April 9, 2001

AEIC Technical Comments Regarding Publication on "Performance Assessment of Rapid Immunological Test"

Performance assessment under field conditions of a rapid immunological test for transgenic soybeans. International Journal of Food Science and Technology, April 2001, 36:4, 357-367.

It has come to our attention that a study conducted and authored by Dr. John Fagan and others will be published in the April issue of the *International Journal of Food Science and Technology*. The following critique by the AEIC discusses the evidence that the experimental design of this study is flawed and as a result, the experimental data generated does not support many of the conclusions as stated by the authors.

The authors' main conclusions of this study were: 1) that the use of lateral flow test strips to analyze unknown samples of whole soybeans for the presence of biotech soybeans resulted in a high incidence of false negative and false positive responses; and 2) "…operator performance, not the inherent characteristics of the kit material, were found to be the primary factor influencing the field performance of the test." It is AEIC's opinion that the study organizers did not: 1) provide the participants with sufficient information to allow them to generate results consistent with the manner in which the authors intended to analyze the data; and 2) the design of the study and the authors' interpretation of the findings are inconsistent with the underlying methodology,resulting in the formation of some erroneous conclusions.

According to the authors, operators at 21 grain handling facilities and two state grain analytical laboratories were given duplicate blind samples of 10,000 soybeans prepared at concentrations of 0, 0.01, 0.1, 0.5, 1.0 and 10% "GM content". The authors did not specify to the labs what concentrations to screen for, but instead, instructed the labs to use their "normal sampling procedure." It is our understanding that, in practice, each user of the test screens at different concentrations determined by their unique business considerations. It is therefore an absolute requirement of the methodology that screening and confidence levels be specified prior to running the test in order to generate meaningful data for a particular concentration. Because the protocol did not specify what levels the samples were to be screened at, participants used sub-sample sizes as low as 50 beans. Using the Poisson probability distribution employed by the test procedure a sample containing 50 beans has 0.5, 4.9, 22, 39 and 99% probability of containing a biotech soybean when the actual concentration of biotech beans is 0.01, 0.1 0.5, 1 and 10% respectively (the study concentrations). Clearly, from the above considerations, a large percentage of 50 bean sub-samples would not contain a biotech soybean given the low concentrations provided by the authors and therefore would correctly result in a negative result. It is important to emphasize that the analytical test result on these sub-samples would be negative without regard to the actual method of analysis (i.e., PCR, microtiter plate ELISA, or strip test).

Although a 50 bean sub-sample may be too small for many applications, the fact is that the appropriate number of beans in the sub-sample is dictated by the specific circumstances. A sample size of 50 beans may be appropriate for certain applications. For example, a soybean producer has a 92% probability of detecting a load containing 5% biotech soybeans (current Japanese threshold for labeling) when testing a single sample of 50 beans using the strip test. If the objective of the method was to have a 99% probability of detecting 0.1% biotech beans then the participants could have used a sample size appropriate for that screening level (5 sub-samples of 1000 beans). As it was, the participants seem to have used sample sizes that were appropriate for their own specific applications – not for detection of the concentrations provided by the authors in this study. The authors do point out that "one facility included in the study used much larger sample sizes than the other facilities (2400 beans)". This facility achieved a perfect accuracy score.

The authors, who had provided the samples in duplicate to the labs, state that "Since sampling procedures were identical for replicates, the inconsistencies in results obtained for replicate analyses is not likely to be related to sampling limitations, but is more likely to be due to operator-related variability." We do not agree with this conclusion. If a sample of 10,000 soybeans contains 1% biotech soybeans, simple probability distributions dictate that there is a 63% probability that a sample of 100 beans will contain a biotech bean

and therefore a 37% probability that it won't. Therefore if 10 sub-samples of 100 were taken from the same sample of 10,000 soybeans it is expected that 3 or 4 of the samples would not contain a single biotech bean and would result in a negative test, while 6 or 7 of the samples would indeed contain biotech soybeans and would result in a positive response. The fact that both positive and negative responses are observed in the same sample is not an "inconsistency" of the analytical method but the expected outcome and is certainly not due to "operator-related variability".

The authors further conclude that the test sensitivity is limited to concentrations above 1%. The test performed best at high concentrations in this study because most users have designed their individual sampling strategies for detection at these levels. Most of the users of these tests are testing soybeans at concentrations around the regulatory thresholds specified by their customers selling into Europe (1%) and Japan (5%) and therefore, their sample sizes are designed around these screening levels – not 0.01 and 0.1%. Failure to use appropriate procedures to detect these low concentrations cannot be viewed as operator error because the labs were instructed by the organizers to use their "normal sampling procedures"

The sensitivity of the method is determined partly by the number of beans in the sample and the number of samples analyzed from the load and can be adjusted to various levels of sensitivity with very high reliability. For example, if a person wanted to determine if a load contained 0.01% biotech soybeans, the analyst could test 10 samples of 1000 beans ground together, and providing the method is always positive when there is a single biotech bean in 1000, then 10 negative tests indicates that there was not a single biotech soybean in 10,000. This strategy can be employed to achieve any detection limit as long as the maximum number of beans in the sample is limited to a size where one biotech bean will always be detected. Ultimately, sensitivity of a method is usually limited by practical considerations like cost, time of analysis, etc. and not by the detection level of the analytical test.

If the laboratories were given the information that the authors intended to evaluate the test's capacity to detect at the level of 0.01, 0.1, or even 0.5%, then the laboratories could have selected the correct sampling strategies to detect these levels. The authors' explicit instructions to the laboratories and lack of information regarding the threshold screening concentrations, confidence levels and intended purpose of the study prevented the laboratories from using the test in a way that they could detect biotech beans in the blind samples.

In summary, it is our opinion that the experimental design of this study is flawed and as a result the experimental data generated does not support many of the conclusions as stated by the authors. This study does not add any substantial scientific information to the literature on biotech testing methods. It simply reinforces the necessity to choose an appropriate sampling strategy and testing method based on the particular testing application. No single sampling strategy or testing method can be used effectively for all applications but we do believe that immunoassay strip test method, as specified by the USDA-GIPSA test method protocol, is appropriate for the designated application.