

Spiroux de Vendomois et al. 2009: FSA internal review

Statistical assessment

The statistical tests applied in the paper are appropriate for the datasets analysed.

Toxicological assessment

It is worth noting that the toxicological tests applied in accordance with the OECD guidelines use predefined statistical methods, and this is a post hoc analysis.

In the opinion of the Agency's Chemical Risk Assessment Unit (CRAU), this paper does not present anything new, it is not infrequent that some parameters in a toxicological study will show statistically significant differences between groups; however in deciding whether they represent an actual (let alone adverse) effect it is essential to look at whether the changes are biologically significant. Of course the larger the study and the more parameters that are measured the greater the number of chance statistical differences.

If there is a statistical difference between groups, the first question is whether the parameter value is outwith the normal range. Group sizes in toxicological studies are a compromise between the ability to detect effects and animal welfare considerations, as such it is not uncommon to see that values of one or more parameters are tightly clustered in different groups and thus a statistical difference can be seen. However when the values of the group mean values are looked at on a population basis, both will be within the normal range and the apparent difference is an artefact rather than a biologically significant event. Even if values are outside the "normal" range this may reflect measurement of the range within a particular laboratory, rather than the true normal range. Hence it is then necessary to look at temporal and spatial consistencies to assess the parameter. Where effects occur at different times and doses, but not consistently, it is difficult to ascribe these to a biologically plausible mechanism. In the view of CRAU none of the statistically significant findings in this paper represent clearly biologically significant changes.

Specific observations on points raised in the paper

The comments on the group sizes in section 2.1 notes that "the investigators claimed that OECD guidelines and standards were followed" but seems to query the group size of 20 per dose (10 per sex per group) and that only half of each group were used. OECD 408 states, "At least 20 animals (ten female and ten male) should be used at each dose level." It is common to measure parameters in a portion of the group. We would agree with the authors that a third dose level would have been desirable, especially for looking at trends in biological observations. It is worth noting even with two dose levels there are around 48000 measurements in these experiments (80 parameters in 10 animals at 2 time points in 10 groups and 3 experiments); therefore the likelihood of chance findings is high.

NK603 results

The blood parameter differences appear to be random and there is no consistent biological pattern. A 29% increase in lymphocytes in males following a decrease at the lower dose is not statistically significant, but a 23% decrease in females following a decrease at the lower dose is; eosinophils at 5 weeks and a 38% increase at the lower dose in males is significant but a 43% increase in females at the same dose is not. All appear to be random differences and may well be in the normal range (these data are not presented). Other changes seen at 5 weeks but not at 14 and with no supporting histopathological changes could (if they were effects, and we might dispute this) be interpreted to represent adaptive changes with no long term consequences.

MON810 results

Again there are no consistencies and similar anomalies are seen with similar results being significant in one sex but not the other. The changes in white blood cells might be attributable to transient infection in the colony.

MON863 results

The kinetic plots of results illustrate the biological folly of the approach; for example in the triglyceride plot the test animals are more tightly clustered than the controls and this over-emphasises an apparent difference caused by the control mean being heavily influenced by three values (1, 4 and 9, two of which are outliers) and appearing to decline, whilst no change is seen in test. Looking at the data, triglyceride varies around a similar value for both with no trends (some go up some down, others don't change). These plots also suggest there may be systematic differences between the groups.

The discussion is riddled with unsupported points and statements that rely more on belief than evidence. Changes are assumed to be adverse without thinking of the actual biology. Differences are attributed solely to the GMO feeding as "the standardized conditions of rat maintenance employed, which are stated to be in accordance with OECD standards, make the diet the only factor of variation in the protocol." But there can still be variations in conditions (the standardised procedures minimise, but do not prevent these) and bias (in all its forms). We would not agree that "the methodology we employed revealed different effects, which completely changed the interpretation of the experimental results," since statistical differences without biological significance do not change the interpretation of the study. The authors also ignore the lack of consistency in effects on biological markers. The arguments that effects may be due to "the presence of residues of Roundup herbicide that are present in GM crop varieties," are unsupported by data on Roundup toxicity and its dose response relationship and are highly speculative. For example the relevance and significance (or lack thereof) of the cited *in vitro* results in embryonic kidney cells to an overall risk assessment is not considered.

In light of the lack of reported functional or histopathological changes then differences in some isolated parameters would not necessarily be considered adverse, especially when they are not consistent or biologically coherent. As such the consensus view would be that these were biologically insignificant and to

discount them (as the review of the MON 863 study by Doull *et al* did). We do not agree that these data “strongly suggests that these GM maize varieties induce a state of hepatorenal toxicity” in the absence of functional or histopathological changes.