Statistical analysis of comparative data on composition and agronomic characteristics: new software tool and recurring issues identified

C. Paoletti*, J. Perry† and H. Broll*
* EFSA GMO Unit; † EFSA GMO Panel

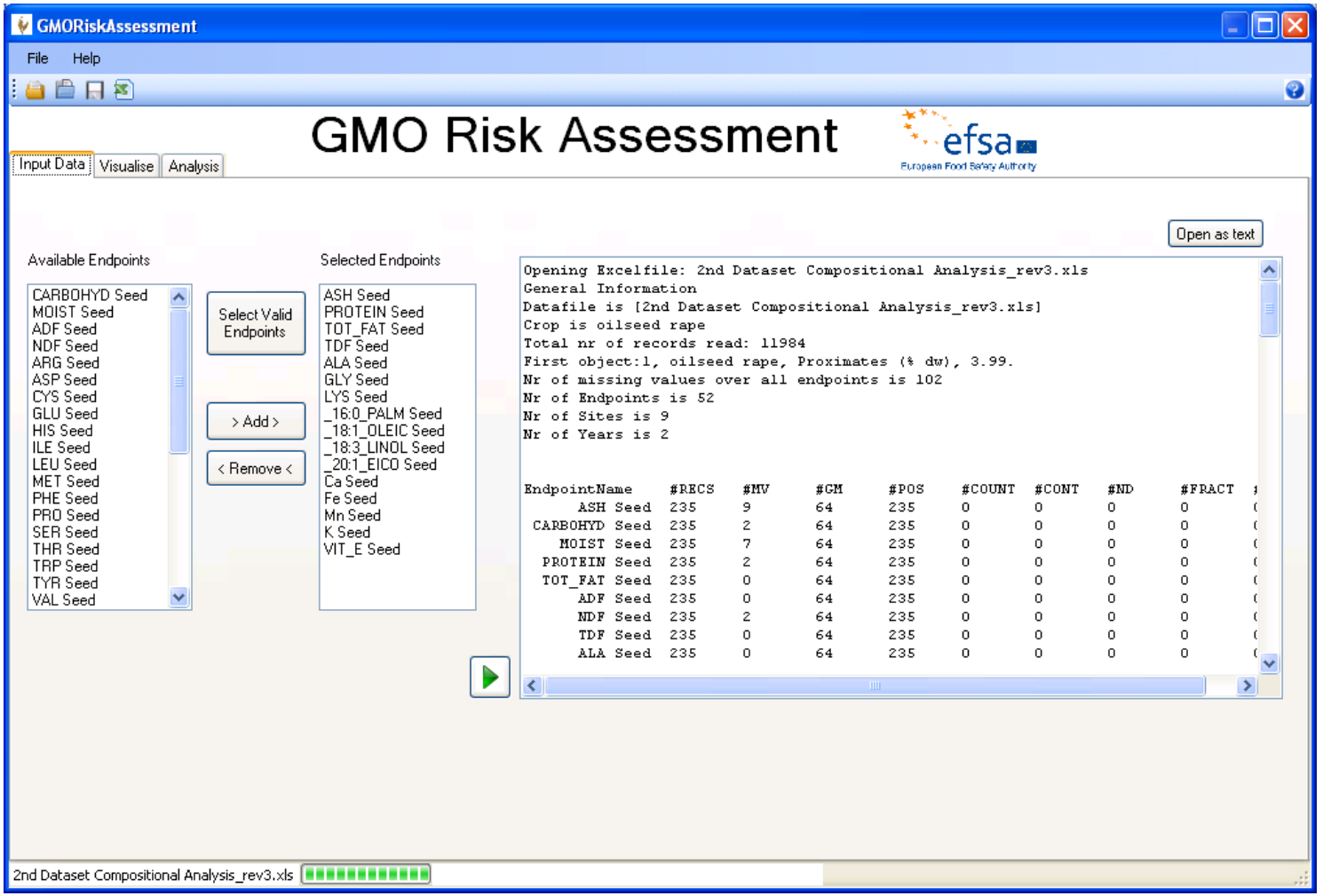
EFSA info session on applications - GMO
Parma, Italy
28 October 2014

GMO software tool – scope of the project

Development of a software for the implementation of the EFSA Guidance for risk assessment of food and feed from GM plants (EFSA, 2011)

- **Tool for stakeholders to comply with EFSA GMO guidelines**
- **Harmonization of statistical analyses for GMO dossiers**
- **Harmonization of reporting outcome of the analyses**

GMO software tool – input data



GMO software tool – access to the software

http://www.efsa.europa.eu/en/gmo/gmoanalysissoftware.htm

GMO software tool – installation requirements

GMO Analysis software for comparative assessment

In the European Union, the use of genetically modified organisms (GMOs) is regulated through a legal framework. According to EU legislation (Regulation (EC) No 1829/2003, Regulation (EU) 503/2013 and Directive 2001/18/EC), GMOs can only be authorized for placing on the EU market following a scientific assessment of any risks that they may pose to human and animal health and the environment.

An integral part of such scientific assessment is the evaluation of field trials performed for the comparative assessment of the compositional, agronomic and phenotypic characteristics of GM plants. In 2009, EFSA's GMO Panel adopted an opinion on Statistical considerations for the safety evaluation of GMOs to provide detailed guidance on the performance of such field trials and their statistical analysis. The principles, concepts and data requirements presented in the above mentioned opinion are endorsed in the GMO Panel's Guidance for risk assessment of food and feed from genetically modified plants and the more recent implementing Regulation (EU) 503/2013.

To provide further practical assistance for the analysis of comparative data for the risk assessment of GM plants, EFSA, in collaboration with Wageningen UR, has developed a software tool to perform the statistical analysis along with EFSA GMO Panel guidelines and Regulation (EU) 503/2013.

The software allows data entry and data analysis using the latest statistical procedures developed and adopted by EFSA's GMO Panel for the risk assessment of GM plants. The software conducts a simultaneous analysis of the GM plant compared to its control (test of difference) and non-GM reference varieties (test of equivalence). Moreover different factors can be introduced and analysed.

An output-file listing all the significant differences and the respective equivalence categories is generated by a single click. Additionally, some settings can be changed, according to the specific needs of the analysis being performed.

The software is user-friendly and data can be imported from different sources (e.g. Excel).

Technical requirements

- Windows
- Microsoft .NET 4 client framework
- R, version 2.15 or higher (http://cran.rstudio.org).

Please refer to the Installation Manual (see below) for detailed instructions on how to set up this software:

- GMO Analysis software: installation manual (0.4 Mb)

The software does not require any license for its installation and utilization by any user.

- GMO Analysis software: installer (11.3 Mb)
- GMO Analysis software: user manual (0.9 Mb)

EFSA welcomes comments which could support improvement of this GMO software tool. Please send your comments at GMO.software@efsa.europa.eu.

Since 2 October available and free of charge

Statistical evaluation of the field trials

- Contract with Hasselt University, Belgium
- Only field trial studies according to the latest EFSA Guidance, 2011
- Dedicated time frame given for the evaluation (1 month)
- Two deliverables were requested from the contractor on each dossier:
  - 1st deliverable: Evaluation report on the field trials highlighting issues for EFSA consideration (preparatory work)
  - 2nd deliverable: Support the evaluation of the applicant's reply to EFSA questions

Experimental design

- No information on randomization and blinding
  - Requested in guidance but not always provided
- Exclusions
  - Site exclusions: e.g. 10 included but only 8 analyzed: no proper explanation
  - Sample exclusions: often not justified
  - Can have implication on the minimum number of replicates

Statistical analysis

- No statistical analysis plan (SAP) provided
  - In line with good statistical practice
- Execution of R and SAS programs
  - Re-analysis can not be carried out due to lack complete directory structure
- Values below Limit of Detection
  - Substitution of ½ the detection limit
  - Statistically, not the most appropriate way to handle non-detects, especially when frequency is medium high (e.g. around 30%)

Statistical analysis

- Discrepancies found between the statistical outcome of the applicant and the re-analysis done by the Hasselt University
  - Different outcome types identified
- No discussion of missing data mechanism (MCAR, MAR, MNAR)
  - In line with good statistical practice
- When no equivalence limits can be established, often biological relevance is not discussed

Biological relevance – example of an insufficient discussion

All observed differences in this statistical analysis conducted under EFSA guidelines were characterized by a lack of compositional relevance from a food or feed perspective.

This lack of relevance from a food or feed perspective is characterized by the small magnitudes of difference between the GM event (treated or not treated) and the control.

Biological relevance – expectations

- The applicant should report and discuss all significant differences observed between the GM plant, its comparator and, where applicable, any other test material, focusing on their biological relevance
- OECD listed endpoints are the basis. They serve as indicators for other parameters not measured but with safety relevance
- Consideration of the specific metabolic function/pathway

Statistical analysis - Verification of assumptions

- Diagnostic Checks
  - Assumption of normality of data and homogeneity of variance should always be checked
    - Sometimes, only graphical ➔ can be subjective
    - Some form of studentized/standardized residual used, but no proper justification is provided for the specific cut-off that is used
    - When outliers found, analysis with and without outliers not always done
    - Goodness-of-fit checks rarely done
    - Linear mixed models have completely different theoretical implications on diagnostic checks than standard (fixed-effects) regression analysis.
    - Conditional vs. marginal residuals
    - Tests on normality, homogeneity of variance
    - Box-Cox procedure

EFSA GMO Unit

gmo@efsa.europa.eu

Thank you for your attention!