



## RIVM-RIKILT FRONT OFFICE VOEDSELVEILIGHEID

### **BEOORDELING INZAKE ARTIKEL CARMAN et al. (2013) in JOURNAL OF ORGANIC SYSTEMS**

#### **Onderwerp**

In het algemeen overleg (AO) landbouwraad van 19 juni 2013 is het Australische onderzoek van Carman et al. (2013) naar voren gebracht, waarin GG-mais en GG-soja gevoerd aan varkens zou leiden tot ernstige maagaandoeningen bij de varkens. De staatssecretaris van Economische Zaken (EZ) heeft in dit overleg met de Tweede Kamer (TK) de toezegging gedaan om het onderzoek te bestuderen, te kijken of vervolgonderzoek nodig is en zo ja, de TK te informeren over wie onderzoek gaat doen. Daarom is Bureau Risicobeoordeling (BuRO) door de directie Plantaardige Agroketens en Voedselkwaliteit (PAV) gevraagd om EZ op korte termijn te adviseren hoe de staatssecretaris dit onderzoek op waarde moet inschatten en welke mogelijke vervolgacties zinvol zijn en wie dat zou kunnen/moeten doen.

BuRO verzoekt het Front Office Voedselveiligheid om de studie te beoordelen.

#### **Vraagstelling**

Wat is het wetenschappelijk oordeel van RIVM-RIKILT over het onderzoek van Carman en medeauteurs (2013)?

De subvragen die BuRO beantwoord zou willen zien, zijn:

1. Voldoet de opzet van de proef aan wetenschappelijke standaarden?
2. Is de proefopzet voldoende discriminatief om uitspraken te kunnen over mogelijke nadelige effecten ten gevolge van
  - a. de insertie van het epsps gen;
  - b. de Cry-eiwitten;
  - c. unintended effects en
  - d. de mate van stacking?
3. Zijn de conclusies voldoende onderbouwd? Met andere woorden zijn de verklaringen die onderzoekers geven ten aanzien van de oorzaak-effect relaties, zoals enerzijds hormoonontregeling door de herbicide resistente GG-soja als gevolg van de insertie en/of anderzijds ontstekingsreacties in het maagepitheel ten gevolge van Cry-eiwitten in het dieet?
4. Zijn de 4 maïsvarianten toegelaten in de EU voor diervoeder en als levensmiddel?
5. Hoe groot is de kans dat als deze GM gewassen alleen aan dieren worden gevoerd, dit ook nog effecten zal hebben op de mens als deze dieren worden geconsumeerd?
6. Geeft de studie aanleiding om de EFSA-risicobeoordelingsrichtsnoeren te herzien?
7. Geeft de studie aanleiding tot twijfel over de veiligheid voor mens en dier van deze varianten?
8. Klopt de bewering van de onderzoekers dat voor praktijkgebruik van GM-gewassen eerst lange termijn voedingsonderzoek bij dieren nodig is?
9. Welke mogelijke vervolgacties zouden zinvol zijn en wie zou dat kunnen/moeten doen?
10. Is deze publicatie bij de EFSA bekend, en wat is de visie/opinie van de autoriteit?

## Conclusie

1. De opzet van de proef voldoet niet aan gepubliceerde wetenschappelijke standaarden voor het uitvoeren van proefdieronderzoek met hele gewassen/voedingsmiddelen.
2. De proefopzet is onvoldoende discriminatief om uitspraken te kunnen doen over mogelijke nadelige effecten ten gevolge van
  - a) De insertie van het epsps-gen
  - b) De Cry-eiwitten
  - c) Unintended effects
  - d) De mate van stacking
3. De conclusies die getrokken worden in de studie van Carman et al. zijn onvoldoende onderbouwd. Dit betreft zowel de oorzaak-effect relaties (hogere uteri-gewichten en ontstekingsreacties in het maagepitheel ten gevolge van het GG-dieet), als de gemelde noodzaak tot chronisch onderzoek van GG-plantvariëteiten.
4. De 4 maisvarianten die gebruikt zijn in de studie zijn toegelaten in de EU voor gebruik in diervoeders en in voedingsmiddelen.
5. Er is geen indicatie dat de GG-variëteiten, die in Europa zijn getoetst op veiligheid voor mens, dier en milieu en veilig zijn bevonden, een indirect, schadelijk effect zouden kunnen hebben bij de mens na consumptie van afgeleide dierlijke producten. Dit aspect wordt impliciet meegenomen in de Europese veiligheidsbeoordeling.
6. De studie geeft geen aanleiding om de EFSA-risicobeoordelingsrichtsnoeren te herzien.
7. De studie geeft geen aanleiding tot twijfel over de veiligheid voor mens en dier van deze GG-variëteiten.
8. Op basis van deze studie is er geen aanleiding om de huidige EFSA-richtlijn hieromtrent te herzien, die aanvullende studies alleen noodzakelijk acht wanneer daar vanuit korter durende studies aanleiding toe zou zijn. De studie van Carman et al. geeft geen aanleiding tot herziening van de veiligheidsbeoordeling door EFSA, die aanvullende studies niet nodig achtte.
9. Op basis van de studie van Carman et al. zijn er vanuit wetenschappelijk perspectief geen nadere acties nodig of wenselijk
10. Voor zover ons bekend is de studie bij EFSA bekend en besproken, maar heeft EFSA nog geen mandaat om de studie nader te evalueren. De EFSA heeft vooralsnog ook geen noodzaak gezien om de studie ongevraagd nader te evalueren.

## Toelichting

### Vraag 1

Voor proefdierstudies met *hele voedingsmiddelen* is er een EFSA-richtlijn beschikbaar voor 90-dagenstudies in knaagdieren, maar niet in landbouwhuisdieren (EFSA, 2011). Dit is de enige gepubliceerde standaard op dit terrein. De opzet van de proef voldoet niet aan deze wetenschappelijke standaard. De principes die in deze EFSA-richtlijn worden beschreven zijn veelal niet toegepast in de studie door Carman et al.

Belangrijke afwijkingen zijn:

- De testdiëten zijn slechts beperkt geanalyseerd op biologische, chemische en microbiële contaminanten. De testdiëten zijn uitsluitend geanalyseerd op de aanwe

zigheid van de verwachte GGO-elementen in mais en soja (niet gekwantificeerd in het geval van mais in het GG-dieet) en op mycotoxinen.

- Zowel de GG-soja en GG-mais als de conventionele soja en mais voor de proefdierdiëten zijn afkomstig uit reguliere productiestromen en zijn dus niet onder vergelijkbare omstandigheden geproduceerd. Naast de te verwachten verschillen in eiwitsamenstelling als gevolg van de genetische modificaties zijn ook andere verschillen tussen GG-soja en GG-mais en de respectievelijke traditionele tegenhangers daarmee waarschijnlijk.
- Gegevens over de opslag van de proefdierdiëten gedurende de duur van de dierproef worden niet vermeld in de publicatie.
- Er is slechts één dosering toegepast, waarin meerdere GGO-variëteiten in sterk variërende concentraties voorkwamen (3.3% - 100%). Het is daarmee niet mogelijk om een dosis-effect relatie vast te stellen.
- De studie is niet onder Good Laboratory Practice (GLP) uitgevoerd.
- Er wordt geen gedetailleerde informatie gegeven ten aanzien van de omstandigheden waaronder de dieren gehouden zijn en de kwalificaties van het personeel, behalve dat het een commerciële Amerikaanse boerderij betreft.
- Er is geen vergelijking gemaakt van de bevindingen in de studie met historische controle-data.
- Een hypothese ontbreekt in de studie, evenals de indicatie of de studie 'confirmatory' of 'exploratory' is. De term 'observational' wordt gebruikt, maar niet in de context van een wetenschappelijke hypothese

NB De duur van de studie, 22,7 weken, komt niet overeen met internationaal geharmoniseerde protocollen voor gestandaardiseerde toxicologische testen. Om die reden is er geen vergelijking mogelijk met data uit eerdere studies.

Er is geen informatie gegeven over de voor mycotoxinen gebruikte analysemethoden en welke toxines daarin wel/niet werden bepaald. Met name zearalenon zou een belangrijke rol kunnen spelen bij de verschillen in uterusgewicht en zou om die reden expliciet genoemd moeten worden.

Ten slotte is de studie gepubliceerd in het tijdschrift 'Journal of Organic Systems', dat geen 'track record' heeft ten aanzien van toxicologische studies. Dit zou kunnen duiden op een beperkte 'peer review' in dit opzicht.

## Vraag 2

Door de gekozen opzet, waarbij GG-plantmateriaal uit reguliere diervoederproductiestromen is gebruikt in een ratio van 90% DK 42-88 RR YG PL (een stacked gene variety van NK603, MON863 en MON810), 3,3% Pannar 5<sup>E</sup>-900RR (NK603), 3,3% Pannar 4<sup>E</sup>-705RR/Bt (stacked gene variety van NK603 en MON810) en 3,3% Producers 5152RR (NK603) ten opzichte van vergelijkbaar (maar niet onder identieke omstandigheden geproduceerd) GGO-vrij materiaal waarin sporen GG-mais (0,4%, niet geïdentificeerd) en GG-soja (1,6%, niet geïdentificeerd) gevonden worden, is het niet mogelijk om een uitspraak te doen over mogelijke nadelige effecten van elk van de aanwezige GG-elementen (epsps-gen, de Cry-eiwitten) of ten aanzien van effecten van stacking.

Doordat de proefdierdiëten niet nader geanalyseerd zijn is het ook niet mogelijk om een uitspraak te doen ten aanzien van mogelijke 'unintended effects' in het algemeen: door de verschillende bronnen van variatie is het mogelijk en waarschijnlijk dat er statistisch relevante verschillen tussen de twee groepen worden gevonden. Hierbij is het in deze opzet niet mogelijk om onderscheid te maken tussen random verschillen en mogelijk toxicologische relevante verschillen. Echter, gezien het feit dat de vier onderliggende GG-variëteiten in goed opgezette studies uitgebreid zijn getoetst zonder dat er indicaties zijn gevonden voor mogelijk toxicologisch relevante onbedoelde neveneffecten van de genetische modificatie, maakt het aannemelijk dat de gevonden verschillen in deze studie random verschillen tussen de groepen betreffen.

### Vraag 3

De conclusie dat de uitkomsten van deze studie aanleiding moeten zijn voor langdurige proefdierstudies met GG plantvariëteiten is onvoldoende onderbouwd. Zie de antwoorden op vragen 1 en 2 voor de gesignaleerde zwaktes in de studie die het niet mogelijk maken om wetenschappelijk verantwoorde conclusies te trekken ten aanzien van mogelijke effecten van GG-diervoeder in het algemeen, dan wel met betrekking tot de gebruikte GG – variëteiten in deze studie. Daarnaast zijn ook de evaluatie en interpretatie van de verkregen data niet goed uitgevoerd.

Verder is het gevonden verschil in de gewichten van de uteri mogelijk te verklaren door de aanwezigheid in verschillende concentraties van mycotoxinen met estrogene werking in de diëten (slechts eenmalig bepaald in het dieet aan het begin van de studie en niet meer tijdens de studie zelf, zodat effecten van opslag niet meegenomen zijn in de analyse van de studie), en/of het feit dat de dieren op een leeftijd van 22,7 weken al in de puberteit kunnen zijn, waardoor verschillen in uterus-gewichten op kunnen treden. Belangrijk hierbij is dat het 'outbred' varkens betreft, die dus genetisch variëren en daarmee ook kunnen variëren ten aanzien van bijvoorbeeld de inzet van de puberteit. Ten slotte is de concentratie fyto-estrogenen in de verschillende soja-partijen ook niet bepaald. Verschillen hierin zouden de gevonden verschillen ook (mede) kunnen verklaren. Hierbij kan nog opgemerkt worden dat de range in de uterus-gewichten groter is in de non-GG groep in vergelijking met de groep dieren die GG-materiaal gevoederd kreeg.

Het vermelde verschil in maagontstekingen is, ook gezien de verhoudingen in 'matige' ontstekingen (die meer voorkomen in de dieren die het conventionele dieet kregen), mogelijk een random effect (zie Bijlage). Maagontsteking, en met name de mate daarvan kan men niet diagnosticeren uitsluitend op basis van een kleurverschil, en zeker niet na 5 maanden. Het is slechts een indicatie om hier nader sampling te doen voor histopathologisch onderzoek. Dit is niet gebeurd. Bovendien past roodheid (hyperaemie) meer bij een acuut effect dan bij een chronische blootstelling. Daarnaast is te verwachten dat bij een chronische maagontsteking er problemen met de voedselopname/vertering optreden. Dit is niet waargenomen bij parameters als groei en voederconversie die meestal zeer gevoelig zijn op dit punt. Het gehanteerde classificatie-systeem is eveneens niet helder beschreven.

Als de waargenomen effecten op de maag al relevant zouden zijn, dan speelt de vraag of de gevonden gehalten fumonisinen en de verschillen tussen de twee diëten een rol kunnen hebben gespeeld. Varkens zijn relatief gevoelig voor deze toxines, hetgeen ook tot uiting komt in de relatief lage Guidance value van 5 ppm (2006/576/EC). In het GM voer werd 3 ppm gemeten, in het non-GM-voer 1,2 ppm, waarbij onduidelijk is of de kwantificering correct is uitgevoerd. Er zijn in de literatuur meldingen te vinden van effecten van fumonisinen op de maag bij varkens, zij het met hogere doseringen, maar wel kortere tijdsduur (EFSA, 2005). Het lijkt op zijn minst onhandig om met dusdanig besmette voeders een dierproef uit te voeren.

### Vraag 4

De 4 maïsvarianten zijn toegelaten in de EU voor diervoeder en als levensmiddel.

### Vraag 5

Binnen Europa (als ook in de Verenigde Staten) worden GG-variëteiten in alle gevallen voor zowel toepassingen in diervoeders als in levensmiddelen beoordeeld. Er is geen reden om aan te nemen dat GG-variëteiten die veilig zijn bevonden voor gebruik in diervoeders indirecte negatieve effecten voor de mens zouden kunnen hebben door de consumptie van de afgeleide dierlijke producten. Dit wordt impliciet meegenomen in de evaluatie van gebruik van de GG-variëteit in diervoeder.

### **Vraag 6**

De studie geeft geen aanleiding om de EFSA-risicobeoordelingsrichtsnoeren te herzien.

### **Vraag 7**

Er is geen aanleiding om te twijfelen over de veiligheid van de gebruikte GG-varianten. De GG-variëteiten die in deze studie zijn gebruikt zijn uitgebreid op veiligheid getoetst binnen het Europese systeem van veiligheidsbeoordeling van GG-variëteiten voorafgaand aan marktintroductie. Het feit dat in deze Europese evaluatie geen enkele indicatie is gevonden van mogelijke onbedoelde negatieve (bij-)effecten van de genetische modificatie, geeft deze in veel opzichten gemankeerde studie geen aanleiding om te twijfelen over de veiligheid van de gebruikte GG-variëteiten.

### **Vraag 8**

Op basis van deze studie is er geen aanleiding om de huidige EFSA-richtlijn hieromtrent te herzien, die aanvullende studies alleen noodzakelijk acht wanneer daar vanuit korter durende studies aanleiding toe zou zijn. De studie van Carman et al. geeft geen aanleiding tot herziening van de veiligheidsbeoordeling door EFSA van de individuele GG-variëteiten, die aanvullende studies niet nodig achtte.

### **Vraag 9**

Op basis van de studie van Carman et al. zijn er vanuit wetenschappelijk perspectief geen nadere acties nodig of wenselijk.

### **Vraag 10**

Voor zover ons bekend is de studie bij EFSA bekend en besproken, maar heeft EFSA nog geen mandaat om de studie nader te evalueren. De EFSA heeft vooralsnog ook geen noodzaak gezien om de studie ongevraagd nader te evalueren.

### **Referenties:**

Carman JA, Vlieger HR, Ver Steeg LJ, Sneller VE, Robinson GW, Clinch-Jones CA, Haynes JI and JW Edwards (2013). A long-term toxicology study on pigs fed a combined genetically modified (GM) soy and GM maize diet. *Journal of Organic Systems* 8 (1): 38-54. Open access full text: <http://www.organic-systems.org/journal/81/8106.pdf>

EFSA Scientific Committee; 2011. EFSA guidance on conducting repeated-dose 90-day oral toxicity study in rodents on whole food/feed. *EFSA Journal* 2011;9(12):2438 [21 pp.] doi: 10.2903/j.efsa.2011.2438. Available online: [www.efsa.europa.eu/efsajournal](http://www.efsa.europa.eu/efsajournal)

EFSA (2005) Opinion of the Scientific Panel on Contaminants in Food Chain on a request from the Commission related to fumonisins as undesirable substances in animal feed. *The EFSA Journal* 235, 1 – 32.

### **Afkortingen:**

GM = Genetically Modified  
GGO = Genetische Gemodificeerd Organisme  
GG = Genetisch Gemodificeerd

### **Colofon:**

Risicobeoordeling aangevraagd door: NVWA-BuRO  
Risicobeoordeling opgesteld door: RIKILT en RIVM  
Datum aanvraag: 21-06-2013  
Datum risicobeoordeling: 27-06-2013 (*definitieve versie*)  
Projectnummers: V/09013/01/AA en V/100181/01/AA

## Bijlage Statistische analyse

### *Statistical analysis of stomach lesions*

The stomach inflammation is tested for each severity category separately. Although this approach is not unusual, it is inappropriate, since the severity categories reflect degrees in the same biological parameter. The observations by Carman et al. illustrate why this is an inappropriate. While the frequency of severe lesions is higher in the treated group, the frequency of moderate lesions is higher in the control group (and practically significant). It should be noted that the borderline between moderate and severe is not clear-cut, and part of the observed severe lesions might be close to moderate and vice versa. As the authors state: "For a severe level of inflammation, almost the whole fundus had to be swollen and cherry-red in colour". The criterion "almost the whole fundus" clearly shows that the borderline between moderate and severe is not very strict, so that misclassification can easily occur.

There is a statistical approach ("latent variable model") that can test ordinal data treating the inflammation as a single parameter while taking the observed severities into account. In this way misclassifications in severity categories is taken into account as well (as "measurement error" ). This test is not in the standard statistical textbooks, and will be briefly discussed.

Brief description of the latent variable model.

The inflammation may be regarded as a gradually increasing phenomenon. One may imagine that even within a given severity category animals differ in severity, and that each animal could in theory be associated with its individual severity level with some quantitative value. However, the individual severity cannot be quantified, and instead the inspector uses categories. So, each category in fact reflects a range of individual severities, separated by certain cut-off points in the minds of the inspectors. The statistical test assumes that the individual severities follow a normal distribution, with unknown cut-off points between severity categories. This normal distribution can be fitted to the observed frequencies of severity categories, where the cut-off points are unknown parameters to be estimated. Next, the likelihood-ratio test can be used to test the hypothesis that the mean of the distributions differ between treatments.

This latent variable model may be applied to the observed frequencies of stomach lesions. Fitting the model assuming that the distribution differs between treatments results in:

```
b- :    -0.2124
th-1 :    1.283
th-2 :   -1.2691
th-3 :   -0.89284
log-likelihood: -183.84
```

where b is an estimate of the difference in means of the latent variable model, and th1, th2, th3 are the estimated cut-off points.

Fitting the model assuming the same distribution for both treatments results in:

```
th-1 :    1.3868
th-2 :   -1.2741
th-3 :   -0.88247
log-likelihood: -184.54
```

Assuming two distributions results in a higher log-likelihood ("better fit") than when assuming the same distribution (because there is one more parameter in the model), but the increase is non-significant (likelihood ratio test;  $P = 0.24$ ).

**Conclusion:** There is no significant difference in stomach inflammation.

### *Effects on the uterus*

Overall, this study results in only one significant effect: uterus weight. However, it needs to be taken into account that in testing a range of parameters, there could easily be a significant outcome just by chance. Therefore, this effect would need to be confirmed in a new study. If so, it is recommendable to create more subgroups (pens) per treatment group, to avoid possible artifacts in the results due to differences among pens.

Note: The uterus weight data show the largest coefficient of variation (CV = 48% in non-GM, CV = 44% in GM) from all organ weights. Data with a CV that large are skewed to the right, i.e. they are not normal but lognormal. Therefore, the procedure of removing outliers first and then check for normality, which is inappropriate as indicated above, might have caused the significant outcome. In other words, the significance might be an artifact due to inappropriate treatment of the data. The (individual) observations would need to be re-analysed to check this.