Project no. 037017

OSIRIS

Optimized Strategies for Risk Assessment of Industrial Chemicals through Integration of Non-Test and Test Information

Integrated Project

Sub-Priority 1.1.6.3: Global Change and Ecosystems

Topic VII.1: Development of Advanced Methodologies for Risk Assessment
VII.1.1: Intelligent testing strategies for chemicals

D4.1.15
Summary of the Fourth Stakeholder Workshop

Due date of deliverable: 31 May 2011

Actual submission date: 31 May 2011

Start date of project: 1 April 2007 Duration: 4,5 years

Lead Contractor: DIALOGIK, Prof. Dr. Dr. h.c Ortwin Renn

Authors:
Prof. Dr. Dr. h.c. Ortwin Renn, Christina Benighaus, Ludger Benighaus, Katrin Alle, DIALOGIK, Stuttgart, Germany

with contributions from:
Emiel Rorije and Dr. Tom Aldenberg, National Institute of Public Health and the Environment (RIVM), the Netherlands
Dr. Sylvia Escher and Dr. Inga Tluczkiewicz, FhG
Dr. Monika Nendza, Analytisches Laboratorium, Luhnstedt, Germany
Anna Lombardo and Dr. Alessandra Roncaglioni, Instituto di Ricerche Farmacologiche "Mario Negri", Milan, Italy
Dr. Dinant Kroese, TNO, Zeist, the Netherlands

Project co-funded by the EU Commission within the Sixth Framework Programme (2002-2006)		
Dissemination Level		
PU	Public	X

Summary of the Fourth Stakeholder Workshop
Integrated Testing Strategies

Authors:

- Prof. Dr. Dr. h.c. Ortwin Renn, Ludger Benighaus, Christina Benighaus, Katrin Alle, DIALOGIK, Stuttgart, Germany

with contributions from:

- Emiel Rorije and Dr. Tom Aldenberg, National Institute of Public Health and the Environment (RIVM), the Netherlands
- Dr. Sylvia Escher and Dr. Inga Tluczkiewicz, FhG
- Dr. Monika Nendza, Analytisches Laboratorium, Luhnstedt, Germany
- Dr. Anna Lombardo and Dr. Alessandra Roncaglioni, Instituto di Ricerche Farmacologiche “Mario Negri”, Milan, Italy
- Dr. Dinant Kroese, TNO, Zeist, the Netherlands

Summary

The fourth and last Stakeholder Workshop was organised by the OSIRIS Partner DIALOGIK in close cooperation with the Helmholtz Centre for Environmental Research (UFZ) in March in 2011 in Leipzig, Germany.
The workshop aimed

- to present the current state of work regarding the development of the webtool and the underlying concepts,
- to explain five ITS endpoints, ITS building blocks and testing steps developed and implemented into the OSIRIS webtool,
- to collect stakeholder feedback on the ITS methodology development and the OSIRIS webtool.

In the workshop 56 key stakeholders from industry, regulatory authorities and academia registered in Leipzig. Most of them are experts from industry or regulation, fewer from NGO or scientific institutions (academia). Following the results of the general questionnaire which was distributed during the workshop most of them were familiar with the REACH context and ITS, but only a few of them already use ITS.
The workshop was divided into five different sessions according to the five ITS endpoints implemented in the OSIRIS webtool and its underlying concepts:

- *Skin Sensitisation*
- *Repeated Dose Toxicity*
- *Bioconcentration Factor (BCF)*
- *Aquatic Toxicity*
- *Mutagenicity and Carcinogenicity*

Every single session included background information on the ITS, discussion in plenum and the demonstration of the ITS webtool with concrete (key) examples followed by practical application and exercises. The participants tested the webtool for all particular endpoints during each workshop sessions. Feedback from stakeholders was conducted through several questionnaires during circles at the end of each workshop session. All together 128 questionnaires were completed and submitted during the workshop.

Webtool
Summing up, the majority of the respondents reported that the OSIRIS webtool is user-friendly and the navigation clear and easy. The results have been clearly documented, readily accessible, understandable and reasonable, but inexperienced users as non-toxicity experts could be overwhelmed by the different options. Helpful would be a general demo version which guides the user through the data processing and explains the structure of the webtool. Advanced users would like to use more short cuts in the navigation and a faster entry as inexperienced users should be able to skip yes/no-answers with alternatives such as "don’t know" and "fill in later".

Some respondents requested to improve the visibility of the workflow, since they have perceived the documentation as unclear and – in general - the information too limited. They see the need to give more background information on probabilistic calculations and to *"indicate the correlation between models"*. An additional tool to view and edit the structure would be helpful, too.
However, some respondents criticized, that the documentation *"wasn´t well structured"*, that is was *"insufficient"* due to a *"lack of information on how the probabilities were (actually) obtained"* and that *"basic assumptions were not displayed in the webtool"*.
A comprehensive output-report document for the assessment was suggested. It should be described which input-parameters are used and the underlying algorithm should be given.
A standard format as OECD, FoxMC e.g. is recommended. Missing features to import and export data and direct links to additional software as "Derek", "ToxTree", "Chemprop", prediction of proteins e.g. should be integrated.
The participants would use the OSIRIS webtool for strategy development and to get a better idea how information sources add statistical "weight of evidence". The OSIRIS webtool is seen more as a supporting instrument than a stand-alone version to avoid animal testing because it *"still relays more on human power rather as a technical tool"*.
Some participants noted that all the good ideas which are developed in the project should be *"preserved and further developed"*. *"Also some modules (Bayes algorithm, RepDose section) should be followed up upon."* Most of the participant explained that they got new insight in ITS and combining toxicity information. They are more confident that ITS will be more implemented in the future.

Endpoint "Skin Sensitisation"
Most of the respondents reported that the statistical approach (Bayesian decision theory) is helpful to "*justify on a scientific basis the choice of acceptability of a final result*". *"Especially if it makes more transparent, how conclusions were obtained and how reliable they are."* However, the use of Klimisch-like scores to "penalize" less reliable results is seen more ambiguous. Some of the respondents supported the score as a good guidance, but a lot of them reported that Klimisch-like scores are *"less practical"* and need a lot of additional guidance.
The majority of the respondents criticized that more in vitro tests or any experimental data should be added to get better "weight of evidence" for skin sensitisation toxicity. *"For the calculation of predicting of in silico methods, these methods were compared with LVNA ...".*

Endpoint "Repeated Dose Toxicity"
According to a large majority of the respondents the results of the assessment have been clearly documented. Some reported that there have been (minor) restrictions and that *"transparency on the detail level* (documentation of underlying assumptions) *could be improved"*. It was stated that it was *"too much a black box of calculation"*. Some of the respondents suggested that *"the judgment should vary on a case-by-case basis"*, between an expert judgment and criteria (QANTOS).

However, concerning the documentation the main aspect claimed was the improvement of the (stepwise!) process documentation.

Endpoint "Bioconcentration Factor (BCF)"

"The methods for in silico evaluation seem to be exhaustive". The respondents recommended to *"add a weight-of-evidence approach for "mixed" data situations"* and the results *"should be more weighted"*. Additionally, *"some more parameter are missing! (Density; molecular weight, ..)"* and should be integrated. The second suggestion was to use more results from in vitro data for WOE approach *("maybe more details on in vitro, how to use results from in vitro options, how to choose"*). 'Respondents would like to *"include Bioconcentration or Biomagnification of metabolism"* and receive *" recommendations for other tests, like a BMF test"*.

Most suggestions refer to the documentation of the assessment. In particular, *"more explanations (on functions) during the assessment workflow"* were expected. Some noted that the *"documentation should be more detailed"*. *"A clear and detailed documentation of the underlying algorithm should be given in the software documentation"* and that *"a description of the uncertainties should be included"*.

Endpoint "Aquatic Toxicity"

Based on the presentation in the workshop the following improvements were suggested: Again the respondents would like to be given more *"scientific background and guidance documents"* and a more detailed *"guidance about waiving, decisions and use of QSARS results and the information of the decision"*. Some of them reported that *"the module too strictly follows a decision tree..."* and should have an enhanced integration and more transparency *"on how validity of QSAR results/ACR prediction is assumed"*. The respondents suggested to focus more on in vitro data and include more *"biodegradation parameter like OECD 301XXX"*.

Endpoint "Mutagenicity & Carcinogenicity"

The predominant suggestion was to focus more *"on in vitro tests and the evolution of data from in vitro"*. A *"more automatic workflow"*, *"default settings whenever possible"*, as well as *"pre-selected values"* should be provided. The respondents suggested that *"the options COECDxy should be adjusted to the selected endpoints"*. Finally, the participants noted that a *"(brief) description of the test"* was missing.

Conclusion

The OSIRIS webtool with the five included endpoints is on the way to be a user-friendly, accessible and transparent webtool which helps the stakeholders to understand the testing strategies and supports the development of integrated and non-testing toxicity test. It is not a stand-alone software and can be combined with additional software as "Derek", "ToxTree" and "ChemProp". An integration of this software would be very practical for users in the future. As a consequence the import and export of data and their format should be standardized and use a standard format (OECD, FoxMC, ...).

Another important point of improvements in the next months should be a more clear and comprehensive documentation and report for each endpoint, the parameters used and the guideline from the EU regulation in the webtool.
The OSIRIS webtool will be finished in autumn 2011. It will be important to secure that the webtool and its huge knowledge behind that will be available for the stakeholders after the project is finished by then.

Contents

1. Introduction

The OSIRIS project aims at developing integrated testing strategies (ITS) fit for REACH that allow to significantly increase the use of non-testing information for regulatory decision making, and to minimise the need for animal testing. By optimally integrating various non-testing methods the goal is to improve efficiency, to reduce overall testing costs, to keep the ambitious time schedule of the REACH system and to improve public acceptance of ITS due to less animal testing.

ITS webtool

In order to make the results of the project available to the public, OSIRIS provided an ITS webtool which contained the following endpoints:

- skin sensitisation
- repeated dose toxicity
- bioconcentration factor
- aquatic toxicity
- mutagenicity & carcinogenicity

As uncertainty reasoning schemes the Bayesian Networks and Dempster-Shafer theory of evidence are integrated in the ITS OSIRIS webtool. Additionally, a Chemical Space Navigation Tool was implemented for prescreening the chemicals.

Stakeholder Workshops

To ensure an optimal implementation of the results achieved in the OSIRIS project and especially the OSIRIS webtool, end-users in industry, academic institutions and regulatory authorities (EU-stakeholders) have been invited to participate in the project by becoming involved in the monitoring of the project by providing specific technical guidance. A key component of the stakeholder involvement strategy is the organisation of four workshops along the basic research steps. A first workshop for setting the agenda and the priorities for research was conducted in 2007 in Stuttgart, Germany[1, 2] The second workshop was scheduled after one and a half years of research to present the preliminary results of selected pillars and to discuss them with the different stakeholders. This event took place in Brussels, Belgium in November 2008. The third workshop was scheduled on the 1st and 2nd of March 2010 in Berlin in cooperation with the German Federal Institute for Risk Assessment (BfR).

[1] The report of the first, second and third Expert Workshop (Report on the comparative review of stakeholder expectations, concerns, and proposals, including an assessment of the impacts of the results on the further structuring of the project) can be downloaded as Deliverables on the OSIRIS webpage (http://www.osiris.ufz.de/).

[2] Being part of the integrated EU-Project OSIRIS the workshop was funded by the Commission within the 6th Framework Programme under the theme "Global Change and Ecosystems". The project is coordinated by Prof. Dr. Gerrit Schüürmann at the Helmholtz Centre for Environmental Research - UFZ, Permoserstr. 15, 04318 Leipzig, Germany.

Major stakeholders participated to provide input to the OSIRIS team and to discuss the contributions of the OSIRIS research to the hazard assessment and risk assessment process for the two endpoints bio-concentration factor (BFC) and skin sensitization.
The fourth and last workshop was conducted on March 8th and 9th in 2011 in Leipzig, Germany. It was organised by the OSIRIS Partner DIALOGIK in close cooperation with the Helmholtz Centre for Environmental Research (UFZ). Key stakeholders and experts from industry, regulatory authorities and science were invited to test the methods and integrated testing strategies (ITS) for the five endpoints implemented in the OSIRIS webtool. The workshop aimed at assessing the practical significance of the OSIRIS webtool, the ITS building blocks and testing steps and collecting stakeholder feedback through questionnaires for each session featuring one particular endpoint.

2. Concept and Method of the Workshop

2.1 Target group

Key stakeholders and potential users of the ITS webtool have been invited from the following areas:

- Industry representatives (individual companies and sector groups), producers/importers
- National and European regulatory authorities such as European Chemicals Agency (ECHA), European Chemicals Bureau (ECB), Federal Institute for Risk Assessment (BfR) and other national agencies
- NGOs such as environmental groups, public health groups, consumer groups
- Experts from universities and research institutes
- Key internal and external OSIRIS partners of the consortium members
- Advisory board members
- Experts from related activities worldwide (OECD, US-EPA, Health Canada)

56 persons from the above mentioned areas registered for the workshop. Most of them are stakeholders from industry or regulation, some from NGO or scientific institutions (academia). As a result of the general questionnaire which was distributed during the workshop most of them were familiar with the REACH context and ITS, but only a few of them already use ITS.

2.2 Objectives and Topics

The fourth stakeholder workshop was aimed at continuing the dialogue between the project and EU stakeholders.

Main objectives

The main objectives of the fourth stakeholder workshop were

- to present the current state of work regarding the development of the webtool and the underlying concepts
- to explain five ITS endpoints, ITS building blocks and testing steps developed and implemented into the OSIRIS webtool
- to collect stakeholder feedback on the ITS methodology development and the OSIRIS webtool

Methods of the workshop

The workshop was divided into five different sessions according to the five following ITS endpoints implemented in the OSIRIS webtool and its underlying concepts:

- *Skin Sensitisation*
- *Repeated Dose Toxicity*
- *Bioconcentration Factor (BCF)*
- *Aquatic Toxicity*
- *Mutagenicity and Carcinogenicity*

All single sessions included background information on the ITS, discussion in plenum and the demonstration of the ITS webtool with concrete (key) examples followed by practical application and exercises. Before evaluating the OSIRIS webtool participants have tested the webtool for all particular endpoints during each workshop sessions. Feedback from stakeholders was conducted through several questionnaires at the end of each workshop session (compare chapter 2.4).

- **Presentations** on the five following ITS endpoints implemented in the OSIRIS webtool and its underlying concepts:
  - *Skin Sensitisation*
  - *Repeated Dose Toxicity*
  - *Bioconcentration Factor (BCF)*
  - *Aquatic Toxicity*
  - *Mutagenicity and Carcinogenicity*
- **Demonstration** of the ITS webtool with concrete examples for each specific endpoint
- **Practical application** of the webtool by workshop participants, exercises and discussions for each specific endpoint
- **Feedback from stakeholders** conducted within several questionnaires, and comments during the discussion

2.3 Agenda of the Workshop

The workshop was divided into five different sessions according to the ITS developed and implemented into the OSIRIS webtool. All single sessions included background information on the ITS, plenary discussions and demonstration of the ITS webtool with concrete examples followed by practical application and exercises. The sessions have been concluded with feedback circles (see 2.4 Feedback Circles and Questionnaires).

Figure 1: Agenda of the Stakeholder Workshop

OSIRIS Stakeholder Workshop Integrated Testing Strategies UFZ, Leipzig, Germany

Tuesday 8 March 2011

8.00 – 9.00	Registration and check of the software implementation on participants' laptops	
INTRODUCTION		
9.00 – 9.30	Welcome Introduction on OSIRIS and ITS	Gerrit Schüürmann, UFZ
9.30 – 9.45	Demo of the OSIRIS Webtool	Eduard Pauné, SIMPPLE
SKIN SENSITISATION		
9.45 – 10.30	Background REACH requirements, ITS Introduction skin sensitisation ITS webtool	Emiel Rorije, RIVM Tom Aldenberg, RIVM
10.30 – 11.00	Coffee Break	
11.00 – 12.45	Demo of skin sensitisation ITS webtool with concrete examples Practical application including exercises by participants	Emiel Rorije, RIVM Tom Aldenberg, RIVM
12.45 – 13.15	Feedback: skin sensitisation ITS (questionnaire distributed in addition)	Emiel Rorije, RIVM Tom Aldenberg, RIVM
13.15 – 14.15	Lunch Break	
REPEATED DOSE TOXICITY		
14.15 – 15.00	Background REACH requirements, ITS Introduction repeated dose toxicity webtool	Sylvia Escher, FhG Inga Tluczkiewicz, FhG
15.00 – 16.00	Demo of repeated dose toxicity ITS webtool with concrete examples Practical application including exercises by participants	Sylvia Escher, FhG Inga Tluczkiewicz, FhG
16.00 – 16.30	Coffee Break	
16.30 – 17.15	Demo of repeated dose toxicity ITS webtool with concrete examples Practical application including exercises by participants (continued)	Sylvia Escher, FhG Inga Tluczkiewicz, FhG
17.15 – 17.45	Feedback: repeated dose toxicity ITS (questionnaire distributed in addition)	Sylvia Escher, FhG Inga Tluczkiewicz, FhG
19:00	Dinner	

Wednesday 9 March 2011

8.00 – 9.00	Registration and check of the software implementation on participants' laptops	
BIOCONCENTRATION FACTOR		
9.00 – 9.45	Background REACH requirements, ITS Introduction BCF webtool	Monika Nendza, AL Alessandra Roncaglioni, IRFMN
9.45 – 10.30	Demo of BCF ITS webtool with concrete examples Practical application including exercises by participants	Anna Lombardo, IRFMN Alessandra Roncaglioni, IRFMN
10.30 – 11.00	Coffee Break	
11.00 – 12.00	Demo of BCF ITS webtool with concrete examples Practical application including exercises by participants (continued)	Anna Lombardo, IRFMN Alessandra Roncaglioni, IRFMN
12.00 – 12.30	Feedback: BCF ITS (questionnaire distributed in addition)	Anna Lombardo, IRFMN Alessandra Roncaglioni, IRFMN
12.30 – 13.30	Lunch Break	
AQUATIC TOXICITY		
13.30 – 14.00	Background REACH requirements, ITS	Alessandra Roncaglioni, IRFMN Anna Lombardo, IRFMN
14.00 – 15.00	Introduction aquatic toxicity webtool and associated ITS workflow	Anna Lombardo, IRFMN Alessandra Roncaglioni, IRFMN
15.00 – 15.15	Feedback: aquatic toxicity ITS (questionnaire distributed in addition)	Alessandra Roncaglioni, IRFMN Anna Lombardo, IRFMN
15.15 – 15.45	Coffee Break	
MUTAGENICITY & CARCINOGENICITY		
15.45 – 16.15	Background REACH requirements, ITS	Dinant Kroese, TNO Emiel Rorije, RIVM
16.15 – 17.15	Introduction mutagenicity & carcinogenicity webtool and associated ITS workflow	Dinant Kroese, TNO Emiel Rorije, RIVM
17.15 – 17.30	Feedback: mutagenicity & carcinogenicity ITS (questionnaire distributed in addition)	Dinant Kroese, TNO Emiel Rorije, RIVM

2.4 Feedback Circles and Questionnaires

Feedback from stakeholders was conducted through several questionnaires during feedback circles that took place in the end of each workshop session. Before evaluating the OSIRIS webtool participants have tested the webtool for all particular endpoints during each workshop session. Conducting stakeholder feedback aimed to improve the ITS practical usage for assessment of testing strategies and the OSIRIS webtool in particular.

The questionnaires were divided into one main general block

- *General Questionnaire*

and five specific blocks for the ITS endpoints implemented into the OSIRIS webtool

- *Questionnaire on Skin Sensitisation*
- *Questionnaire on Repeated Dose Toxicity*
- *Questionnaire on Bioconcentration Factor (BCF)*
- *Questionnaire on Aquatic Toxicity*
- *Questionnaire on Mutagenicity & Carcinogenicity*

The general questionnaire consisted of open questions concerning workshop participants "overall impression" on navigation, usage and structure of the OSIRIS webtool. In addition, they were asked if they have faced technical problems and what kind of problems while testing the webtool. The participants gave feedback if they have missed any features and if there were any features still to add or improve. The general questionnaire included also questions on personal data about workshop participants. It was filled out in the end of the 3rd workshop session on Bioconcentration Factor.
Within each of the five specific questionnaires participants were first asked about their expectations on the ITS webtool, then to assess the practical significance and the ITS building blocks (open questions). In addition, closed-ended questions were designed for the evaluation of different testing steps of assessing substances using the OSIRIS webtool (except for the Questionnaire on Mutagenicity & Carcinogenicity because the feature is still under development).
Moreover respondents were asked to suggest further improvements as well as to add any comments on both, the underlying theoretical concepts of ITS and the implementation into the webtool for particular endpoints.

3. Results of the Workshop

As described above each workshop session included one presentation on the ITS developed and implemented into the OSIRIS webtool (see Table 1: Agenda). Subsequently, feedback circles were held in the end of each session (see chapter 2.4 Feedback Circles and Questionnaires). In total, 128 questionnaires were submitted by participants during the workshop. The following table gives an overview of the numbers of questionnaires completed within each feedback circle.

Table 1: Overview of Feedback Questionnaires

Topic of the Questionnaire	Number of Questionnaires completed
General Questionnaire	21
Questionnaire Skin Sensitisation	30
Questionnaire Repeated Dose Toxicity	25
Questionnaire Bioconcentration Factor (BCF)	21
Questionnaire Aquatic Toxicity	20
Questionnaire Mutagenicity & Carcinogenicity	11
	128 in total

3.1 OSIRIS ITS webtool

The general questionnaire was designed to capture participants´ "overall impression" of the webtool concerning its development, usage and applications. 21 Questionnaires were completed.

Webtool navigation

Questions: What is your overall impression concerning the navigation (user interface) and interactivity of the OSIRIS webtool?

Most respondents´ impressions concerning the navigation (user interface) and interactivity of the OSIRIS webtool were *"overall good"*.

- *"the interactivity and user interface is fine"*
- *"good; easy to work with", "overall good"*
- *"Navigation is quite easy", "I think it is relatively clear and easy"*
- *"Overall a good interface."*
- *"Usability / interactivity quite good."*
- *"Satisfying but could be better."*
- *"Basic idea is good but for some blocks to much expert knowledge is needed to understand the program. If the expert knowledge is available, the program is not necessary."*

Most found it "*user-friendly*", "*clear and easy*", but there were concerns about inexperienced users that haven´t been trained.

- *"it´s ok once you know it, but it is not obvious for a fresh user who hasn´t been given training"*

- *"I think it is relatively clear and easy. An inexperienced user would be overwhelmed by all the different options."*
- *"An inexperienced user would be overwhelmed by all the different options, but since s/he is not the person addressed I think it´s okay."*

Technical problems

Questions: Have you faced any technical problems with the OSIRIS webtool? Could you describe them?

Most of the respondents have faced "no" or only "few errors". The errors occurred were mainly "service errors".

- *E.g."once in a while service errors (2-3 times)"*
- *"Upload of ChemProp-generated data did not work, because the substance (same name and CAS) created as "new substance" in the webtool had received a different internal ID"*
- *"Yes. My substances page did not automatically update (needed refresh page). Some for entering estimated LOGKOW in the BCF assessment"*
- *"some errors occurred yesterday with SS"*
- *"entering scores for different modules is not consistent, for example skin sensitization QSAR=klimisch 1, but BCF QSARS klimisch 2"*
- *"Consensus Tool: not working; Navigation Tool: not working (3D Navigation)"*
- *"sometimes the workflow produced screens of other types then the speaker or other participants had (probably lack of caution?)"*

Help Button, Guidance

Questions: Did you use the Info/Help button? If yes, what did you search for? And was it helpful?

Most respondents did not use the help button, but testing the feature users often could not find the information/guidance needed.

- *"I tried to find out how to change the user password. I didn´t find it. The help is very basic."*
- *"I was looking for a "general demo" that guides me through the data processing. It should include screenshots like the demos given in the demonstration during the workshop."*
- *"I have used it, but I didn´t find the search keyword 'hydrolysis'."*
- *"I searched for a documentation of the ITS algorithm, but didn´t find it."*

Throughout the questionnaires a lack of online help, sufficient instructions/guidance or a user manual was consistently emphasized.

- *"yes, I was looking for a "general demo" that guides me through the data processing. It should include screenshots like the demos given in the demonstration during the workshop"*
- *"Guidance in the program must be improved if it should be used by non tox/ecotox experts e.g. regulators"*
- *"Help / online help / tooltips"* are missing
- *"Maybe think about an on-line help / guidance to be provided."*

Suggested improvements

Suggested improvements concerning the navigation were as follows:

- *"It is a little bit annoying that the user needs to scroll down a lot and to go back to questions that have been answered in previous assessment"*
- *"tool to view or edit the structure" is missing*
- *"more short cuts / better "go back" option" would be helpful*
- *"Graphical appeal could be improved"*
- *"entry of more the one test/QSAR value at the same time; expert modus for fast data entry;"*
- *"should be able to skip yes/no answers with "don´t know" and fill gaps later? Perhaps let users enter data into e.g. Excel and import when all info collected."*

Additional comments from the following sessions:

Respondents reported that they have missed features like a

- *"save button"*
- *"help-symbol for the abbreviations" / "more descriptions of the functions (what is meant) of the webtool" in order to improve the usability of the webtool"*

Webtool structure

Question: What is your overall impression about the structure of the OSIRIS webtool?

Concerning the webtool structure respondents´ impression were more ambiguous. Answers vary from:

- *"structure is quite clear, simple and easily to understand and to use"*
- *"It is well structured and simple, which is the advantage of the software."*
- *"it is ok, but requires some training"*
- *"you have to get used to it"* (e.g. *"confusing for different endpoints when data should be added"*)
- *"it was puzzling, not self-explaining"* (e.g. *"entry of experimental data a bit confusing and time consuming")*

Questions: Which features should be improved? Can you describe your ideas in detail, please?

Suggested improvements were mainly concerning the clear and comprehensive documentation and report for each endpoint, the parameters used and the guideline from the EU regulation which is missing in the webtool:

- *"there must be comprehensive output-reports, which clearly document the assessment", "A comprehensive output report should be generated automatically."*
- *"more descriptions on results"*
- *"an extensive reporting tool is missing"*
- *"the documentation within the webtool: It must be made very clear, which input-parameters are used. An ambiguity regarding units is not acceptable."*
- *"for each endpoint, detailed documentation of the underlying algorithm should be given in the software"*
- *"Export of results; e.g. as 'report'"*
- *"detailed documentation of algorithms → important to ensure transparency"*
- *"An extensive reporting tool is missing. The current webtool output is insufficient to provide essential information for regulators."*
- *"Guidelines for the meaning of a result should be better implemented if there will ever be strict guidelines from the administrative (E.U.)."*
- *"References and links, when available, should be given for all in silico model proposed. It would then be easier for users to really try and test them."*

Additional comments from the following sessions:

- *"more complete information on the results would be helpful (statistics, argumentation, why such result…)"*
- *"an option to save results in printable format (e.g. .txt files)"*
- *"option to output to report format that could be included with official dossier/regulation submission, etc."*
- *"import of external data"*
- *"it would be nice if OSIRIS could import itself in silico data"*
- *"Question that require "a/b"- answers should not be posed in a yes/no – manner" and "Sometimes tool demands yes/no answer, where might be "unknown"; might be better to read multiple values, assign weights and have software derive weighted value to use rather than picking highest/lowest arithmetic mean";*
- *"avoid double negation questions"*

The participants suggested entering and export of data and format should be more integrated and standardized:

- *"The flow of entering data. For the "user input" stages this becomes very repetitive. Could a single page of questions be asked?"*
- *"Fine, the ChemProp interface is very weak, maybe use at least a local web service to transfer data. Use a standard format (OECD, FoxMC, ...) to interchange data, don´t invent the wheel."*
- *"Export in IUCLID, so it could be used for other assessments not only for REACH.*

Additional comments from the following sessions:

- *"It would be nice to have something like an endpoint summary that can be copied into the UCLID endpoint summary."*

Question: Which features of the OSIRIS webtool are missing and should be added? Why are they important for risk assessment?

Suggested improvements concerning the structure were the following:

- *"As not every user has access to some data, e.g. commercial QSAR results, this should be somehow taken into account by the webtool. More flexibility in input."*
- *"...but data entry seems strongly tied to following a decision tree. Users may compile dossier/file over time and add specific values as they arise? At least decision tree highlight missing data gaps."*

Suggestions concerning missing features which should be integrated as "Derek", "ToxTree", "ChemProp" and prediction of proteins.

- *"would be nice to get SMILES direct → link to web;"*
- *"Displaying and drawing chemical structures. More PhysProp calculators like ToxTree, etc"*
- *"Prediction of protein/XXX binding (a plasma protein binding/bioavailability and a hepatic clearance), and reactive metabolites (very much more complicated I know); Endocrine disruption?"*
- *"the log B → log BCF issue should be addressed; How can "different"/"new" QSARS be interpreted in the future? E.g. Lhasa ("Derek") are working on an XMC interchange for rats as a standard, can this or any other be integrated?"*
- *"Important additional software (ChemProp) must be installed locally, it should run on a server with secure access.", "ChemProp should run as a web services."*

Usage

Questions: Do you think you will use the OSIRIS webtool in your daily work? If so, what are the main reasons for you to switch to this webtool? If no, what are the main reasons for not using the webtool?

Concerning the usage the answers were ambiguous and varied from using the webtool "definitely", "maybe", to "not sure", "I will give a try" or "not in the current status":

- *"Definitely yes"*
- *"Maybe I would use it to get a "better idea", how outcomes of different information sources add statistical "weight of evidence""*
- *"I am going to try it. An advantage would be a standardized procedure (especially if it would be reported). One disadvantage is that not all of the relevant REACH endpoints are provided"*
- *"for sure cases I will use it"*
- *"I would use it for strategy development evolution, but not to avoid tests so far"*
- *"I will try to use it within TP under REACH when repeated dose test are more commonly proposed tests up to now."*
- *"I assume that OSIRIS webtool could be used as a supporting tool for the evaluation, testing strategy etc."*
- *"could be useful, but would need to dedicate time to building databases, running QSARS etc.. Perhaps EU project to compile info on known components with reported data (text mining?)"*
- *"not sure"*
- *"no, too expensive for university; no (fate) module included"*
- *"Not in the current status of development, for all endpoints would prefer an output which tells when the WOF is sufficient to XXX a REACH endpoint."*

Questions: Have you learned something about ITS? If yes, please write down what you find most promising about ITS?

Most of the participants explained that they got new insight in ITS and combing toxicity information. They are more confident that ITS will be more implemented in the future.

- *"The idea to combine information from different sources, to come to a conclusion convinces me. How this is implemented is another question."*
- *"The interpretation of QSAR for "basic" XXX with decision tree XXX XXX alerts by Bayesian Statistics X sound and promising;"*
- *" I like that it uses different existing models and combines the advantages and disadvantages;"*
- *"Be more trusting in the usefulness of alternative methods and schemes."*
- *"nothing really new"*

Summing up, the majority of the respondents reported that the OSIRIS webtool is user-friendly and the navigation clear and easy. The results have been clearly documented, readily accessible, understandable and reasonable, but inexperienced users as non-toxicity experts could be overwhelmed by the different options. Helpful would be a general demo version which guides the user through the data processing and explains the structure of the webtool. Advanced users would like to use more short cuts in the navigation and a faster entry as inexperienced users should be able to skip yes/ no-answers with alternatives such as "don't know" and "fill in later".

Some respondents requested to improve the visibility of the workflow, since they have perceived the documentation as unclear and – in general - the information too limited. They see the need to give more background information on probabilistic calculations and to *"indicate the correlation between models"*. An additional tool to view and edit the structure would be helpful, too.

However, some respondents criticized, that the documentation *"wasn´t well structured"*, that is was *"insufficient"* due to a *"lack of information on how the probabilities were (actually) obtained"* and that *"basic assumptions were not displayed in the webtool"*.

A comprehensive output-report document for the assessment was suggested. It should be described which input-parameters are used and the underlying algorithm should be given.

A standard format as OECD, FoxMC e.g. is recommended. Missing features to import and export data and direct links to additional software as "Derek", "ToxTree", "ChemProp", prediction of proteins e.g. should be integrated.

The participants would use the OSIRIS webtool for strategy development and to get a better idea how information sources add statistical "weight of evidence". The OSIRIS webtool is seen more as a supporting instrument than a stand-alone version to avoid animal testing because it *"still relays more on human power rather as a technical tool"*.

Some participants noted that all the good ideas which are developed in the project should be *"preserved and further developed"*. *"Also some modules (Bayes algorithm, RepDose section) should be followed up upon."* Most of the participants explained that they got new insight in ITS and combining toxicity information. They are more confident that ITS will be more implemented in the future.

3.2 Endpoint “Skin Sensitisation”

3.2.1 Summary of Presentation "OSIRIS ITS for Skin Sensitization"

By Emiel Rorije and Dr. Tom Aldenberg, National Institute of Public Health and the Environment (RIVM), the Netherlands

Most approaches to Integrated Testing Strategies (ITS) comprise of flow schemes leading to a certain decision (i.e. classification) when specific information is present (i.e. a positive result in an in-vivo test for a certain toxicological endpoint). However, REACH requires registrants for each REACH relevant information requirement to “gather and share all available information”, compare this information with the information requirement for the applicable tonnage band, identify the information gaps, and subsequently generate new data or propose a testing strategy to ECHA (REACH Annex VI, Chemical Safety Assessment).
Taking into account ALL available information with respect to a certain information requirement under REACH calls for a procedure that evaluates all information in parallel, not sequentially (as is often the case in the flow schemes i.e. for a number of endpoints in the REACH Guidance Documents. Especially when different sources of information are - on their own - not sufficient to fulfill the REACH information requirements, or if contradicting information is available, a Weight of Evidence (WoE) procedure is required. In order to perform such a WoE procedure in a transparent, reproducible, objective and flexible way we have developed in the OSIRIS project a *quantitative* WoE approach, making use of Bayesian statistics, using what is called a Bayesian Inference networks.
The (general) weight of each information source in such a network is defined by specifying the sensitivity and specificity (Cooper statistics) of each methodology, model, *in vitro* test, or other information source towards a Gold Standard. In the ITS presented here (REACH Skin Sensitization) the Gold Standard is one of the *in vivo* test results indicated by the REACH guidance to fulfill the information requirement (at the Annex VII level, and above); the Local Lymph Node Assay (LLNA). In this ITS the cooper statistics of 5 different QSAR models (DEREKfW, SMARTs, TopKat, TIMES-SS and DK-EPA Multicase model), 2 *in vivo* tests (LLNA and GPMT), 1 *in vitro* test (hClat), and 1 human endpoint (Human patch test) are determined, in order to be able to “add” probabilities from evidence (positive or negative result) from these tests, for the endpoint of Skin Sensitization.
Bayesian Inference Networks allow to update the probability that a given hypothesis is true (i.e. substance is a Skin Sensitizer), given specific information. Without any information we assume that there is a 50% probability that a substance is a skin sensitizer (uninformed prior probability). When we now “add” a positive result from a QSAR model, the probability that this substance is a skin sensitizer increases to 70% (based on the sensitivity and specificity of that specific QSAR model). By comparing this probability to a predetermined probability threshold, which represents the REACH information requirement, one can quantitatively determine whether the information requirement has been met, or if there is still a data gap. Furthermore, it allows the user to identify the most efficient way to fill a data gap, by

selecting the most cost effective methodology that is able to increase the probability to the required level. The probability level as required by REACH has been determined from the fact that both the *in vivo* GPMT as well as the LLNA test are acceptable information to fulfill the REACH requirements. The probability that one can predict the outcome of the LLNA test with a given result of the GPMT test is therefore apparently an acceptable level of probability for REACH. From this assumption we have determined that the WoE probability of a substance being a skin sensitizer (positive) should be above 80%, whereas the WoE probability of a substance being a non-sensitizer should be above 90%, in order to be acceptable in REACH.
The OSIRIS scheme allows for individual modification of the influence (weight) of a piece of information by adding (user assigned) Quality Factors. In the simplest implementation the Klimisch, or Klimisch-like codes are used and translated to a quantitative Quality Factor. This is what has been implemented as the Bayesian Weight of Evidence procedure for Skin Sensitization in the OSIRIS webtool.

3.2.2 Feedback and Recommendations

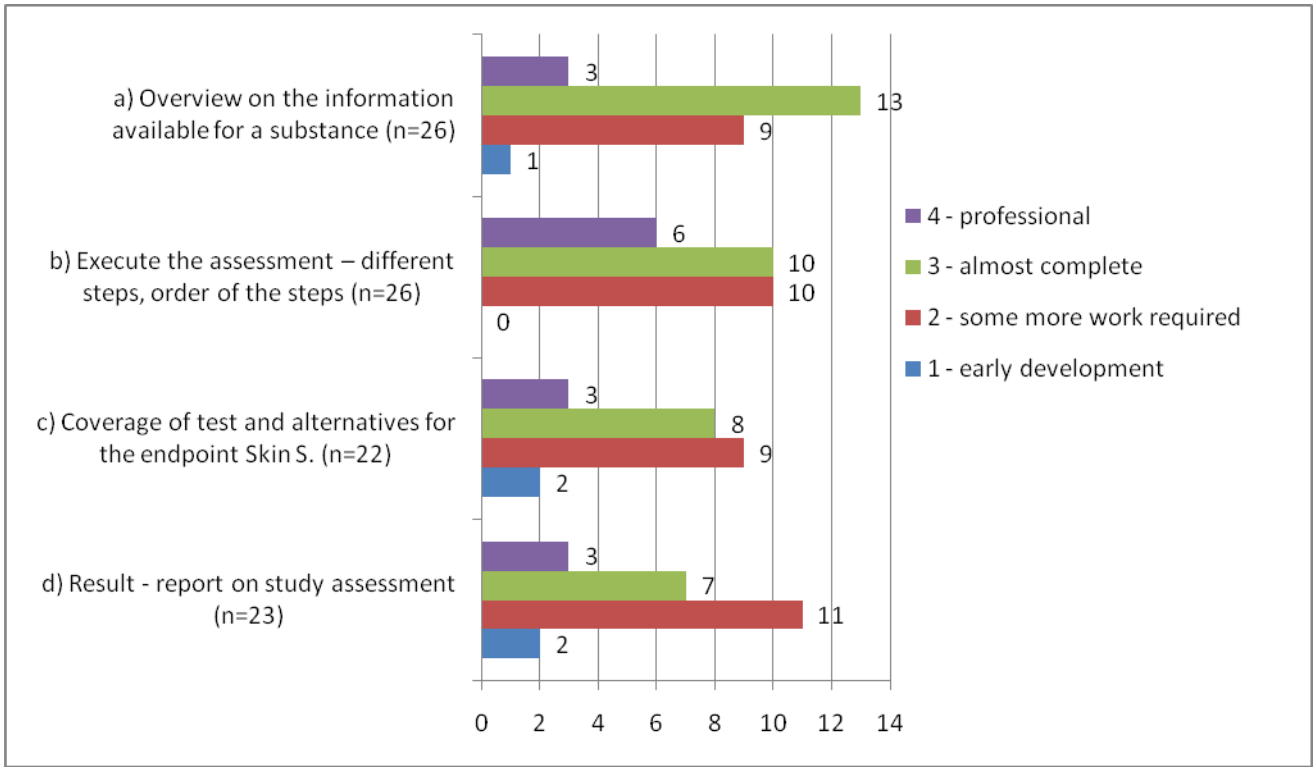
Question: What do you expect from the ITS Skin Sensitisation in the OSIRIS webtool?

According to the 30 questionnaires completed within the skin sensitisation session the respondents mainly expected the following from the OSIRIS webtool:

- *a "structured workflow"*
- *a transparent process and transparent reporting*
- *a "large database representing all types of substances"*
- *a (reliable) "prediction of skin sensitisation" and thereby acceptance with regulation*
- *"recommendations for successful (further) testing strategies"*
- *support for decision making while fulfilling REACH requirements*

Evaluation of different assessment steps in the OSIRIS webtool
After trying to assess several substances during the skin sensitization session, respondents were asked to evaluate different steps of the assessment. Figure 2 demonstrates the evaluation of each single step as compiled within the questionnaire. Absolute frequencies are shown.

Figure 2: Evaluation of different steps of assessing substances (N=30)



Overall, most of the respondents see the development of assessment steps in the middle range in the continuous scale of one to four. That means, that the assessment steps are considered as *„almost complete"* or that there is still *„some more work required to make a complete application"*. Altogether only very few respondents – from none to two respondents - considered the assessing steps as being in an *„early development*" status. Slightly more have considered them as *„professional*" whereby the step "*assessment execution*" was rated best.

Comments to building blocks: How is your impression of the building blocks? Do you have further comments regarding the background of the main building blocks?

Most of the respondents reported that the statistical approach (Bayesian decision theory) is helpful to "*justify on a scientific basis the choice of acceptability of a final result*". *"Especially if it makes more transparent, how conclusions were obtained and how reliable they are."* However, the use of Klimisch-like scores to "penalize" less reliable results is seen more ambiguous. Some of the respondents supported the score as a good guidance, but a lot of them reported that Klimisch-like scores are *"less practical"* and need a lot of additional guidance.
The majority of the respondents criticized that more in vitro tests or any experimental data should be added to get better "weight of evidence" for skin sensitisation toxicity. *"For the calculation of predicting of in silico methods, these methods were compared with LVNA ...".*

Furthermore respondents suggested:

- *"more self-explaining fields/entries"*
- *"open the tool for user to define their own tests and allow user to adjust predicting..."*
- an option to *"enter the uncertainty of each result of the models entered"*
- *"further integration with QSAR software"*

3.3 Endpoint “Repeated Dose Toxicity”

3.3.1 Summary of Presentation "ITS Repeated Dose Toxicity"

Presentation by Sylvia Escher and Inga Tluczkiewicz, FhG

The ITS RepDose focuses on the assessment of the reliability/relevance of available testing data. Within REACH, the toxicity of existing chemicals has to be assessed. Numerous of these chemicals have been on the market for decades. A typical problem under REACH is therefore, that toxicity data are derived from “old” studies which were conducted before the publication of OECD guidelines in 1981. Today, these studies are considered as non-guideline which indicates that study results are not reliable or only reliable with restrictions. However, also “old” studies have a certain probability to detect target organs and NOEL values similar to those detected in guideline studies. The probability or the validity of the "old" study depends on the scope of examination and the overall quality of the study.
So the basic question for the RepDose ITS is to assess to which extent and under which conditions old studies can be used for risk assessment. In OSIRIS two methods have been developed which assess the scope of examination, called coverage approach and the overall study quality, called QUANTOS. Both methods are explained in more detail in the following:

Coverage approach
Studies which have been conducted before guideline may differ in scope of examination and thus there is only a certain probability that they "cover" all relevant examinations and have thus identified the correct study LOEL. If so, targets and organs have to be identified that trigger the study LOEL in guideline studies to assess the relevance of non-examined organs.
It has been shown that for about 90% of the subchronic guideline studies in rats the study LOEL was triggered by six main target organs (liver, kidney, body weight, clinical chemistry, clinical symptoms, and hematology). The relative importance of these organs and their predictivity to the study LOEL was evaluated using the Poisson regression.

QUANTOS
Beside coverage also the overall quality of the study will influence the reliability of study results. In non-guideline studies number of animals and study durations may differs from guideline. Recently ECHA has published the ToxRtool, which can be used to assign a quality score, the Klimisch code, to experimental studies and thus allows the more or less automated assessment of study quality. The about 23 questions of the ToxRtool were used as starting point to develop a tool which assesses the quality of a non-guideline study in the ITS RepDose. 4 questions were identified to be crucial for repeated dose toxicity studies, and termed knock out criteria. If one these questions is missing, the study quality can not be assessed. Five further questions are included in the **QUANTOS** tool, **QU**alitative **A**ssessment of **N**on guideline **TO**xicity Studies. With the use of an intra- and inter attribute

weighting, QAUTNOS concludes on study quality giving a fraction, with a maximum of 1 (full reliable).

Alternative methods

Alternative methods for repeated dose toxicity studies are limited. The current ITS RepDose proposes read across/category approach if a data gap is identified. If read across/category approach is not possible exposure based waiving is an option, where the threshold of the TTC concept will be used to define exposure limits below which no risk for human health is assumed. Also QSAR methods could be used to predict the toxicity of compounds. Within OSIRIS the performance of the TopKat model for substances in and out of its applicability domain was evaluated with the RepDose DB. TopKat predictions were not considered to be reliable and thus not included into the ITS RepDose.

3.3.2 Feedback and Recommendations

Question: What do you expect from the ITS Repeated Dose Toxicity in the OSIRIS webtool?

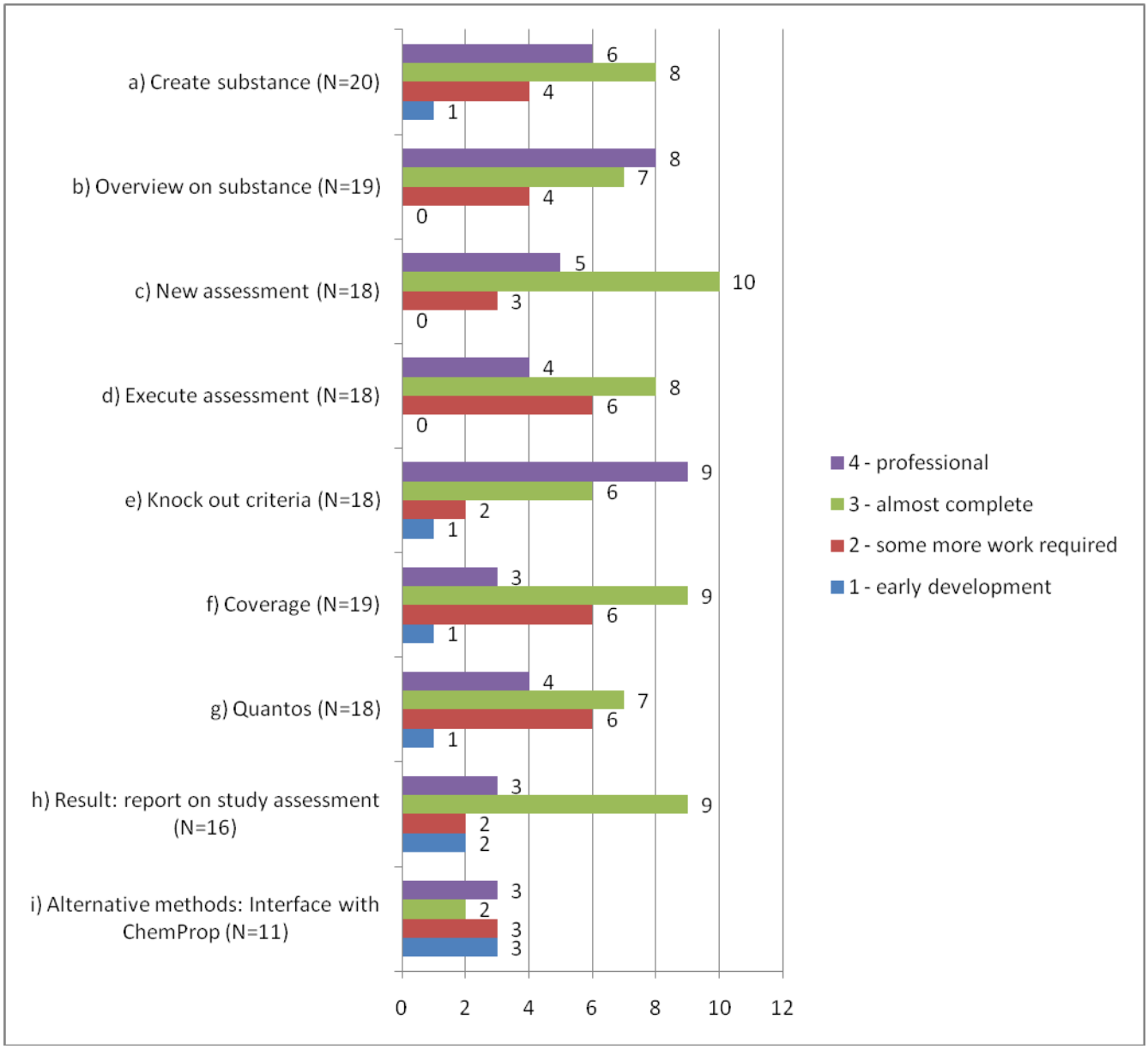
25 Questionnaires were completed for the repeated dose toxicity session. The respondents mainly expected that the webtool in order to predict toxicity

- provides risk assessment guidance
- provides a framework for assessment and a *"guided workflow with detailed explanation on each topic"*
- *provides "a transparent account of the studies or tests available including their reliability and deficits"*
- enables a *"comparison between different results", "Evaluation use of "non-reliable" data"*
- *"I would expect more focus on new in vitro developments."*

Evaluation of different assessment steps in the OSIRIS webtool

After trying to assess several substances during the repeated dose toxicity session, respondents were asked to evaluate the assessment steps. Figure 3 shows how workshop participants have evaluated particular steps in the assessment of substances. Absolute frequencies are shown.

Figure 3: Evaluation of different steps of assessing substances (N=25)



Respondents were similarly satisfied with the steps "*create substance, "overview on substance", "new assessment" and "execute assessment*". There are no outlying responses. Concerning the first steps of the assessment were most responses within the „*almost complete*" or the "*professional*" answering option, but some indicated that there is still "*more work required to make a complete application*". Only one respondent evaluated one of the first steps *("create substance")* as being in an "*early development*" status.

The evaluation of the following steps was more ambiguous. The step „*Knock out criteria*" is overall the best rated and displays a very high grade of contentment. "*Coverage*" and "*Quantos*" meet a similar degree of contentment. They are considered as "*almost complete*" by most of the respondents. Again, similar to the previous steps, early development

judgments were near to nothing. Quite ambiguous are the responses concerning the interface with ChemProp, since they are almost uniformly distributed.

Comments to building blocks: How is your impression of the building blocks? Do you have further comments regarding the background of the main building blocks?
According to a large majority of the respondents the results of the assessment have been clearly documented. Some reported that there have been (minor) restrictions and that *"transparency on the detail level* (documentation of underlying assumptions) *could be improved"*. It was stated that it was *"too much a black box of calculation"*. Some of the respondent suggested that *"the judgment should vary on a case-by-case basis"*, between an expert judgment and criteria (QANTOS).
However, concerning the documentation the main aspect claimed was the improvement of the (stepwise!) process documentation.

Further suggestions were

- to *"take read-across data into consideration"*
- data input should be more flexible
- *"new in vitro developments"* should be focused more
- *"exposure must come first!"*.
- *"to calculate TTC risk at the very beginning, before entering all study details"*
- *"copy/export possibility for coverage and QUANTOS data as entry for IUCLID reliability justification in a summarized form"*
- more descriptions of questions and definitions for clinical chemistry and symptoms are needed
- *"In further development, the tool should also take read-across data into consideration;"*

3.4 Endpoint "Bioconcentration Factor (BCF)"

3.4.1 Summary of Presentation "Bioconcentration Factor (BCF)"

By M. Nendza, Analytisches Laboratorium, Luhnstedt, Germany

The efficient assessment of the bioaccumulation potential of chemicals under REACH with integrated test strategies (ITS) requires multiple tools. Existing data have to be searched and information from chemical structures and physico-chemical properties need to be evaluated prior to considering to conduct *in-vivo* experiments with vertebrates. The OSIRIS inventory of chemistry-driven and *in-silico* BCF modules for ITS compiles:

- Sources of existing data
- Computational methods
    - B/nonB classification models
    - QSARs
    - Physiological models
    - Exposure models
    - Read across
- *in-vitro* tools
- 3R (Refine, Reduce, Replace) modules

The ITS components for bioaccumulation listed in the ECHA Guidance on information requirements and chemical safety assessment[3,4] have been extended with new knowledge generated in OSIRIS and complemented with feedback from stakeholders on the actual problems in using ITS for chemical registration. The alternative ITS modules share three major objectives to save time and money by reducing the number of experimental animals required to come to a conclusion about the bioaccumulation potential of chemicals under REACH:

- Classification of non-B/B/vB-compounds
- Omission of BCF studies, that are scientifically unnecessary or technically not feasible
- Waiving of BCF studies, that provide no risk-relevant information

The OSIRIS ITS for bioaccumulation will be publicly available (webtool) after further refinement based on stakeholder feedback. Its concepts and modules, as well as validation results, are presented in detail in a dedicated poster corner at SETAC Europe, 2011.

[3] ECHA. 2008. Guidance on information requirements and chemical safety assessment Chapter R.11: PBT assessment. (http://echa.europa.eu/reach_en.asp). Accessed 6 March 2009.

[4] ECHA. 2008. Guidance on information requirements and chemical safety assessment Chapter R.7c: Endpoint specific guidance. (http://echa.europa.eu/reach_en.asp). Accessed 6 March 2009.

3.4.2 Feedback and Recommendations

Question: Was the OSIRIS webtool for BCF in line with your expectations?

Within the ITS bioconcentration factor (BCF) session 21 Questionnaires were completed. The webtool for BCF met most of the respondents' expectations:

- *"it is round and good approach"*
- *"clear definition for what group of substances the tool can be applied"*
- *"Very important and useful tool to avoid some tests"*

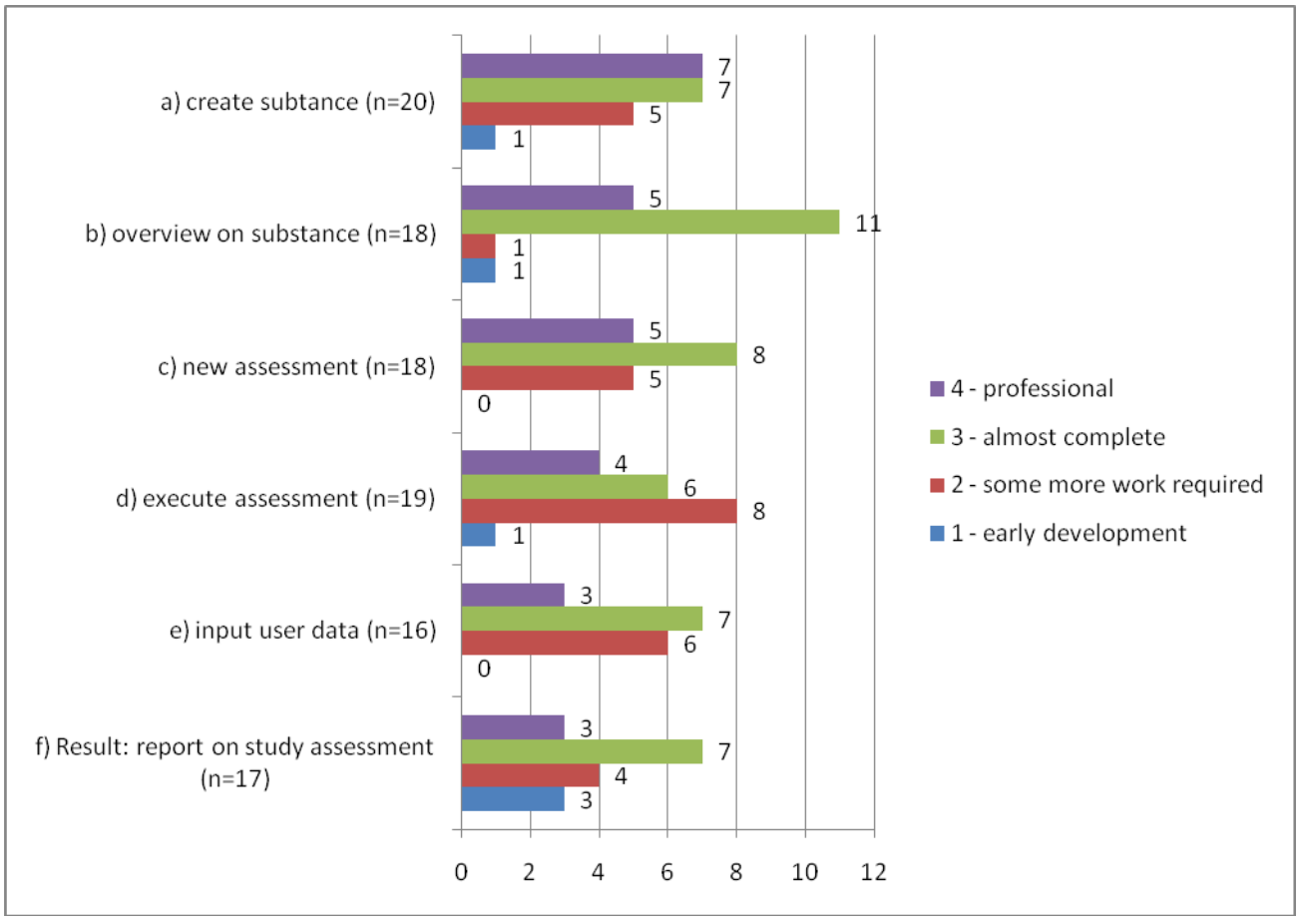
A few respondents expected minor improvements. They have criticised/suggested that:

- there were *"no advantages compared to a workflow diagram"*
- the "webtool *"follows too simply the decision tree, there were no weight-of-evidence approach for "mixed" data situations"*
- *"links to QSAR models"* should be provided

Evaluation of different assessment steps in the OSIRIS webtool

Figure 4 demonstrates how workshop participants have evaluated particular steps in the assessment of substances. Absolute frequencies are shown.

Figure 4: Evaluation of different steps of assessing substances (N=21)



Evaluating different steps of assessing substances, respondents were generally satisfied with the webtool: The second step *"overview on substance"* was clearly considered as *"almost complete"*, *w*hereas responses on the other assessment steps are similarly distributed over the scale from *"professional"* to *"some more work required"*. As before, in the webtool evaluation of other endpoints only very few respondents regarded the assessment steps as being in an early stage of development, except for the documentation, *"report on study assessment"*. Responses to this step are comparatively ambiguously.

Questions: Comments to constituting elements of the webtool: what is your impression of these elements? Do you have further comments regarding the scientific background of the main constituting elements?

"The methods for in silico evaluation seem to be exhaustive". The respondents recommended to *"add a weight-of-evidence approach for "mixed" data situations"* and the results *"should be more weighted"*. Additionally, *"some more parameter are missing! (Density; molecular weight, ..)"* and should be integrated. The second suggestion was to use more results from in vitro data for WOE approach *("maybe more details on in vitro, how to use results from in vitro options, how to choose"*). Respondents would like to *"include Bioconcentration or Biomagnification of metabolism"* and receive *"... recommendations for other tests, like a BMF test"*.

Questions: What should be improved? What is missing? Is there something misleading? Please consider this question both in terms of usability of the webtool as well as for the scientific concepts behind the current webtool workflow.

Most suggestions refer to the documentation of the assessment. In particular, *"more explanations (on functions) during the assessment workflow"* were expected. Some noted that the *"documentation should be more detailed"*. *"A clear and detailed documentation of the underlying algorithm should be given in the software documentation"* and that *"a description of the uncertainties should be included"*.

3.5 Endpoint “Aquatic Toxicity”

3.5.1 Summary of Presentation "Aquatic Toxicity"

By Anna Lombardo and Alessandra Roncaglioni, Instituto di Ricerche Farmacologiche “Mario Negri”, Milan, Italy

REACH requirements together with rules for Classification and Labelling, to assess the T criterion of PBT and to perform the Chemical Safety Assessment were used to built an ITS for Aquatic toxicity. The rationale of the scheme includes four steps: 1. justification for waiving the tests (degradation of the chemical and presence of mitigating factors), 2. a tonnage evaluation to satisfy requirements for CLP, REACH annexes and preliminary CSA, 3. Investigate T criterion for PBT assessment and 4. an evaluation to decide if CSA refinement is needed and how to satisfy its requirements. For each step some methods were proposed.
To better understand how the scheme works two substances were used. The first substance (CAS No. 58430-94-7) were supposed to be produced or imported below 10 tonnes/y. The real water solubility (above the threshold for poorly soluble substances) was used firstly. In this case, it is possible to reach a conclusion on the basis of acute toxicity data of algae/aquatic plant and invertebrates. Secondly it was assumed that on the contrary the substance is poorly soluble so both chronic and acute values for invertebrates were necessary (over information for algae/aquatic plant). In this case the acute toxicity values for invertebrates should be obtained using non-testing methods.
A second substance (CAS No. 84-74-2) was supposed to be produced or imported above 10 tonnes/y. Also in this case both real water solubility and an hypothetic one were used. If the substance is poorly soluble acute toxicity data for algae/aquatic plant and invertebrates and chronic toxicity data for invertebrates and fish are necessary to satisfy C&L and annexes requirements. Otherwise, only acute dare are required. Above 10 tonnes/y also PBT assessment and CSA are necessary. To explore this part of the scheme three different combination of toxicity data were used: 1. only acute data, 2. acute data for the three trophic levels and chronic data for fish and 3. acute data for algae/aquatic plant and both acute and chronic for invertebrates and fish. To satisfy PBT and CSA requirements chronic toxicity data for invertebrates and fish were needed for all the three cases. To reduce in-vivo tests, alternative methods (such as in-silico, in-vitro or limited in-vivo test) were always suggested before in-vivo test.

3.5.2 Feedback and Recommendations

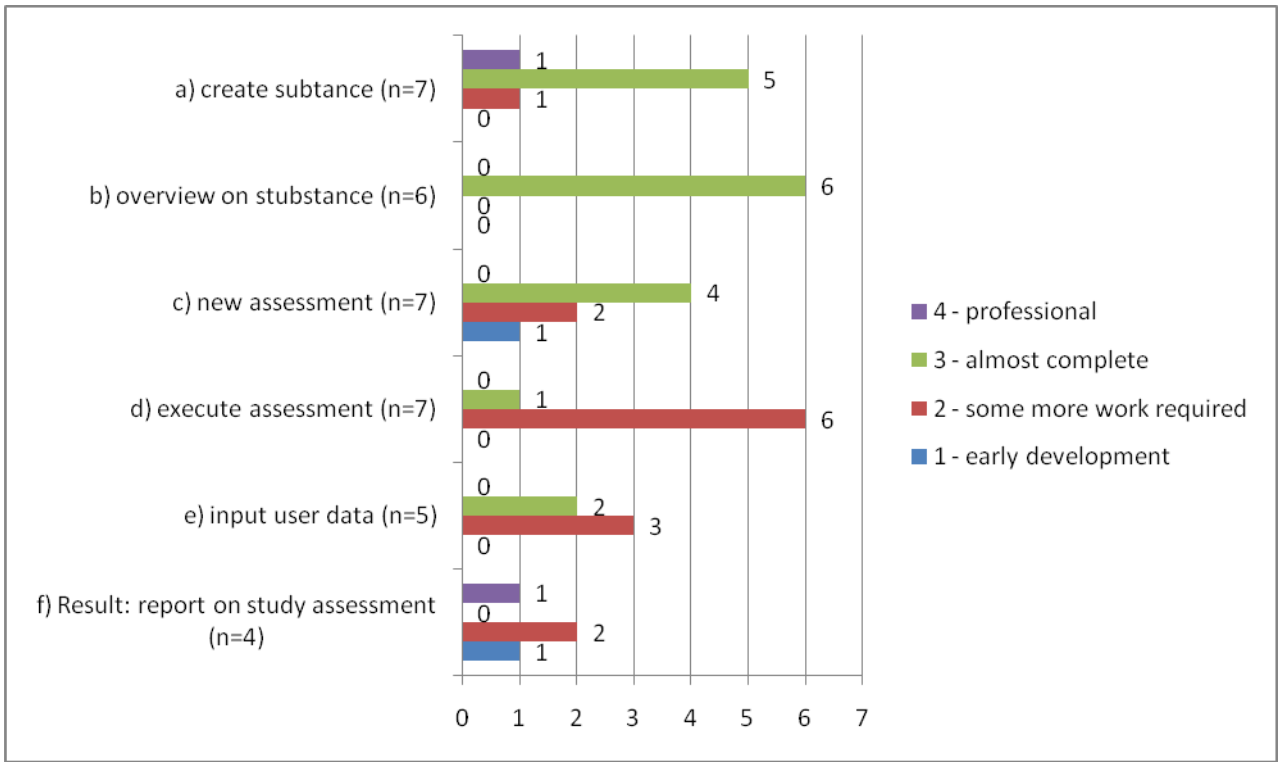
Question: Was the OSIRIS webtool for aquatic toxicity in line with your expectations?

Even though 20 Questionnaires were completed for the ITS aquatic toxicity session the non-response rate within these questionnaires is considerably higher compared to other sessions.

This is due to the fact that the webtool haven´t been implemented into the webtool yet. Therefore evaluation and judgements were nearly impossible.
However, based on the presentation in the workshop, the approach to the ITS aquatic toxicity is according to some respondents *"overall good"* and in line with their expectations, while others noted that they have expected *"a shorter workflow"* and that *"some work remains necessary"*.

Evaluation of different assessment steps in the OSIRIS webtool
Again, being asked to evaluate different steps of the assessment, it was the case that there were only a fewe comments – based on the presentation only - since workshop participants' haven´t had the chance to test the tool (see Figure 5). The total number of respondents ranges from only four for the item *"result"* to seven the maximum for other steps. These responses can hardly be interpreted, but possibly we can see a trend towards the better development of the first steps compared to the last ones.

Figure 5: Evaluation of different steps of assessing substances (N=20)



Questions: Comments to constituting elements of the webtool: what is your impression of these elements? And what should be improved? What is missing? Is there something misleading?

Based on the presentation in the workshop the following improvements were suggested: Again the respondents would like to be provided with more *"scientific background and*

guidance documents" and a more detailed *"guidance about waiving, decisions and use of QSARS results and the information of the decision"*. Some of them reported that *"the module too strictly follows a decision tree..."* and should have an enhanced integration and more transparency *"on how validity of QSAR results/ACR prediction is assumed"*. The respondents suggested to focus more on in vitro data and include more *"biodegradation parameter like OECD 301XXX"*.

3.6 Endpoint "Mutagenicity & Carcinogenicity"

3.6.1 Summary of Presentation "Mutagenicity & Carcinogenicity"

By Dinant Kroese, TNO, Zeist, the Netherlands

Concepts of Integrated Testing Strategies (ITS) have been described in REACH Guidance documents that provide a general description of how to obtain the information required. Currently, more guidance is needed on how to facilitate an efficient gathering and use of all relevant information to provide a sound basis for classification & labeling and risk assessment: in fact, what preferably is needed is a quantitative weight-of-evidence approach that helps assessors to conclude when information does satisfy the information requirements of REACH.

Being a categorical endpoint in nature, the approach for assessing weight of available information for mutagenicity will be similar to what has been described before for skin sensitization: i.e. all available sources of information need to be defined by specifying their sensitivity and specificity towards the Gold Standards (GS) for this endpoint, and the probability that a substance is negative or positive will be determined by these characteristics and the prediction for the substance by this specific test. Contrary to the situation with skin sensitization, for mutagenicity *five* Gold Standard tests are defined as this endpoint itself in fact represents three distinct endpoints: gene mutations, chromosomal aberrations, and aneuploidy. These GS are: Ames test for bacterial gene mutation tests, the micronucleus or chromosomal aberration test as *in vitro* cytogenicity test in mammalian cells, the mouse lymphoma test as the *in vitro* gene mutation test in mammalian cells, and the *in vivo* tests for micronuclei in bone marrow, and induction of UDS in rodent liver. Also contrary to the situation for skin sensitization, and being unfortunate, is the situation that for none of these GS databases of alternative test results are available, making the described approach not applicable in this specific situation. On the other hand, quite a number of SAR models can be used for predicting bacterial mutagenicity: e.g. CAESAR, and Benigni-Bossa to name just two. Clearly, these models may not be completely independent in their prediction; for these two models the dependency of their predictions will be illustrated. Also, a methodology of weighing an incomplete Ames test, i.e. with one or a few Salmonella strains missing, will be illustrated. Although not yet implemented into the webtool, the ITS for carcinogenicity will be shortly discussed as well: especially the availability and usefulness of alternative sources of information and how to weight an imperfect carcinogenicity study, i.e. one lacking examination of a few organs or tissues.

The webtool for mutagenicity will be 'tested' with two case-examples; glutaraldehyde, and 4-methylcyclohexanol.

3.6.2 Feedback and Recommendations

For the last workshop session on Mutagenicity and Carcinogenicity were 11 feedback questionnaires submitted. Respondents overall expected

- *"a stringent workframe for supporting testing"* from the webtool
- *"to discover what was available of in vitro tests and what was taken into account for statistical methods used"*

The webtool even exceeded some expectations

- *"I was pleasant surprised with the tool", "good schematic workflow", "surprised with the Bayesian approach but this concept was as well XXX and made sense to use this method. Reduces uncertainty at least in my opinion"*
- *"I think the REACH ITS for C&M are so complex that I didn´t expect the OSIRIS project to be able to reproduce it satisfying within a "simple" webtool"*

However, it was criticized by two respondents that there were

- *"still a lot of technical difficulties"* and that the ITS was in a *"too early state of development"*

Suggested Improvements for the webtool

The predominant suggestion was to focus more *"on in vitro tests and the evolution of data from in vitro"*. A *"more automatic workflow"*, *"default settings whenever possible"*, as well as *"pre-selected values"* should be provided. The respondents suggested that *"the options COECDxy should be adjusted to the selected endpoints"*. Finally, the participants noted that a *"(brief) description of the test"* was missing.

4. Conclusion

The ITS webtool tool offers an implementation of the methods and ITS developed within the OSIRIS project. The ITS are implemented as user-interactive workflows that weight different types of data (obtained from databases, computed with models or input by the user). The ITS goal is to conclude if there is sufficient information for Classification & Labelling and Risk Assessment, and to suggest the appropriate test in case that not enough information is available for decision making. So far, five endpoints are included in the webtool (skin sensitization, repeated dose toxicity, bioconcentration factor, aquatic toxicity, mutagenicity and carcinogenicity).

The fourth OSIRIS Stakeholder Workshop aimed to present the current state of work regarding the development of the webtool and its underlying concepts, to explain five ITS endpoints, ITS building blocks and testing steps developed and implemented into the webtool. It also aimed at the practical application of the OSIRIS webtool as well as exercises and discussions for each specific endpoint by workshop participants. Hence, all invited stakeholders and experts from industry, regulatory authorities and academica have tested the methods and integrated testing strategies (ITS) for the five endpoints implemented in the webtool. Their feedback was conducted through questionnaires, one on the “general” application of the webtool and one on each endpoint.

The evaluation of the stakeholder feedback has revealed that the OSIRIS webtool with the five included endpoints is on the way to be a user-friendly, accessible and transparent webtool. It helps stakeholders to understand the testing strategies and supports the development of integrated and non-testing toxicity tests. Most of the workshop participants explained that they got new insight in ITS and combing toxicity information. They became more confident that ITS will be more implemented in the future. The OSIRIS webtool is not a stand-alone software and can be combined with additional software as "Derek", "ToxTree" and "ChemProp". As the participants' feedback has shown, an integration of this software would be appreciated. This would be very practical for users in the future. As a consequence the import and export of data and their format should be standardized, and a standard format (OECD, FoxMC, ...) be used. Another important point of improvements in the next months should be a more clear and comprehensive documentation and report for each endpoint, the parameters used and the guideline from the EU regulation in the webtool.

The OSIRIS webtool will be finished in autumn 2011. It will be important to secure that the webtool and its huge knowledge behind will be available for the stakeholders after the project is finished.

5. Annex: List of Participants

Participants of the Workshop

No	Name	Institution	City	Country
1.	Ade, Nadège	L'Oréal	Paris	France
2.	Dr. Aicher, Lothar	Swiss Centre for Applied Human Toxicology (University of Basel)	Basel	Switzerland
3.	Dr. Aldenberg, Tom	RIVM	Bilthoven	The Netherlands
4.	Alle, Katrin	Dialogik	Stuttgart	Germany
5.	Prof. Dr. Allgaier, Clemens	ACA-pharma concept GmbH	Leipzig	Germany
6.	Belkhiria, Sami	Dow Corning Europe SA	Seneffe	Belgium
7.	Benighaus, Ludger	Dialogik	Heidelberg	Germany
8.	Böhnhardt, Anna	Umweltbundesamt	Dessau	Germany
9.	Dr. Bringezu, Frank	Merck Serono	Darmstadt	Germany
10.	Dr. Bucior, Katarzyna	Dr. Knoell Consult GmbH	Mannheim	Germany
11.	Dr. Buist, Harrie	TNO	Zeist	The Netherlands
12.	Dr. Casalegno, Carlotta	British Union for the Abolition of Vivisection (BUAV)	London	United Kingdom
13.	Ciffroy, Philippe	EDF	Chatou	France
14.	Ebert, Ralf-Uwe	UFZ	Leipzig	Germany
15.	Dr. Ensenbach, Uwe	Clariant Produkte GmbH	Sulzbach	Germany
16.	Dr. Escher, Sylvia	Fraunhofer-Institut für Toxikologie und Experimentelle Medizin (ITEM)	Hannover	Germany
17.	Fritzsche, Eric	Umweltbundesamt	Dessau	Germany
18.	Dr. Geldsetzer, Felix	Bayerisches Landesamt für Umwelt	Augsburg	Germany
19.	Gomes, Charles	L'Oréal	Paris	France
20.	Gómez Contretras, Jeannette	RIVM	Bilthoven	The Netherlands
21.	Dr. Griem, Peter	Clariant Produkte GmbH	Sulzbach	Germany

No	Name	Institution	City	Country
22.	Dr. Hassold, Enken	Umweltbundesamt	Dessau	Germany
23.	Dr. Hennes, Christa	ECETOC AISBL	Brussels	Belgium
24.	Dr. Herzler, Matthias	Bundesinstitut für Risikobewertung (BfR)	Berlin	Germany
25.	Dr. Hoffmann-Doerr, Simone	Cognis GmbH	Düsseldorf	Germany
26.	Hynes, Geoffrey	Givaudan	Kent	United Kingdom
27.	Dr. Jenner, Karen	Givaudan	Kent	United Kingdom
28.	Jöhncke, Ulrich	Umweltbundesamt	Dessau	Germany
29.	Dr. Karl, Wolfgang	MC-Bauchemie Müller GmbH & Co. KG	Bottrop	Germany
30.	Dr. Kroese, Dinant	TNO	Zeist	The Netherlands
31.	Dr. Kühne, Ralph	UFZ	Leipzig	Germany
32.	Kunkel, Uwe	Universität Bayreuth	Bayreuth	Germany
33.	Lombardo, Anna	IRFMN	Milano	Italy
34.	Dr. Nendza, Monika	Analytisches Laboratorium	Luhnsted	Germany
35.	Dr. Neumann, Michael	Umweltbundesamt	Dessau	Germany
36.	Pauné, Eduard	SIMPPLE	Tarragona	Spain
37.	Poopal, Rama Krishnan	Bharathiar University	Coimbatore	India
38.	Dr. Reader, Stuart	Givaudan	Kent	United Kingdom
39.	Richard, Bertille	EDF	Chatou	France
40.	Dr. Richarz, Andrea	UFZ	Leipzig	Germany
41.	Dr. Roncaglioni, Alessandra	IRFMN	Milano	Italy
42.	Rorije, Emiel	RIVM	Bilthoven	The Netherlands
43.	Dr. Rücker, Thomas	ENVIRON Germany GmbH	München	Germany
44.	Prof. Dr. Schüürmann, Gerrit	UFZ	Leipzig	Germany
45.	Shemotyuk, Lidiya	CFCS-Consult GmbH	Essen	Germany
46.	Dr. Stock, Frauke	Umweltbundesamt	Dessau	Germany
47.	Dr. Szymoszek, Andrzej	Dr. Knoell Consult GmbH	Mannheim	Germany
48.	Teetz, Wolfram	Helmholtz Zentrum München	Neuherberg	Germany

No	Name	Institution	City	Country
49.	Thomas, Marie	L'Oréal	Paris	France
50.	Thüns, Sabine	Universität Bayreuth	Bayreuth	Germany
51.	Tluczkiewicz, Inga	Fraunhofer-Institut für Toxikologie und Experimentelle Medizin (ITEM)	Hannover	Germany
52.	van Elsacker, Paul	Federal Public Service (FPS) Health	Brussels	Belgium
53.	Dr. Voigt, Matthias	Universität Konstanz	Konstanz	Germany
54.	Wagner, Barbara	UFZ	Leipzig	Germany
55.	Dr. Welch, Jonathan	Charles River	Tranent	Scotland
56.	Dr. Wichard, Jörg	Bayer AG	Berlin	Germany