EXTERNAL SCIENTIFIC REPORT

APPROVED: 30 November 2015 PUBLISHED: 27 January 2016

MCRA made scalable

for large cumulative assessment groups

Hilko van der Voet, Waldo J. de Boer, Johannes W. Kruisselbrink (Biometris, Wageningen UR), Gerda van Donkersgoed, Jacob D. van Klaveren (RIVM)

Abstract

A scalable version of the Monte Carlo Risk Assessment (MCRA) software was developed and implemented in MCRA 8.1, available at https://mcra.rivm.nl, where also the MCRA Reference manual can be found (MCRA 2015).

Scalability was addressed in two ways: 1) by improving the computational infrastructure, and 2) by adding specialised algorithms.

First, MCRA 8.1. has been implemented in the flexible environment for high-performance computing at RIVM, allowing the use of a flexible number (currently, 10) simulation worker services to address simultaneously submitted jobs in parallel.

Second, an algorithmic approach was developed to handle large Cumulative Assessment Groups (CAGs). Two unique features of MCRA are: (1) contributions to the exposure results can be seen both in terms of food-as-eaten (e.g. white bread) and foods-as-measured (e.g. wheat), and (2) a drill-down can be made into the exact foods and compounds contributing for simulated individuals or individual-days in the upper tail. The number of combinations of simulation, compound, food-as-measured and food-as-eaten can be very large. To avoid memory problems with very large datasets, an additional optional modelling step, named Screening, was added to MCRA 8.1. Screening should be used if the data dimensions are too large for a direct analysis, i.e. when a direct analysis would run into computational problems due to insufficient memory availability. Screening identifies risk drivers. A full analysis based on screened risk drivers will still retain all food/compound combinations in the exposure calculation, and will therefore produce exactly the same cumulative exposure distribution, and allow to see contributions of all compounds and all foods-as-measured. Details with respect to foods-as-eaten are however restricted to the risk drivers selected in the screening step.

Using test data provided by EFSA it was shown that screening allowed to run a large CAG of 96 compounds. The screening approach was tested for the optimistic and pessimistic models as defined in the EFSA guidance for probabilistic exposure assessments (EFSA 2012). In terms of identifying risk drivers screening performed well when using the EFSA optimistic model, but not so good for assessments using the EFSA pessimistic model. The reason is that in the EFSA pessimistic model all non-detects are imputed with a positive value (the limit of reporting), leading to an immense and unrealistic number of food/compound combinations that contribute to the estimated exposure. Further work on the assumptions used in the pessimistic model is foreseen in a European working group.

Using test data provided by EFSA the feasibility of the approach was validated. A set of 40 assessments, consisting of 20 acute assessments for a CAG of 65 compounds plus 20 chronic assessments for a CAG of 127 compounds, all with an uncertainty analysis based on 100 resampled

datasets, was completed in the MCRA computational environment within a period of 6 hours, with a maximum memory usage of around 4GB.

In conclusion, MCRA 8.1 will handle datasets with up to at least 100 compounds in a cumulative assessment group (CAG) and at least four million concentration records reported with the EFSA Standard Sample Description (SSD) model. With or without screening MCRA 8.1 will produce the same estimated cumulative exposure distribution summarized by percentiles and exceedance percentages, and allow to see contributions of all compounds and all foods-as-measured. After screening, contributions related to food-as-eaten are available for the risk drivers.

An implementation report of this specific action was delivered to EFSA.

Key words: combined exposure, cumulative exposure groups, probabilistic, Monte Carlo Risk Assessment, scalability, risk drivers

Question number: EFSA-Q-2014-00622

Correspondence: pras.secretariat@efsa.europa.eu

Contractor:
Dutch National Institute for Public Health and the Environment (RIVM), with subcontractor Wageningen University and Research centre (WUR)

Contract:
FRAMEWORK PARTNERSHIP AGREEMENT 'Support to the Regulatory Implementation of Cumulative Risk Assessment of Pesticides'

Grant number: GP/EFSA/PRAS/2014/02

Specific agreement number 1: Scalability and data organisation

Acknowledgements: The help of EFSA staff members Paula Medina, Davide Arcella and Luc Mohimont is greatly appreciated.

Suggested citation: van der Voet H, de Boer WJ, Kruisselbrink JW, van Donkersgoed G, van Klaveren JD, 2016. MCRA made scalable for large cumulative assessment groups. EFSA supporting publication 2016:EN-910. 38 pp.