

**THIS IS AN UNOFFICIAL TRANSLATION PREPARED ON BEHALF OF
THE UK FOOD STANDARDS AGENCY**

the original French text is published at <http://www.ogm.gouv.fr/>

High Council of Biotechnologies

Scientific Committee

Opinion relating to the deposition of 15 December 2009 by the Member of Parliament, François Grosdidier, as to the conclusions of the study entitled “*A comparison of the effects of three GM corn varieties on mammalian health*” by J. Spiroux de Vendômois, F. Roullier, D. Cellier and G.E. Séralini, *Int. J. Biol. Sci*, 2009: 5(7) : 706-726.

Summary of the opinion:

The High Council of Biotechnologies (HCB) has adopted the findings to the effect that weaknesses in statistical analysis make it impossible to conclude with sufficient certainty that there are no health or environmental risks related to GMOs. These weaknesses are now accepted by the assessment authorities (EFSA).

The study by J. Spiroux de Vendômois *et al.*, like that presented by Monsanto for the same GMOs, limits itself to determining whether the differences observed between the control and treatment groups are statistically significant. Although the two studies differ in their interpretation of whether the differences found are significant, the weakness of the statistical analysis presented makes it impossible to reach a definitive conclusion.

In their analysis based on the initial data supplied by Monsanto, J. Spiroux de Vendômois *et al.* demonstrate that very few of the differences observed can be regarded as statistically significant. Nonetheless, the authors attempt to prove that these differences are biologically significant.

Only the arguments of a statistical nature, unacceptable to a statistician or toxicologist, are held up to justify the conclusions of this study. At this point it is important to reiterate the role of statistics: they are a means to assist decision-making but not a decision-making tool. It is not statistics that make it possible to conclude whether a GMO is hazardous to human health or not. Statistics are there to assist the toxicologist in correctly assessing the risks of reaching the wrong conclusions as to the absence or presence of negative effects. A confluence of events without statistical significance might still lead the toxicologist to assess a substance as potentially toxic to a target tissue or organ.

However, although the authors are right to take up these arguments, their conclusions are based solely on an interpretation of certain isolated changes in haematological and biochemical constants. No link is made to other key parameters in the assessment of toxicity, for example the results of macroscopic and histological examinations.

In conclusion, the Scientific Committee (SC) of the High Council of Biotechnologies (HCB) indicates that the study by J. Spiroux de Vendômois *et al.* presents no admissible scientific element likely to ascribe any haematological, hepatic or renal toxicity to the three re-analysed GMOs.

A. Working methodology

The deposition was validated at the meeting of the HCB bureau on 21 December 2009. In order to respond to this request, the SC relied on the expertise of three of its members. The document was reviewed by the chairman and vice chairman of the SC before being distributed to all of the members of the SC.

B. Comments on the statistical methodology

Preamble: towards more persuasive statistical processing

As the HCB has already indicated, not only the protocol but also the statistical analysis traditionally used by a number of petitioners, including Monsanto, show certain weaknesses that make it impossible to conclude with sufficient certainty that there are no health and environmental risks associated with GMOs. These weaknesses are now largely accepted outside the HCB. Indeed,

- The EFSA has acknowledged the need to review its guidelines:
http://www.efsa.europa.eu/EFSA/efsa_locale-1178620753812_1211902768517.htm
- The French Food Safety Agency (AFSSA) has set up a statistical working group in order to propose new methods for processing the results of studies on the toxicity of GMOs.
- The 4th Biosafenet seminar, 12-15 January 2009 (Italy), on the theme of "*The statistical design and analysis of field trials for assessing the risk associated with GM plants, with a focus on Non-Target Organisms*" had the objective to define new statistical procedures to assess the environmental risks associated with GMOs.
- All of the recommendations for the assessment of the environmental risks associated with GMOs are described in an article published at the end of the seminar, in the journal *Environmental Biosafety Research* (Vol. 8, no. 2) "*Statistical aspects of environmental risk assessment of GM plants for effects on non-target organisms*".

The HCB has adopted these new rules, which should make it possible to improve the assessment of the health and environmental risks associated with GMOs. The HCB analyses any study claiming to demonstrate the absence or presence of negative effects associated with GMOs systematically and with the same scientific rigour.

Comments specific to the article and concerning the statistical methodology

The study by J. Spiroux de Vendômois *et al.* suffers from the same idiosyncracies and weaknesses as the studies presented by Monsanto in its licence application dossier. Although the two studies differ in their interpretation of whether the differences found are significant, the authors limit themselves to determining whether the differences observed between control and treatment groups are statistically significant. The objective of these comparative tests is to test the existence of effects due to diet, but the extent of those effects is not considered. Yet any change in diet will have an effect, however negligible. The right question to ask would be: “is the effect of the GMO sufficiently great as to indicate potential toxicity?”. Equivalence tests are therefore a useful tool, not to answer this question but to assess the risks of giving the wrong answer.

The analysis presented in the Monsanto dossier uses historic data to explain that the differences observed are not biologically significant. However, the statistical methodology lacks rigour and is not acceptable to the statistician.

The statistical study contained in the article by J. Spiroux de Vendômois *et al.* does not call for lengthy comments either, since it limits itself to the performance of numerous comparative tests and analysis of principal components for each of the three corn varieties studied. Although the statistical methodology is sometimes debatable, it is essentially the interpretation of the results of the statistical analysis that is most questionable.

For example:

- The analysis of principal components shows that there are physiological differences between males and females... is that surprising? What is it supposed to demonstrate?
- Figures 3 to 7 present the variation of parameters between weeks 5 and 14. For NK603, for example, the authors present just one parameter chosen from several hundred: the clearance of creatinine. They note a faster decrease in the GMO group than in the control group, but it converges in the two groups at week 14. It may be wise to consider the differences between weeks 5 and 14, provided that:
 - i)* a more rigorous statistical methodology is applied, taking account of a mixed model: in this case the fixed effects are the population parameters, including possible effects of gender, diet, dose etc., whereas the random effects follow the inter-subject variability among the population,
 - ii)* a persuasive biological interpretation is put forward (as mentioned below in the paragraph on toxicology) that creatinaemia and uraemia, which are associated with glomerular function, usually increase in the case of renal failure and not the other way around. Similarly, the clearance of creatinine reduces in case of renal impairment and the variations noted in the article do not suggest any toxicity, since the GMO group has better glomerular filtration than the non-GMO group in week 5 and the values converge in week 14.
- Calculating a large number of p-values (*i.e.* degrees of significance¹), retaining only a few (those that correspond to differences greater than 5%) then analysing them retrospectively according to gender or dose may have perverse effects (these issues of causality in epidemiology are well known, see Gauvrit, 2007 for example). Concluding

¹

Probability that any difference observed is not due to the diet but to chance.

HIGH COUNCIL OF BIOTECHNOLOGIES

that the GMO has an effect on just one gender, at only one observation time or at just one level of exposure (different for each parameter considered) implicitly means that there are numerous interactions between diet, gender, dose and time. Given the small sample size, this model is far too complex to be regarded as having the slightest predictive value: to do so would contradict the principle of parsimony, which prefers “simple” models with few interactions to complex models.

- The results presented in this study concerning MON863 are not new as they have already been published in 2007 by Séralini *et al.*. This preceding study included some methodological errors. In particular, the p-values calculated for the comparison of the weight curves were incorrect and the differences presented as very significant² are in fact absolutely not (a correct calculation gives p-values all above 0.20).
- Other remarks concerning more technical aspects:
 - i) the study applies a very large number of parametric and non-parametric tests; for example, there is little point in testing the normality of 10 pieces of data using a Shapiro test, as this is not very effective with small samples. Uniformly spread data will therefore be regarded as Gaussian in more than 90% of cases!
 - ii) It is correct to adjust the p-values using techniques like FDR, but the Benjamini-Yekutieli method presumes quite strong hypotheses that there is no reason to verify here (there is no systematically positive correlation between the different statistics).

C. Comments specific to the article and concerning the toxicological interpretation

In their analysis based on initial data from Monsanto, J. Spiroux de Vendômois *et al.*, attempt to show that these differences are biologically significant, but the arguments used are not acceptable to a statistician or toxicologist.

General observations

- The authors note that a certain number of parameters (CYP 450, endocrinal parameters) were not taken into account. These parameters are not demanded by the OECD protocol. However, a Working Group set up by the AFSSA in March 2009 will soon make proposal in order to improve the relevance of the parameters considered.
- The authors point out that the OECD protocol demands 3 doses, whereas the GMO is assessed only at two doses. Although this remark is well founded, it is still necessary to take account of the discussion of the “choice of doses” in the toxicity studies. In this case, the maximum dose is dictated by the observance of dietary balance between nutritional intakes (33% in the case of these corn varieties). A lower dose makes it possible to assess the possible relationship between a high dose and a lower dose. The relevance of a third, even lower dose is discussed on a case-by-case basis.
- The authors indicate that the “deleterious” effects are less pronounced at the start of exposure than after more prolonged treatment. That clearly depends on the nature of the possible toxic effects and the dose, as some changes may be seen from the first doses onwards. This is particularly the case for hepatotoxic or nephrotoxic substances, chemicals or heavy metals, but also medicinal plants such as the common germander

² “However, we clearly proved very significant differences in weight growths for both males and females”, Séralini *et al.*, 2007

HIGH COUNCIL OF BIOTECHNOLOGIES

(Laliberté and Villeneuve, 1996) or green tea (Santé Canada, 2007; Bonkovsky, 2006; Gloro *et al.*, 2005), which has led to liver transplants, or snakeroots responsible for nephrotoxicity (Chen, 2000; FDA, 2001).

- The authors mention that gender differences are often used to reject significant facts. This is a totally incorrect point of view. If an event is regarded as biologically significant for only one gender, it is obviously taken into account. If an event is statistically significant for only one gender but not biologically significant, the fact that it is not observed for the other gender (neither statistically speaking nor biologically speaking) makes it possible to declare that the phenomenon has no biological or toxicological significance and can therefore be regarded as isolated.
- The authors write that *“the absence of a linear dose-effect graph is used to draw conclusions about the safety of the GMO, which is nonsensical given the application of 2 doses”* (reference § 2.3. p 708). The first part of this statement is false. What is taken into account is a possible increase in the effects observed between a low and high dose. An effect at the low dose that is not confirmed at the high dose has virtually no toxicological significance. For example, a significant reduction in serum albumine after 14 weeks of treatment may suggest a hepatotoxic effect. However, the fact that the reduction occurs only at the lowest dose and in one gender, whereas one observes an insignificant effect in both genders at the highest dose, has no toxicological relevance (table 2 line 4, MON863 of the previous commented article (Séralini *et al.* 2007)).
- Although the authors are right to take up the toxicologists' arguments that *“the isolated signs of toxicity do not constitute proof of adverse effects on health”*, their conclusions are based solely on an interpretation of certain isolated changes in haematological and biochemical constants. No link is made to other key parameters in the assessment of toxicity, for example the results of macroscopic and histological examinations.
- According to the publications of J. Spiroux de Vendômois, the criteria for the assessment of hepatotoxicity are not those found in the guidelines of the *“Non-clinical guideline on drug-induced hepatotoxicity”* published by the EMEA in 2008 (cf. Guillemain 2009). The EMEA guideline indicates that *“Increases in the levels of the liver enzymes alanine aminotransferase (ALT) and aspartate aminotransferase (AST) in serum, in combination with increased bilirubin levels, are actually considered to be the most relevant signal of liver toxicity. Macroscopic and in particular histopathological observations will allow confirmation of the occurrence of liver toxicity and will provide further evidence of the type of liver toxicity.”* The guidelines specify that *“Total protein, albumin, triglycerides, cholesterol, glucose and blood urea nitrogen, activated partial thromboplastin time (APTT) and prothrombin time (PT) can be used as supplementary tests for hepatic synthetic functions.”*
- Lastly, only arguments of a statistical nature are put forward to justify the conclusions of this study. However, it is important to reiterate that the role of statistics is to be a tool to aid decision-making, not a decision-making tool! It is not statistics that make it possible to conclude whether a GMO is hazardous to human health or not. Statistics are there to assist the toxicologist in correctly assessing the risks of reaching the wrong conclusions as to the absence or presence of negative effects. A confluence of events without statistical significance might still lead the toxicologist to assess a substance as potentially toxic to a target tissue or organ.

Data analysis

MON863 corn: As in the case of MON 810, the authors indicate that several differences reported are not statistically significant once the FDR adjustment is applied. It is therefore irrelevant to mention slight differences such as the weights of the animals (around 3%) that are not statistically significant.

Whereas an increase in the transaminases ALT and/or AST is regarded as a dominant marker of cellular hepatotoxicity, the only significant change previously reported (Séralini *et al.*, 2007) is the reverse, namely a reduction in ALT at the lowest dose of the GMO. For that matter, the high dose has no effect in the male and female animals... It should be noted that tables 1 and 2 of the publication (Spiroux de Vendômois *et al.*, 2009) make no mention of these dominant parameters either for NK603 or MON 810. It is therefore strange to regard an increase in serum albumine as a warning value for liver impairment (table 2 line 3, MON863, 2007) when in fact the opposite would be expected... It will also be noted that the values observed are not very different: an average of 4.850 g/dl for MON863 and 4.600 g/dl for isogenic corn, whereas the values for two commercial varieties are 4.890 and 4.820 g/dl. In this case, one may think that the relativity and biological relevance of the variation are disregarded.

NK603 corn: A significant difference to the control corn varieties (isogenic and/or reference varieties) is detected in 39 cases. There are only 23 differences greater than 5%, including 18 among the males, which prompts the authors to conclude that there is a gender effect – which might be due to greater sensitivity in the males. Such a conclusion is not admissible in toxicological terms, in as much as heterogeneous results are added without the slightest interpretation of toxicological plausibility. For example, the increase in eosinophils at 5 weeks, followed by a reduction at 14 weeks, or indeed the unexplained reduction in blood urea. Indeed, whilst chronic nephrotoxicity is usually accompanied by an increase in blood urea, table 1 referring to NK603 reports significant effects at both doses in the male animals. The direction of the variation is the reverse of that expected in case of chronic nephrotoxicity (table 1, line 6 *kidney*). Note that this effect is reported for week 5 only and it is not reported at the end of the 13 weeks of treatment, which must at the very least raise questions for any toxicologist as to the direction and biological consequences of such a variation.

Moreover, the argument of a relationship with a dose-dependent effect appears to be based solely on a greater frequency of significant differences at the higher dose, whatever the observation time), whilst the authors themselves emphasise that “*the absence of a dose-effect relationship is inevitable as only two doses are administered.*” Lastly, no reference is made to the data from the histopathological examinations, even though they would be crucial to establish toxicity to organs and tissues.

MON810 corn: Of the 450 differences calculated, it is statistically expected that one will observe 22 or 23 statistically significant differences under the null hypothesis. To observe 29 of them for MON810 (of which only 15 are greater than 5%) is not surprising and it proves absolutely nothing. Furthermore, the authors recognise that when an FDR adjustment is applied to the calculation of the p-values, none of the differences between the test samples and control samples is statistically significant. Nevertheless, it is legitimate to ask oneself whether these statistically insignificant differences have any biological significance, but the

HIGH COUNCIL OF BIOTECHNOLOGIES

authors provide absolutely no evidence of any toxicity to a target tissue or organ. Mentioning a significant reduction in serum albumine at the high dose of MON810 in the male animals (table 2, lines 1 and 2 *liver*), with no change to any other biochemical parameter (in particular the markers regarded as dominant, ALT and AST) and no report of any histological alteration, is absolutely no basis on which to draw conclusions on liver toxicity.

Similarly, mentioning a slight increase in the weight of the kidneys at the lowest dose for one gender (table 2, lines 1 and 2 *kidney*), without a simultaneous change in one or more biochemical parameters or any detection of this effect at the high dose, has no toxicological significance. It is worth noting that the authors had in fact emphasised the significant reduction in the weight of the kidneys in the MON863 study to indicate nephrotoxicity. Incidentally, these effects were not observed in the studies with the hybrids MON863 x NK603 and MON863 x MON810 x NK603. This demonstrates that if isolated effects must be considered, they must also be likened to events contributing to the notion of a set of arguments.

The variations in blood count (*White blood cells* and *lymphocyte*) in the females at the high dose cannot be interpreted as they are, given the positive and then negative variations depending on dose. These variations are observed at 5 weeks and not reported at the end of the study. As for the variations in organ weight, these relate only to the lowest dose for one gender and do not correspond to any biological effect.

D. Summary and conclusion

The approach developed in the article by J. Spiroux de Vendômois *et al.* focuses on the statistical differences between various genetically modified corn varieties and isogenic controls or commercial varieties.

This publication is merely a list of the differences, with no attempt at biological or toxicological interpretation. As is repeatedly emphasised by the international institutions charged with assessing toxicological risks, a significant statistical difference does not necessarily prove the existence of a biological disorder. As a consequence, the argument involving a list of significant differences between the exposed animals and control animals is not admissible. Furthermore, in most cases the “differences” observed relate to just one gender, one observation time or one level of exposure, without the slightest tendency to link the variation to a particular intensity and period of exposure. No hypothesis is presented to demonstrate that these variations by gender are the result of a gender-dependent effect relating to an endocrinal disturbance. What is more, some of the variations cited in the publications are the reverse of those usually regarded as evidence of a toxic effect, particularly on the liver or kidneys.

As a result of the approximations, deficiencies and errors of interpretation in the article by J. Spiroux de Vendômois *et al.*, it is impossible to conclude that the three GMOs re-analysed on the basis of the initial data supplied by the petitioner have any haematological, hepatotoxic or nephrotoxic effect.

In conclusion, the HCB indicates that the study by Spiroux de Vendômois *et al.*, 2009, like a preceding study (Séralini *et al.*, 2007), presents no admissible scientific element likely to ascribe any haematological, hepatic or renal toxicity to the three re-analysed GMOs.

Please note:

It will also be noted that the authors' absence of conflict of interest, which is mentioned at the end of the article, might be questioned. On 5 January 2010, the body to which the authors belong continues to display on its public website the results of studies – including those of the Austrian study of November 2008 – claiming to demonstrate negative effects of MON810 on reproduction, even though those results have been recognised as erroneous by the authors of the study themselves.

Bibliography:

Bonkovsky, H. L. (2006). « Hepatotoxicity associated with supplements containing Chinese green tea (*Camellia sinensis*) ». *Ann Intern Med*, 144, 68-71.

Chen, J. (2000). Nephrotoxicity associated with the use of *Aristolochia* ». *Herbal Gram*, 48, 48-49.

EMA: *Non-clinical guideline on drug-induced hepatotoxicity*
EMA/CHMP/SWP/150115/2006, 24 January 2008.

Environmental Biosafety Research (Vol 8, n° 2) "*Statistical aspects of environmental risk assessment of GM plants for effects on non-target organisms*"

Food and Drug Administration (2001). *FDA warns consumers to discontinue use of botanical products that contains aristolochic acid*. U.S. Food and Drug Administration.

Gauvrit, N. (2007), *Statistiques, Méfiez-vous !* Paris : Ellipses

Gloro, R. and coll. (2005). « Fulminant hepatitis during self-medication with hydroalcoholic extract of green tea. ». *Eur J Gastroenterol.Hepatol.*, 17, 1135-1137.

Guillemain J. Evolution des protocoles et critères de preuve dans les tests toxicologiques. Colloque INRA « Les OGM face aux nouveaux paradigmes de la biologie, Paris, février 2009.

Laliberté, L. and Villeneuve, J. P. (1996). Hepatitis after the use of germander, a herbal remedy. *CMAJ*, 154, 1689-1692.

Santé Canada (2007). « Extrait de thé vert (Green Lite) : lien soupçonné avec une hépatotoxicité ». *Bulletin canadien des effets indésirables*, 17.

Séralini G.E., Cellier D., Spiroux de Vendômois J. *Arch. Environ. Contam. Toxicol.* (2007). New analysis of a rat feeding study with a genetically modified maize reveals signs of hepatorenal toxicity. 52, (4) pp 596-602.

Spiroux de Vendômois J., Roullier F., Cellier D., Séralini, G.E. Aison of the effects of three GM corn varieties on mammalian health. *Int. J. Biol. Sci.*, 2009, 5(7), 706-726.