

The statistical theory and analysis of GMO enforcement

Roy Macarthur and Alistair W. A. Murray

Central Science Laboratory, Sand Hutton, York, YO41 1LZ

A novel statistical simulation model has been developed that combines sampling and testing for GMOs in one framework, incorporating the uncertainties in each. This model offers policy makers and regulators an evidence-based tool that can be used to inform decision taking. The model has been used to critically review the performance of current practice in sampling and testing to detect reliably the presence of genetic modification (GM) in oilseed rape at permitted thresholds under a range of commonly encountered situations. The model structure explicitly includes all stages of the sampling and testing process. Its mathematical and statistical calculations are derived from the best available and generally accepted theory. The input parameters of the model relate directly to information commonly available in the course of collecting samples and conducting laboratory tests. Heterogeneous distribution of GM seeds in bulks is important because, as the degree of heterogeneity increases, the choice of appropriate sampling plan to ensure samples are representative of the material being analysed becomes increasingly important. Heterogeneity can be modelled by considering the bulk/lot as composed of a mixture of 'sub lots' each characterised by a mean proportion of GM seed, and specifying how the lot is partitioned into these 'sub lots'. A general model has been developed where each 'sub lot' conforms to the binomial distribution and the mixture of these (i.e. whole lot) is modelled by the beta distribution – this is the 'beta-binomial' model. Heterogeneity is input to these models as the proportion of the bulk / lot expected to contain 100% or 95%, respectively, of any GM seeds present. This is an intuitive parameter that could be estimated by expert opinion where measurements are not available. The application of the model has confirmed the interdependence of sampling and testing in determining the reliability of the combined process. Results from the range of practical scenarios for oilseed rape modelled show that current best practice is generally reliable in assuring that existing thresholds down to 0.1% can be met by the analysis of samples (of 3000 seeds) from a small number of increments taken from a lot, even in the presence of heterogeneity, provided that suitable analytical replication is employed. It is essential to combine both sampling and testing uncertainty in any model of the reliability of GM seed testing. Heterogeneity is a crucial factor in such a model and should be incorporated in a physically-plausible manner. Results from simulations under a range of commonly-encountered scenarios for GM inspection and enforcement of oilseed rape seed show that good practice can ensure the reliability of the sampling and testing process in meeting the requirements of legislation on thresholds and confidence levels. This study has objectively demonstrated the information required to ascertain the reliability of claimed results. The method is readily applied to any sampling + detection/measurement system.