

This is a draft paper for discussion. It should not be quoted, cited or reproduced.

CC/07/17

**COMMITTEE ON CARCINOGENICITY OF CHEMICALS IN FOOD,
CONSUMER PRODUCTS AND THE ENVIRONMENT**

Papers on QSAR predictive models for carcinogens

1. Two papers have recently been published on predictive toxicology using QSAR models. The first paper, by Valerio et al (2007) from the US FDA, is attached. This paper describes an evaluation of the performance of a particular software, MDL-QSAR predictive discriminant analysis modeling of rodent carcinogenicity, to estimate the carcinogenic potential of small, organic, naturally occurring chemicals found in the human diet. The predictive performance for this group of chemicals and a control group of 19 known synthetic dietary constituents was 97% sensitivity and 53% specificity.

2. The second paper, by Benigni et al (2007), describes and discusses preliminary findings from a project on (Q)SARs for mutagenicity and carcinogenicity initiated by the European Chemicals Bureau and carried out by the Istituto Superiore di Sanita', in the context of the forthcoming REACH legislation. It is being carried out in two parts: Phase I consists of collection of models from the literature, and a preliminary evaluation based largely on information reported by authors of the models; the most promising models are then being identified in Phase II, where they can be further characterised, including by assessment of predictive ability using independent test sets. Phase II is still in progress and the paper reports the results of Phase I and some results from Phase II. The project is only evaluating non-commercial models of direct relevance; commercial models excluded.

3. The initial phase of the evaluation showed that the models cover a large variety of approaches and techniques e.g. congeneric and non-congeneric series modeling, use of structural alerts as well as more quantitative chemical descriptors, and use of interpretable physico-chemical parameters, as well as descriptors selected by purely statistical means. Conclusions drawn by the authors so far include:

- any attempt to draw conclusions about the behaviour of certain compounds from the behaviour of other (similar) compounds provides only a model that reflects reality to some, largely unknown extent. The key problem is that there is no generally valid definition of similarity, so it is difficult to assess the appropriate congenericity of a data set.
- predictivity of carcinogenesis is poorer than for mutagenesis, where fewer mechanisms are involved.
- the OECD principles for (Q)SAR validation provide the best currently available harmonized framework to assist in judging suitability of approaches for regulatory use. A number of available (Q)SARs for carcinogenicity comply with these principles to different extents, the main problem being missing information on the domain of applicability.
- different methods of model predictivity exist and they have advantages and limitations that preclude a uniform preference.

This is a draft paper for discussion. It should not be quoted, cited or reproduced.

- the derivation and evaluation of (Q)SAR models should not only depend on statistical criteria, but should also take into account mechanistic aspects and interpretability.

4. Do Members consider that it would be useful to embark on a detailed review of QSAR for prediction of carcinogenicity, either in 2008 or when the full results of the Istituto Superiore de Sanita' project are published?

Secretariat
November 2007

References

Benigni R, Netzeva TI, Benfenati E, Bossa C, Franke R, Helma C, Hulzebos E, Marchant C, Richard A, Woo Yin-Tak and Yang C (2007). The expanding role of predictive toxicology: An update on the (Q)SAR models for mutagens and carcinogens. *J of Envir Science and Health, Part C* 25: 53-97.

Valerio LG, Arvidson KB, Chanderbhan RF and Contrera JF (2007). Prediction of rodent carcinogenic potential of naturally occurring chemicals in the human diet using high-throughput QSAR predictive modelling. *Tox Appl Pharmacol* 222: 1-16.

Note: The appendix is not being made publicly available because the references which it contains cannot be distributed for copyright reasons.