UNCERTAINTY AS A PARAMETER TO COMPARE AND INTEGRATE DATA OF HETEROGENEOUS SOURCES

Emilio Benfenati[1], Elena Boriani[1], Theo Vermeire[2] , Tom Aldemberg[2] , Alberto Fernandez[3], Alberto Manganaro[1]

[1] Mario Negri Institute - Via Giuseppe La Masa 19 - 20156 Milano - Italy
[2] National Institute for Public Health & Environment - 3720 BA Bilthoven - The Netherlands
[3] Universitat Rovira i Virgili - Carrer de l'Escorxador - Tarragona - Spain

INTRODUCTION

The efforts of new legislations such as REACH are aimed to reduce the impact on human health and the environment of chemicals and to promote risk assessment analysis for the evaluation of the health status of the environment. Due to the present lack of data about chemicals in Europe and to the limitations of assays, REACH supports a full use of all type of data (*in vivo*, *in vitro*, *in silico*). REACH promotes innovation (with the enhancement of non-testing methods), in order to cover as much as possible the knowledge gap that currently exists in the assessment of the safety of chemicals and to avoid repetition in terms of test and animal use.

It now becomes necessary to **assess how heterogeneous data can be managed within a unified approach**, suitable for the risk characterisation of the chemicals.

Uncertainty is a typical characteristic of all data, of any nature. This refers to all values for the exposure and effect related risk assessment. Risk refinement tool, as developed by Verdonck at al.[1], are useful to reduce the uncertainty related to the conservatism of worst-case assumptions, and the uncertainty of using assessment factors. This requires defining and modulating the uncertainty when data comes from one or more sources, as experimental, *in vitro* and *in silico*.

OECD and ECHA recently produced guidelines addressing the **issue of reliability and uncertainty for exposure models**, for *in silico* models and also for *in vivo* studies[2,3]. Internal organization, such as ECVAM in Europe, are devoted to assess the validity of alternative methods; at present most of the evaluation of the data uncertainty has been done for the *in vitro* studies. Most QSAR studies have produced models described in their fitness properties; more recently the *in silico* methods have been studied as tools to predict toxicity properties of chemical compounds for regulatory purposes, and some example appeared for instance for the specific use within the EC regulation for pesticides[4].

MATERIALS & METHODS

CAESAR EU project databases and models[5] for *Bioconcentration factor* (*BCF*), a quantitative model, and *Mutagenicity*, a qualitative model, are considered as examples.

BCF: The uncertainty of *in vivo* studies is about 0.5 in log units[6]. The standard deviation error in prediction for the predicted values of *in silico* methods is about 0.5-0.6 in log units; these values refer to the results of the CAESAR model on a set of compounds not used to build up the model, and somehow higher values have been obtained using the EPI Suite[7] model, but in this case the values may also include calculated values for chemicals which were present on the training set. See *fig. 1*.

Mutagenicity: the experimental reproducibility is about 85%, when the same experiment is carried out by different laboratories[8,9]. Within the project CAESAR the accuracy of the prediction was about 85%, on the basis of a large data set (more than 800 chemicals).

A possible way to assess the relevance of a method and conduct a comparison among methods is to consider their uncertainty, checked on the basis of specific tests. The uncertainty relative to the alternative method should be compared for the specific uncertainty of the experimental method. This gives a realistic expectation of the possibility to use a certain value, for the specific endpoint. The weight of the relevance of a method in the use for the final risk characterization may be expressed in relation to its uncertainty (*fig. 1* and *2*).

The above refers to the comparison of the performance of the models and methods, on the average, for a set of compounds. However, this general approach can be modulated on the basis of specific compounds. Indeed, uncertainty can be easily assessed at the level of chemical groups, or of individual compounds.

For instance, within CAESAR we measured a higher uncertainty for chemicals with some fragments and we introduced warnings in the predictions obtained with the CAESAR models, thus making possible to have a more precise evaluation of the uncertainty for a single compound, besides the evaluation of the general accuracy of the model (*fig. 3*).

Furthermore, at the individual level, the potential error can be evaluated with the similarity tool which is implemented in the CAESAR models (*fig. 4*).

These examples highlight a particular positive aspect of *in-silico* models, as they can give a precise measurement of the uncertainty associated with the prediction, thus allowing the regulator to have an explicit evaluation of the quality of data.

DISCUSSION & CONCLUSIONS

Advantages of the uncertainty comparison among methods:

- Different methods can be easily compared.
- A clear, objective procedure is defined which can be applied unambiguously.
- The scientific basis relies on fundamental criteria for scientific measurements, which have been incorporated into official guidelines for the evaluation of alternative methods, both for *in vitro* and *in silico* methods[10].
- Weighing factors of different methods is not fixed on the basis of subjective opinions based on performance of previous models; they may evolve with time as a result of new improved models.
- The approach may be also suitable to manage data from exposure models.

Acknowledgements

EU project 022674 SSPI - **CAESAR**
http:// www.caesar-project.eu

EU project 037017 GOCE - **OSIRIS**
http://www.osiris-reach.eu

CONTINUOUS VALUES

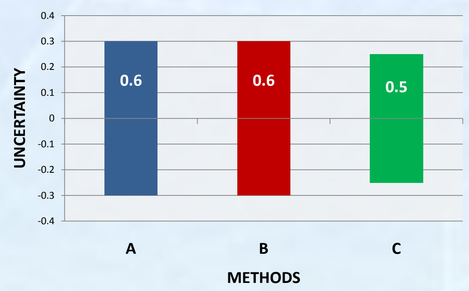

Figure 1
For *continuous values* (**BCF endpoint**) three methods are compared in their performances relatively to the measured uncertainty: A (***in vivo* model**), B (***in silico* EPI Suite model**), C (***in silico* CAESAR model**). If the uncertainty is similar the methods are similar (see *Materials & Methods*).

CATEGORY VALUES

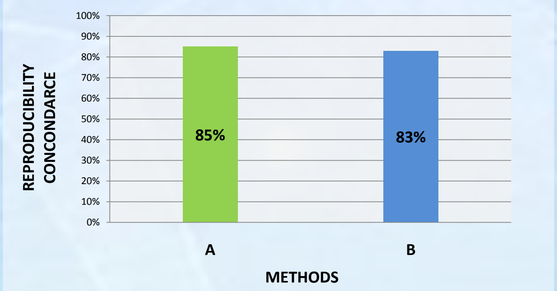

Figure 2
For *categorical methods* (**Mutagenicity endpoint**) the reproducibility as concordance is compared: A (***in vitro* test**), B (***in silico*** CAESAR model). If these values are similar, the two methods are similar. If one of the method is an *in-silico* model, its accuracy can be used to be compared with the concordance of the others.

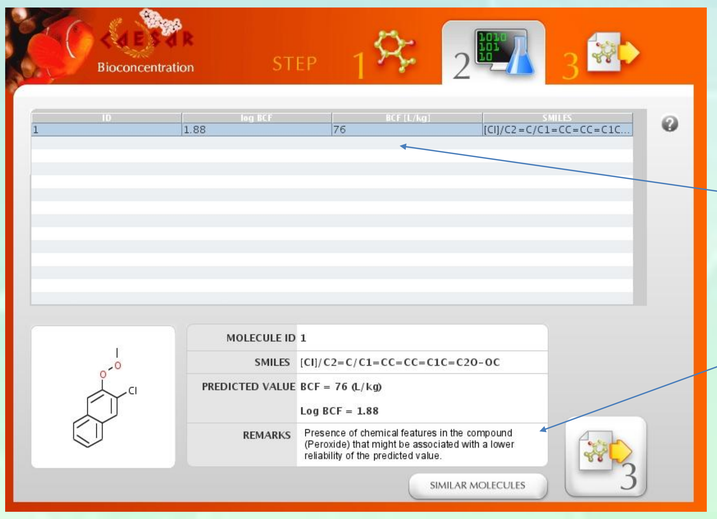

Figure 3

The result window of **CAESAR model for BCF** shows at the top all molecules submitted (as SMILES) with the predicted property value(s).

The presence of *eventual remark for critical compounds* is indicated, giving the user a warning about the uncertainty of the prediction on the selected compound.

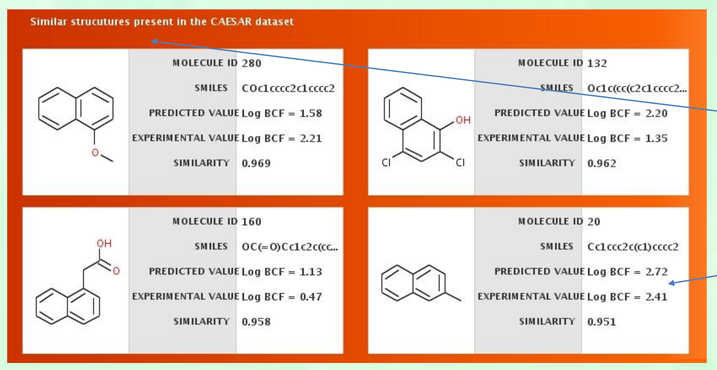

Figure 4
In addition to the prediction of the endpoint property, the server automatically retrieves *the most similar molecules to the molecule submitted*, found in the whole dataset used for the model building and testing.
The experimental and predicted values are shown, allowing measurements of the error for related compounds.

References

1. F. A. M. Verdonck, P. A. Van Sprang, P. A. Vanrollenghem. An intelligent data collection tool for chemical safety/risk assessment. Chemosphere, 70, 1818-1826.
2. OECD Guidelines for testing of chemicals. http://www.oecd.org/document/40/0,3343,en_2649_34377_37051368_1_1_1_1,00.html
3. Guidance on information requirements and chemical safety assessment, Part B: Hazard Assessment, updated October 2008, R19, updated May 2008.
4. E. Benfenati E. (Ed.), 2007. Quantitative Structure-Activity Relationship (QSAR) for Pesticide Regulatory Purposes. Amsterdam: Elsevier.
5. http://www.caesar-project.eu/software/index.htm
6. C. Zhao, E. Boriani, A. Chana, A. Roncaglioni, E. Benfenati, 2008. A New Hybrid QSAR Model for Predicting Bioconcentration Factor (BCF). Chemosphere, 73, 1701-1707.
7. Estimation Programme Interface (EPI) Suite. US EPA. http://www.epa.gov/opptintr/exposure/pubs/episuite.html
8. E. Benfenati, R. Benigni, D. M. DeMarini, C. Helma, D. Kirkland, T. M. Martin, P. Mazzatorta, G. Ouédraogo-Arras, A. M. Richard, B. Schilter, W. G. E. J. Schoonen, R. D. Snyder, C. Yang. Predictive models for carcinogenicity and mutagenicity: frameworks, state-of-the-art, and perspectives. Journal of Environmental Science and Health Part C, 27: 57-90, 2009.
9. J. Kazius, R. McGuire R, R. Bursi, 2005. Derivation and validation of toxicophores for mutagenicity prediction. L. Med. Chem 48 (1), 312-20.
10. A. Fernandez, R. Rallo, F. Giralt. 2009. Uncertainty Reduction in Environmental Data with Conflicting Information. Environmental Science & Technology. In press