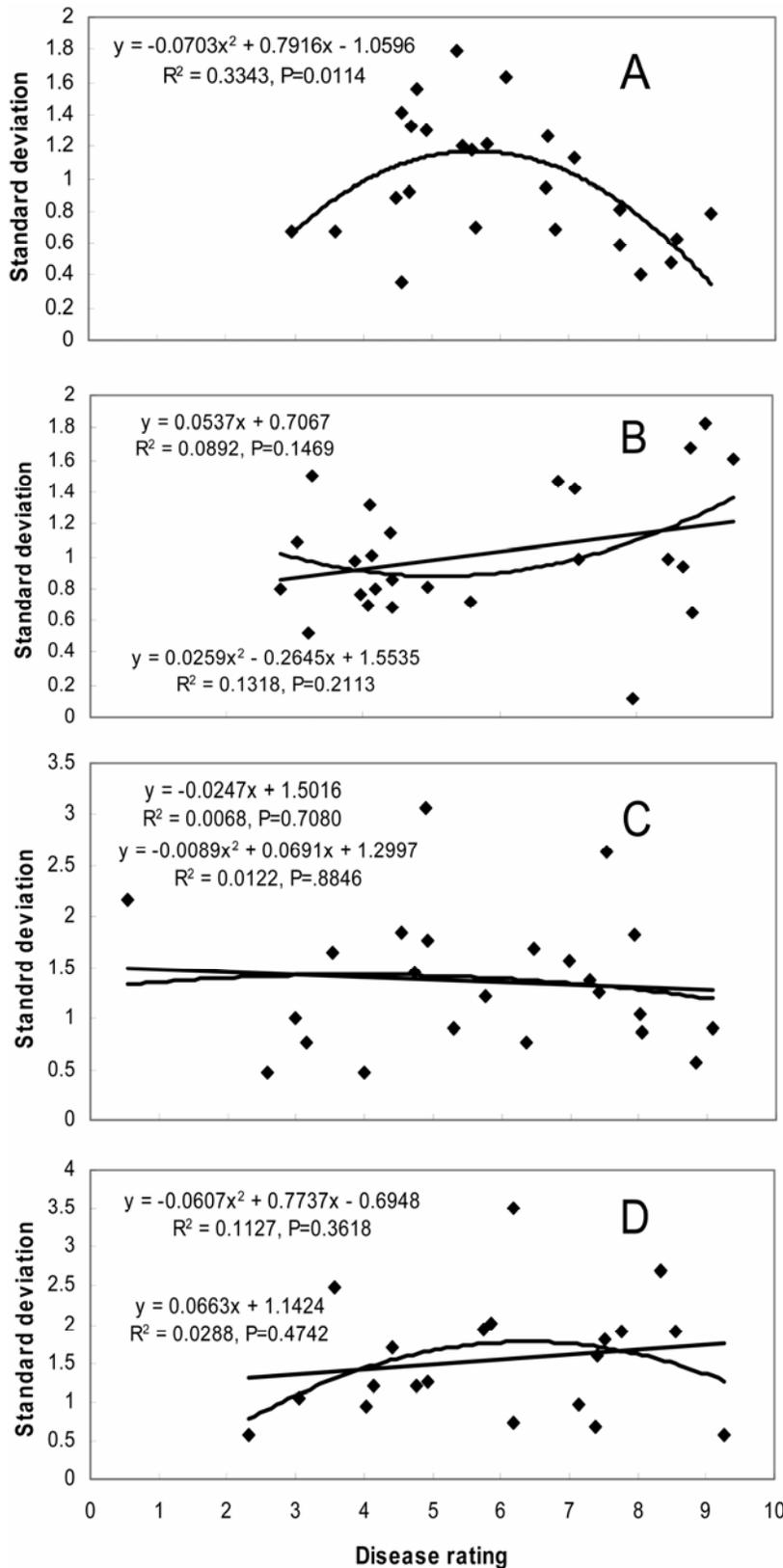
Substantial variation among experiments can occur in the measured phenotype for a cultivar. For example, the standard deviation for the reaction of one cultivar to *Stagonospora* approached 3.5 on a 1-to-9 scale (Fig. 1D). This variation can be due to several factors, including the accuracy of the inoculation procedure, the accuracy of the person doing the visual disease assessment, the virulence of the pathogen, and environmental effects on phenotype.

Experimental protocols used in these experiments represent typical procedures that have been used successfully by many researchers (19,30,33,37). Nevertheless, variations can occur between experiments in such parameters as the inoculum potential of the spores and uniformity of the inoculation of plants. Campbell and Lipps (10) and Fuentes et al. (21) reported some of the types of variation that can occur when screening for Fusarium head blight. All such variations would add to the observed deviations in measuring disease phenotypes of cultivars. Therefore, any modification of procedures that would result in more uniformity among experiments would reduce the interexperiment variation.

Several different people did the disease evaluations used in this research. Although their competence probably varied somewhat, we assume that the accuracy of the phenotype data is similar to what would be achieved by an average evaluator in any location. It has been our experience that any person may be quickly trained to evaluate wheat plants for percent disease severity and that they can produce reasonably accurate data. Furthermore, the phenotype data obtained at KSU for the Northern Uniform Winter Wheat Scab Nursery had the highest correlation with the overall average of all eight rating sites over a 2-year period (18). Therefore, we believe that the recommendations presented here to achieve a certain level of accuracy are good estimates for what other researchers will experience. However, it must be pointed out that raters with above-average accuracy would be expected to lower the number of experimental repetitions needed to determine the phenotype of wheat lines.

For the four diseases used in these experiments, changes in fungal physiological specialization were not responsible for the observed variation among experiments reported here. Although races have been reported for tan spot (11) and *Septoria tritici* (20), all of the isolates used in these experiments represented the most virulent that have been identified to date in Kansas. Nevertheless, if a difference in pathogen virulence does occur among experiments, it obviously would introduce large variations in the disease phenotype for cultivars.

Environmental differences among experiments can have a large impact on the disease phenotype. For example, tempera-



**Fig. 1.** Standard deviation of disease phenotype ratings as a function of the disease rating (1-to-9 scale) for individual winter wheat cultivars for four different diseases. **A**, Fusarium head blight; **B**, tan spot; **C**, Septoria leaf blotch; and **D**, Stagonospora leaf blotch). Linear and quadratic models were fit to the data. Note: the y-axis scales are different among the figures.

ture-sensitive resistance has been reported for *Stagonospora* leaf blotch (16,24). The temperature effect reported by Kim and Bockus (24) first was detected because widely differing results were obtained depending on the time of year that cv. AGSECO 7853 was tested in the greenhouse. Similarly, more subtle differences undoubtedly occur with some cultivars as the environment fluctuates from experiment to experiment, contributing to the observed phenotype variability. If predictions for increased temperature fluctuations in many parts of the world prove to be correct, this may contribute to greater variability in the effectiveness of temperature-sensitive resistance genes in trials and in production systems (12,22).

We originally hypothesized that cultivars at either the resistant or susceptible ends of the scale would have lower standard deviations than cultivars with scale values in the middle. However, in only one instance (*Fusarium* head blight; Fig. 1A) was this pattern observed. With the other three pathogens, there were no significant linear or quadratic relationships detected when standard deviation was regressed against the disease scale value (Fig. 1B to D). Nevertheless, with *Fusarium* head blight, it would take fewer experiments (three versus five, with a margin of error of  $\pm 1.5$ ) to accurately determine the disease phenotype for cultivars that are either very resistant or very susceptible compared with cultivars in the middle of the range.

For obvious reasons, accurate disease phenotype data are important. They allow wheat producers to choose which cultivars to plant and allow breeders to decide which lines to advance or discard in their program. A main question is: what level of accuracy is acceptable? For a 1-to-9 scale, such as the one used here and by KSU Research and Extension, how many experiments must be conducted so that the reported value is identical to the actual value? Such accuracy, using the testing methods reported here, would require a margin of error of  $\pm 0.5$ , so that the obtained average value would be rounded to the correct scale value. Depending upon the disease, that level of accuracy would require between 20 and 47 evaluations of a cultivar (Fig. 2). In only rare instances will a cultivar have been tested that often; therefore, such accuracy usually will not be obtained.

If the desired level of accuracy is such that a reported value can be as much as 2 units from the actual value (margin of error = 2.5), only three to four evaluation experiments are needed (Fig. 2). However, that level of accuracy is not very useful because a cultivar or line with an actual value of 5 could have a reported value of between 3 (in the resistant group) and 7 (in the susceptible group). Although requiring relatively few experiments, this level of accuracy is too low to be of use in most situations.

We believe that, when reporting disease-phenotype data to wheat producers or breeders, an accuracy of  $\pm 1$  unit on the 1-to-9 scale is acceptable. For this to occur, enough experiments must be conducted so that the average of those experiments is  $\pm 1.5$  units, so that, when the average is rounded, it is within the  $\pm 1$ -unit range. For *Fusarium* head blight, tan spot, *Septoria* leaf blotch, and *Stagonospora* leaf blotch, that would require a minimum of 4.4, 4.5, 6.4, and 7.5 experiments, respectively (Fig. 2). Therefore, for an average cultivar, five to eight evaluations of its reaction to a disease must be conducted before the desired level of accuracy is achieved. It is our belief that this number of experiments is more than usually might be used by most scientists. According to our data, lower levels of repetition can result in significant errors between the published resistance level of a cultivar to a disease and its actual level. Therefore, extension personnel should be aware of the number of experiments upon which they are basing their publications and adjust any disclaimers accordingly. Similarly, wheat breeders should be aware of the inherent variability in phenotyping the four diseases discussed here and make any appropriate adjustments to their selection schemes.

The numbers of experiments needed to achieve certain levels of accuracy reported above are for an "average" cultivar or breeding line. However, standard deviations of disease phenotype ratings for individual cultivars ranged from about 3.5 to less than 0.5 on the 1-to-9 scale (Fig. 1). Cultivars with relatively high standard deviations would have less reproducible results and require more experiments. Conversely, cultivars with relatively low standard deviations would have more reproducible results and require fewer experiments. Although possible to calculate,

a confidence interval about the estimate of repetition is not provided. The reason is the statistical difficulty of constructing a useful confidence interval when the number of experiments is small ( $n < 10$ ), such as for some of the cultivar-disease situations in these experiments. Additionally, some cultivars may have a large variance due to an important reason besides random variation between experiments. As an example, the cultivar with the highest standard deviation (3.5) was AGSECO 7853 for *Stagonospora* leaf blotch. As reported above, when the reason for this large variation was investigated, it was shown to be due to resistance that was temperature sensitive (24). If the greenhouse was cooler or warmer than normal, very different results were obtained. Because the standard deviation of a cultivar is not known until it has been tested many times, the numbers of experiments for an "average" cultivar reported above should be used by researchers as preliminary guidelines and appropriate adjustments made as data are gathered.

The phenotype data reported here are for four "necrotrophic" pathogens. Plants were inoculated with spore suspensions, incubated under conditions conducive to disease, and the percent disease severity determined. For some diseases, there may be procedures available where much less variation is encountered. For example, cultivar-specific toxins produced by the causal agent of tan spot have been reported (11). When determining phenotypes to one of these toxins, accurate data may be obtained using as few as two replications of five plants in a single experiment (26). In all likelihood, controlled conditions involving preparations of this toxin would even allow selections on a single-plant basis. Similarly, for "biotrophic" pathogens such as leaf rust, phenotype data are

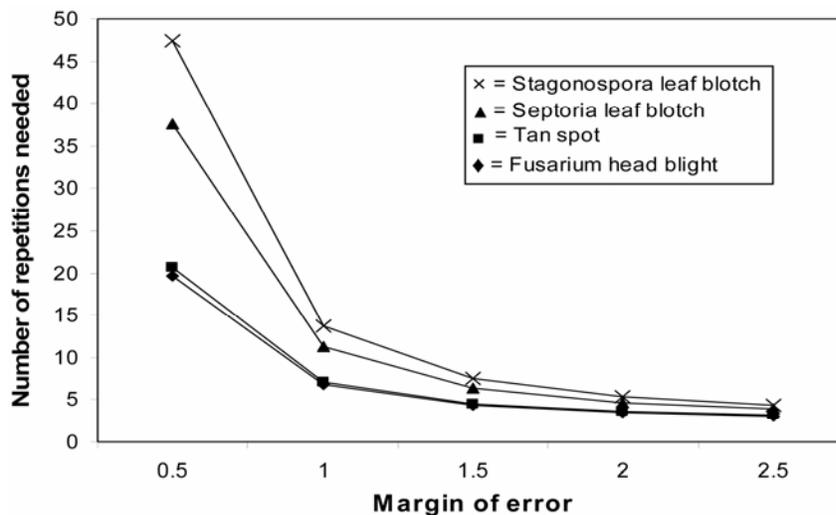


Fig. 2. Number of experiment repetitions needed to achieve a specified margin of error (MOE) for phenotyping four wheat diseases using a 1-to-9 scale. The LSMEAN of standard deviations for the curves, followed by mean separation letters, are 1.54 a for *Stagonospora* leaf blotch, 1.36 a for *Septoria* leaf blotch, 1.01 b for tan spot, and 0.98 b for *Fusarium* head blight.

more reproducible and fewer experiments are needed to obtain accurate values, sometimes involving the reaction from only a single plant (31). However, determining disease phenotypes for adult plant resistance toward biotrophic pathogens may require more plants and multiple experiments.

Disease observations from commercial fields, extension demonstration plots, and breeder's nurseries rely on the natural occurrence of epidemics. Additionally, farmer participatory research can supply useful information about the performance of breeding materials in a wide range of environments (28,29,36). On the other hand, the use of inoculum in greenhouse and field testing procedures has the advantage of producing consistent disease levels, so that good data may be obtained on a regular basis. However, such experiments are useful only if they produce data that reflect the reaction of cultivars when grown under commercial conditions. For three of the wheat diseases reported here, KSU Research and Extension ratings are available (23,32). These latter ratings are mainly from observations collected in producer's fields and extension demonstration plots; therefore, they represent disease phenotypes seen under naturally occurring conditions. Average phenotype values from the inoculated experiments reported here, obtained from a minimum of five experiments, had very high correlations (0.813 to 0.922,  $P < 0.0001$ ) with the KSU Research and Extension values. Therefore, values obtained from these types of experiments are useful because they mimic those observed under commercial conditions. They can be used by producers, county agents, crop consultants, and researchers to predict how a cultivar will perform when certain diseases occur. The assumption is, however, that the disease phenotype values have been calculated from enough experiments. We believe that, for the four wheat diseases discussed here, a minimum of five experiments should be averaged to obtain a reasonable level of accuracy.

#### ACKNOWLEDGMENTS

Funding for this research was from the Kansas Wheat Commission, Kansas Crop Improvement Association, Kansas State Research and Extension, and the United States Department of Agriculture, under agreement numbers 58-5430-2-323 and 5430-21000-005-245. The latter are cooperative projects with the United States Wheat & Barley Scab Initiative. We thank M. A. Davis and B. Jedlicka for technical assistance.

#### LITERATURE CITED

- Appel, J., Stack, J., Bowden, R. L., and Bockus, W. W. 2005. Kansas Cooperative Plant Disease Survey Report; Preliminary 2005 Kansas Wheat Disease Loss Estimates, Online, Kansas Department of Agriculture.
- Bai, G. H., and Shaner, G. 1996. Variation in *Fusarium graminearum* and cultivar resistance to wheat scab. *Plant Dis.* 80:975-979.
- Bockus, W. W. 1996. Reaction of selected winter wheat cultivars to *Nodorum* leaf and glume blotch, 1995. *Biol. Cult. Tests Control Plant Dis.* 11:84.
- Bockus, W. W., Appel, J. A., Bowden, R. L., Fritz, A. K., Gill, B. S., Martin, T. J., Sears, R. G., Seifers, D. L., Brown-Guedira, G. L., and Eversmeyer, M. G. 2001. Success stories: Breeding for wheat disease resistance in Kansas. *Plant Dis.* 85:453-461.
- Bockus, W. W., and Davis, M. A. 2004. Reaction of selected winter wheat cultivars to tan spot and leaf rust, 2003. *Biol. Cult. Tests Control Plant Dis. Rep.* 19:FC005, DOI:10.1094/BC19, The American Phytopathological Society, St. Paul, MN.
- Bockus, W. W., and Davis, M. A. 2005. Reaction of selected winter wheat accessions to *Fusarium* head blight, 2004. *Biol. Cult. Tests Control Plant Dis. Rep.* 20:FC026, DOI:10.1094/BC20, The American Phytopathological Society, St. Paul, MN.
- Bockus, W. W., Fritz, A. K., and Martin, T. J. 2005. Reaction of the 2004 Kansas Intrastate Nursery to *Fusarium* head blight, 2004. *Biol. Cult. Tests Control Plant Dis. Rep.* 20:FC025, DOI:10.1094/BC20, The American Phytopathological Society, St. Paul, MN.
- Bockus, W. W., and Niblett, C. L. 1984. A procedure to identify resistance to wheat soilborne mosaic in wheat seedlings. *Plant Dis.* 68:123-124.
- Brown-Guedira, G. L., Gill, B. S., Bockus, W. W., Cox, T. S., Hatchett, J. H., Leath, S., Peterson, C. J., Thomas, J. B., and Zwer, P. K. 1996. Evaluation of a collection of wild timopheevi wheat for resistance to disease and arthropod pests. *Plant Dis.* 80:928-933.
- Campbell, K. A. G., and Lipps, P. E. 1998. Allocation of resources: sources of variation in *Fusarium* head blight screening nurseries. *Phytopathology* 88:1078-1086.
- Ciuffetti, L. M., Francl, L. J., Ballance, G. M., Bockus, W. W., Lamari, L., Meinhardt, S. W., and Rasmussen, J. B. 1998. Standardization of toxin nomenclature in the *Pyrenophora tritici-repentis*/wheat interaction. *Can. J. Plant Pathol.* 20:421-424.
- Coakley, S. M., Scherm, H., and Chakraborty, S. 1999. Climate change and plant disease management. *Annu. Rev. Phytopathol.* 37:399-426.
- Cox, C., Bockus, W., Garrett, K., Cox, T. S., and Peters, T. 2005. Reaction of selected perennial grass accessions to barley yellow dwarf, 2004. *Biol. Cult. Tests Control Plant Dis. Rep.* 20:FC028, DOI:10.1094/BC20, The American Phytopathological Society, St. Paul, MN.
- Cox, C. M., Garrett, K. A., Cox, T. S., Bockus, W. W., and Peters, T. 2005. Reactions of perennial grain accessions to four major cereal pathogens of the Great Plains. *Plant Dis.* 89:1235-1240.
- Cox, T. S., Bockus, W. W., Gill, B. S., Sears, R. G., Harvey, T. L., Leath, S., and Brown-Guedira, G. L. 1999. Registration of KS96WGRC40 hard red winter wheat germplasm resistant to wheat curl mite, stagonospora leaf blotch, and septoria leaf blotch. *Crop Sci.* 39:597.
- da Luz, W. C., and Bergstrom, G. C. 1986. Temperature alteration of phenotypic expression of spring wheat resistance to *Stagonospora nodorum* spot. *Pesqui. Agropecu. Bras.* 21(6):625-629.
- Davis, M. A., and Bockus, W. W. 2005. Reaction of selected winter wheat cultivars to *Fusarium* head blight, 2004. *Biol. Cult. Tests Control Plant Dis. Rep.* 20:FC029, DOI:10.1094/BC20, The American Phytopathological Society, St. Paul, MN.
- Davis, M. A., Bockus, W. W., and Bowden, R. L. 2000. Reproducibility of results from field and greenhouse evaluations of resistance to *Fusarium* head blight on winter wheat. Page 251 in: *Proc. 2000 Natl. Fusarium Head Blight Forum.*
- Eyal, Z., and Scharen, A. L. 1977. A quantitative method for the inoculation of wheat seedlings with pycnidiospores of *Septoria nodorum*. *Phytopathology* 67:712-714.
- Eyal, Z., Scharen, A. L., Huffman, M. D., and Prescott, J. M. 1985. Global insights into virulence frequencies of *Mycosphaerella graminicola*. *Phytopathology* 75:1456-1462.
- Fuentes, R. G., Mickelson, H. R., Busch, R. H., Dill-Macky, R., Evans, C. K., Thompson, W. G., Wiersma, J. V., Xie, W., Dong, Y., and Anderson, J. A. 2005. Resource allocation and cultivar stability in breeding for *Fusarium* head blight resistance in spring wheat. *Crop Sci.* 45:1965-1972.
- Garrett, K. A., Dendy, S. P., Frank, E. E., Rouse, M. N., and Travers, S. E. 2006. Climate change effects on plant disease: Genomes to ecosystems. *Annu. Rev. Phytopathol.* 44:489-509.
- Jardine, D. J., and Sloderbeck, P. E. 2006. Wheat Variety Disease and Insect Ratings 2004. Kansas Cooperative Extension Service publication MF-991.
- Kim, Y., and Bockus, W. W. 2003. Temperature-sensitive reaction of winter wheat cultivar AGSECO 7853 to *Stagonospora nodorum*. *Plant Dis.* 87:1125-1128.
- Kim, Y., Friebe, B., and Bockus, W. W. 2003. Resistance to take-all root rot is not expressed in wheat-alien chromosome addition and substitution lines. Online. *Plant Health Progress* doi:10.1094/PHP-2003-11XX-01-HN.
- Lamari, L., McCallum, B. D., and DePauw, R. M. 2005. Focenic pathology of Canadian bread wheat: The case of tan spot. *Phytopathology* 95:144-152.
- Martin, T. J. 1978. Procedures for evaluating wheat streak mosaic virus resistance. *Plant Dis. Rep.* 62:1062-1066.
- Nelson, R., Mundt, C., Orrego, R., Ortiz, O., Fredrix, M., Tenorio, J., and Vien, N. V. 2001. Working with resource-poor farmers to manage plant diseases. *Plant Dis.* 85:684-695.
- Ortiz, O., Garrett, K. A., Heath, J. J., Orrego, R., and Nelson, R. J. 2004. Management of potato late blight in the Peruvian highlands: Evaluating the benefits of farmer field schools and farmer participatory research. *Plant Dis.* 88:565-571.
- Raymond, P. J., Bockus, W. W., and Norman, B. L. 1985. Tan spot of winter wheat: procedures to determine host response. *Phytopathology* 75:686-690.
- Roelfs, A. P., Singh, R. P., and Saari, E. E. 1992. Rust Diseases of Wheat: Concepts and methods of disease management. CIMMYT, Mexico, D.F.
- Roozeboom, K., Fritz, A., Stack, J., Whitworth, J., Evans, P., Long, J., Martin, T. J., Schlegel, A., Spangler, M., Claassen, M., Gordon, W. B., Heer, W., Kimball, J., Maddux, L., Parker, E., Seabourn, B., Knapp, M., Bennett, R., Bockus, B., and Shroyer, J. 2005. 2005 Performance tests with winter wheat varieties. Pages 4-28 in: 2005 Kansas Wheat Seed Book. Kansas AES Report of Progress 947. Kansas Agricultural Experiment Station, Manhattan.
- Rufty, R. C., Hebert, T. T., and Murphy, C. F. 1981. Evaluation of resistance to *Septoria nodorum* in wheat. *Plant Dis.* 65:406-409.
- Stack, R. W., and McMullen, M. P. 1995. A visual scale to estimate severity of *Fusarium* head blight in wheat. N.D. Ext. Serv. *Plant Pathol. Bull.* 1095, Fargo.
- Wang, Y. Z., Yong, X. N., and Xiao, Q. P. 1982. The improvement of identification technique for scab resistance of wheat and the development of resistant sources. *Sci. Agric. Sin.* 5:67-77.
- Witcombe, J. R., Joshi, A., Joshi, K. D., and Sthapit, B. R. 1996. Farmer participatory crop improvement. I. Varietal selection and breeding methods and their impact on biodiversity. *Exp. Agric.* 32:445-460.
- Yechilevich-Auster, M., Levi, E., and Eyal, Z. 1983. Assessment of interactions between cultivated and wild wheats and *Septoria tritici*. *Phytopathology* 73:1077-1083.