

# HAUT CONSEIL DES BIOTECHNOLOGIES

---

## COMITE SCIENTIFIQUE

Paris, le 2 août 2011

REPONSE A LA CONSULTATION PUBLIQUE DE L'AESA SUR LE  
PROJET DE DOCUMENT D'ORIENTATION RELATIF AUX ETUDES  
DE TOXICITE SUB-CHRONIQUE DE 90 JOURS SUR LES  
DENREES ALIMENTAIRES ET ALIMENTS POUR ANIMAUX MENEES  
CHEZ LES RONGEURS

Le Haut Conseil des biotechnologies (HCB) a été sollicité par la Direction générale de la prévention des risques du Ministère de l'écologie, du développement durable, des transports et du logement pour examiner le projet de document d'orientation de l'Autorité européenne de sécurité des aliments (AESA) relatif aux études de toxicité sub-chronique de 90 jours sur les denrées alimentaires et aliments pour animaux menées chez les rongeurs, et pour l'inviter à répondre à la consultation publique correspondante de l'AESA.

Le Comité scientifique (CS)<sup>1</sup> du HCB a procédé à l'examen de ce document par voie électronique sous la présidence de Jean-Christophe Pagès. Les commentaires du HCB à destination de l'AESA sont transmis par ce rapport aux autorités françaises.

---

<sup>1</sup> La composition du CS est indiquée dans l'Annexe 2.

## TABLE DES MATIERES

<b>1. INTRODUCTION.....</b>	<b>3</b>
<b>2. COMMENTAIRES SUR LE DOCUMENT D'ORIENTATION DE L'AESA SUR LA MISE EN ŒUVRE DES ÉTUDES DE TOXICITÉ SUB-CHRONIQUE DE 90 JOURS SUR LES DENRÉES ALIMENTAIRES ET ALIMENTS POUR ANIMAUX MENÉES CHEZ LES RONGEURS .....</b>	<b>4</b>
2.1 REMARQUES GÉNÉRALES .....	4
2.2 COMMENTAIRES PAR SECTION .....	4
<b>ANNEXE 1 : ELABORATION DES COMMENTAIRES .....</b>	<b>6</b>
<b>ANNEXE 2 : COMPILATION DES COMMENTAIRES TRANSMIS À L'AESA EN ANGLAIS .7</b>	<b>7</b>
1. GENERAL COMMENTS .....	7
2. COMMENTS PER SECTION .....	7

## 1. Introduction

En réponse à une demande de la Commission européenne (question EFSA-Q-2009-00941), le Comité scientifique de l'Autorité européenne de sécurité des aliments (AESA) a élaboré ce document d'orientation dans l'objectif d'aider les pétitionnaires à concevoir, mettre en œuvre, analyser, interpréter et rendre compte des études de toxicité subchronique de 90 jours sur les denrées alimentaires et aliments pour animaux chez les rongeurs. Cette demande s'inscrit dans le cadre du Règlement (CE) n° 1829/2003<sup>2</sup>, relatif aux denrées alimentaires et aliments pour animaux provenant d'organismes génétiquement modifiés, et dans le cadre du Règlement (CE) n° 258-97<sup>3</sup> sur les nouveaux aliments et ingrédients alimentaires (en cours de révision).

Jusqu'à présent, les pétitionnaires étaient invités à se conformer aux lignes directrices de l'OCDE sur les essais de produits chimiques (OECD, 1998)<sup>4</sup> pour évaluer les risques associés aux OGM et aux nouveaux aliments et ingrédients alimentaires. Leurs analyses variaient de façon significative sur le design des expériences, les régimes testés, les doses employées, les paramètres biologiques à mesurer et les approches statistiques.

Récemment, l'Anses<sup>5</sup> a publié des recommandations sur l'analyse statistique de ces études dans le contexte des demandes d'autorisation de mise sur le marché d'organismes génétiquement modifiés (Anses 2011)<sup>6</sup>, qui ont contribué utilement à l'élaboration de ce document d'orientation par l'AESA.

Par souci de transparence et en vue d'améliorer ce document en concertation avec la communauté scientifique et d'autres parties prenantes qui souhaiteraient apporter leur contribution, l'AESA a ouvert une consultation publique<sup>7</sup> sur une première version de ce document.

Le Comité scientifique (CS)<sup>8</sup> du Haut Conseil des biotechnologies (HCB) a procédé à l'examen de ce document par voie électronique sous la présidence de Jean-Christophe Pagès. Par souci d'efficacité et de complémentarité avec l'analyse de l'Anses sur les aspects de toxicité du document, seuls les commentaires du HCB concernant l'analyse statistique ont été retenus. Ces commentaires sont transmis par ce rapport aux autorités françaises ; leur traduction en anglais (Annexe 2) a été postée dans les formulaires de réponse mis à disposition sur le site de l'AESA pour recueillir les commentaires du public.

---

<sup>2</sup> Règlement (CE) n° 1829/2003 du Parlement européen et du Conseil du 22 septembre 2003 concernant les denrées alimentaires et les aliments, consistant en, ou contenant des, ou issus d'organismes génétiquement modifiés, pour l'alimentation humaine et animale.  
<http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:32003R1829:FR:HTML>.

<sup>3</sup> Règlement (CE) n° 258/97 du Parlement européen et du Conseil du 27 janvier 1997 relatif aux nouveaux aliments et aux nouveaux ingrédients alimentaires.  
<http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:31997R0258:fr:HTML>.

<sup>4</sup> OECD Guideline for the Testing of Chemicals – Repeated Dose 90-day Oral Toxicity Study in 1007 Rodents, 408, 1998. <http://browse.oecdbookshop.org/oecd/pdfs/free/9740801e.pdf>.

<sup>5</sup> Anses : Agence nationale de sécurité sanitaire de l'alimentation, de l'environnement et du travail.

<sup>6</sup> Anses, 2011. Recommandations pour la mise en œuvre de l'analyse statistique des données issues des études de toxicité sub-chronique de 90 jours chez le rat dans le cadre des demandes d'autorisation de mise sur le marché d'OGM. Saisine n° 2009-SA-0285.

<sup>7</sup> La consultation publique est ouverte sur le site de l'AESA. Le document soumis à consultation et la procédure de réponse à la consultation sont disponibles sur le lien : <http://www.efsa.europa.eu/fr/consultations/call/110707.htm>.

<sup>8</sup> La composition du CS est indiquée dans l'Annexe 1.

## 2. Commentaires sur le document d'orientation de l'AESA relatif aux études de toxicité sub-chronique de 90 jours sur les denrées alimentaires et aliments pour animaux menées chez les rongeurs

### 2.1 Remarques générales

Le HCB se réjouit de la décision de l'AESA de publier un document d'orientation relatif aux études de toxicité sub-chronique de 90 jours sur les denrées alimentaires et aliments pour animaux chez les rongeurs. En effet, les méthodologies utilisées jusqu'à présent par les pétitionnaires présentaient des lacunes, ce qui ne permettait pas toujours de valider les conclusions formulées dans les dossiers soumis pour évaluation.

Le HCB considère que plusieurs aspects méthodologiques importants sont maintenant correctement abordés dans ce document d'orientation :

- Seuls les tests d'équivalence autorisent à conclure à une éventuelle équivalence entre la plante génétiquement modifiée et son comparateur.
- Ce ne sont pas uniquement les valeurs des *p-values* mais surtout les tailles d'effet qu'il convient d'examiner attentivement. En effet, des différences statistiquement significatives ne traduisent pas nécessairement un effet biologique significatif, et réciproquement.
- La multiplicité des tests mérite d'être correctement prise en compte, au moyen d'outils tels que le FDR (*False Discovery Rate*).
- Le choix de la taille des échantillons doit être motivé par des analyses de puissance des tests statistiques mis en œuvre.

### 2.2 Commentaires par sections définies par l'AESA

*N.B.* : Les titres soulignés correspondent aux sections de dossier définies par l'AESA, et aux différents formulaires mis à disposition par l'AESA pour la collecte de commentaires en ligne. Seules les sections pour lesquelles le HCB transmet des commentaires sont indiquées ici. Chaque commentaire est écrit de manière indépendante. La somme des commentaires n'est pas destinée à constituer un texte en soi.

## APPENDICES

### APPENDIX 1 – STATISTICAL PRINCIPLES AND GOOD EXPERIMENTAL DESIGN

#### 4. Considerations when designing an experiment

##### 4.2. High power

##### 4.2.4. Increasing the significance level

p. 30, l. 1159-1165 :

*“By convention most investigators use a 5 % significance level. If it were to be increased to 10 % any real effects which just failed to reach significance at the 5 % level would now be judged “significant”. However, this will also increase the number of false positive results (Type I errors). As toxicity tests usually involve many outcomes, there is in any case a problem with an excess of false positive results. For the assessment, strong emphasis should be placed on the biological relevance of any observed differences whether or not they reach the chosen level of statistical significance. This is best done by looking at the point and interval (e.g. confidence) estimates and not by just focussing on the P-value.”*

Ce paragraphe évoque le risque de détecter de nombreux faux-positifs en augmentant le niveau du test. Ceci est vrai dans un cadre théorique, lorsqu'il n'existe absolument aucune différence pour plusieurs paramètres (les négatifs) et des différences pour d'autres (les positifs) et que ces paramètres sont indépendants. Le problème ne se pose pas ici

exactement en ces termes : tous les paramètres peuvent être des positifs (il n'y a aucune raison de penser qu'une différence de régime ne produit absolument aucun effet) mais il s'agit de différencier les différences anodines pour la santé (les « négatifs ») de celles indiquant un effet biologiquement significatif (les « positifs »). De plus, une telle analyse nécessite une approche globale, c'est-à-dire un examen de l'ensemble des effets.

## 6. Statistical analysis

p. 34, l. 1359-1367 :

*“The separate analysis of many parameters, most of which are not expected to differ between treatment groups, may result in a large number of statistical tests. In order to control the number of false positive results the use of false discovery rate (FDR) methods have been suggested (Kall, 2008), although their use in the analysis of toxicity tests is not well established. The FDR is the estimated proportion of false positives among all the significant hypotheses tested. However, this technique is not applicable in experiments where there are no strongly positive responses. Should there be no real differences between the groups being compared across many parameters, then all the positive results will be false positives and the FDR will be 100 %. Therefore, this method is only recommended when there are some strong and statistically highly significant differences between the groups.”*

Ce paragraphe sur les tests multiples et l'utilisation de correction de type FDR est assez confus. En l'absence de différences, il est attendu qu'une proportion  $\alpha$  des *p-values* **non corrigées** soient inférieures à  $\alpha$  (on s'attend donc à une proportion  $\alpha$  de faux positifs). Par contre, toujours en l'absence de différences, la probabilité qu'au moins une *p-value* **corrigée** soit inférieure à  $\alpha$  est précisément  $\alpha$  (on s'attend donc à ne pas détecter de faux positifs avec une probabilité  $1-\alpha$ ).

## **Annexe 1 : Elaboration des commentaires**

Ces commentaires ont été élaborés par le CS du HCB, composé de :

Jean-Christophe Pagès, Président, Jean-Jacques Leguay, Vice-Président,

et par ordre alphabétique des noms de famille : Yves Bertheau, Pascal Boireau, Denis Bourguet, Florence Coignard, François-Christophe Coléno, Jean-Luc Darlix, Elie Dassa, Maryse Deguergue, Hubert de Verneuil, Robert Drillien, Anne Dubart-Kupperchmitt, Nicolas Ferry, Claudine Franche, Philippe Guerche, Joël Guillemain, Mireille Jacquemond, André Jestin, Bernard Klonjkowski, Marc Lavielle, Jane Lecomte, Olivier Le Gall, Yvon Le Maho, Stéphane Lemarié, Didier Lereclus, Rémy Maximilien, Antoine Messéan, Bertrand Ney, Jacques Pagès, Daniel Parzy, Catherine Regnault-Roger, Pierre Rougé, Patrick Saindrenan, Pascal Simonet, Virginie Tournay, Bernard Vaissière, Jean-Luc Vilotte.

## Annexe 2 : Compilation des commentaires transmis à l'AESA en anglais

Cette annexe est une compilation des commentaires transmis à l'AESA en anglais. Le format de ces commentaires a été adapté pour une transmission sur le site de l'AESA par l'intermédiaire de formulaires distincts par section.

### 1. General comments

For the sake of conciseness and to complement the analysis of Anses (the French Agency for Food, Environmental and Occupational Health & Safety) on the toxicity aspects of the document, only those comments of the High Council for biotechnologies (HCB) concerning statistical analysis are submitted to EFSA.

HCB welcomes the decision of EFSA to publish a guidance document on 90-day sub-chronic oral toxicity study on whole food/feed in rodents. Indeed, some of the methodologies used so far by the notifiers were deficient, which did not always permit to validate the conclusions formulated in the dossiers submitted for evaluation.

HCB considers that several important methodological aspects are now correctly dealt with in this guidance document:

- Only equivalence tests permit to conclude on a possible equivalence between the genetically modified plant and its comparator.
- The effect sizes, rather than just the p-values, must be carefully analyzed. Indeed, statistically significant differences do not necessarily indicate biologically relevant effects, and vice versa.
- The test multiplicity should be correctly taken into account, using tools such as FDR (False Discovery Rate) approach.
- The choice of sample sizes must be determined by a power analysis of the statistical tests implemented.

### 2. Comments per section

#### APPENDICES

#### APPENDIX 1 – STATISTICAL PRINCIPLES AND GOOD EXPERIMENTAL DESIGN

#### 4. Considerations when designing an experiment

#### 4.2. High power

#### 4.2.4. Increasing the significance level

p. 30, l. 1159-1165 :

*“By convention most investigators use a 5 % significance level. If it were to be increased to 10 % any real effects which just failed to reach significance at the 5 % level would now be judged “significant”. However, this will also increase the number of false positive results (Type I errors). As toxicity tests usually involve many outcomes, there is in any case a problem with an excess of false positive results. For the assessment, strong emphasis should be placed on the biological relevance of any observed differences whether or not they reach the chosen level of statistical significance. This is best done by looking at the point and interval (e.g. confidence) estimates and not by just focussing on the P-value.”*

Paragraph 4.2.4 mentions the risk of detecting many false-positives by increasing the level of the test. This is true in a theoretical framework, where there is absolutely no difference for several parameters (the negatives) and differences in others (the positives) and where all these parameters are independent. Here, the problem does not arise exactly in these terms: all parameters can be positive (there is no reason to believe that a difference in diet produces

absolutely no effect) but it is a matter of distinguishing the differences that are harmless to health (the "negatives") from those showing a significant biological effect (the "positives"). Moreover, such an analysis requires a global approach, that is to say, a simultaneous review of all effects.

## 6. Statistical analysis

p. 34, l. 1359-1367 :

*"The separate analysis of many parameters, most of which are not expected to differ between treatment groups, may result in a large number of statistical tests. In order to control the number of false positive results the use of false discovery rate (FDR) methods have been suggested (Kall, 2008), although their use in the analysis of toxicity tests is not well established. The FDR is the estimated proportion of false positives among all the significant hypotheses tested. However, this technique is not applicable in experiments where there are no strongly positive responses. Should there be no real differences between the groups being compared across many parameters, then all the positive results will be false positives and the FDR will be 100 %. Therefore, this method is only recommended when there are some strong and statistically highly significant differences between the groups."*

This paragraph (l. 1359-1367) about multiple tests and the use of the FDR correction is a little confusing. When there are no differences, it is expected that a proportion  $\alpha$  of the non-corrected p-values be less than  $\alpha$  (i.e., a proportion  $\alpha$  of false positives is expected). On the contrary, in the same situation where there are no differences, the probability for at least one corrected p-value to be less than  $\alpha$  is precisely  $\alpha$  (it is therefore expected not to detect false positives with probability  $1-\alpha$ ).