Training Needs for Toxicity Testing in the 21st Century: A Survey-informed Analysis[a]

Silvia Lapenna,[1b] Silke Gabbert[2] and Andrew Worth[1]

[1]Systems Toxicology Unit, Institute for Health and Consumer Protection, Joint Research Centre, European Commission, Ispra, Varese, Italy; [2]Department of Social Sciences, Environmental Economics and Natural Resources Group, Wageningen University, Wageningen, The Netherlands

Summary — Current training needs on the use of alternative methods in predictive toxicology, including new approaches based on mode-of-action (MoA) and adverse outcome pathway (AOP) concepts, are expected to evolve rapidly. In order to gain insight into stakeholder preferences for training, the European Commission's Joint Research Centre (JRC) conducted a single-question survey with twelve experts in regulatory agencies, industry, national research organisations, NGOs and consultancies. Stakeholder responses were evaluated by means of theory-based qualitative data analysis. Overall, a set of training topics were identified that relate both to general background information and to guidance for applying alternative testing methods. In particular, for the use of *in silico* methods, stakeholders emphasised the need for training on data integration and evaluation, in order to increase confidence in applying these methods for regulatory purposes. Although the survey does not claim to offer an exhaustive overview of the training requirements, its findings support the conclusion that the development of well-targeted and tailor-made training opportunities that inform about the usefulness of alternative methods, in particular those that offer practical experience in the application of *in silico* methods, deserves more attention. This should be complemented by transparent information and guidance on the interpretation of the results generated by these methods and software tools.

Key words: *alternative method, hazard and risk assessment of chemicals, qualitative data analysis, stakeholder requirements, training*.

Address for correspondence: *Andrew Worth, Systems Toxicology Unit, Institute for Health and Consumer Protection, Joint Research Centre, European Commission, Ispra, Varese, Italy.*
E-mail: *andrew.worth@ec.europa.eu*

Introduction

Over the past decade, attempts to describe and predict the biological and toxicological effects of chemicals have increasingly taken mechanistic considerations into account. For example, the report of the US National Academy of Sciences and National Research Council (NRC) "Toxicity Testing in the 21st Century: A Vision and a Strategy" (1) has attracted considerable attention and triggered a worldwide, controversial debate about the need for a "paradigm shift" in the hazard and risk assessment of chemicals (2–5). The NRC report was commissioned by the US Environmental Protection Agency (EPA), and envisions a transformation of the current way of conducting toxicity testing, from a system based on phenotypic responses in animals toward an approach that relies on a comprehensive understanding of the molecular mechanisms of toxicant effects in human cells and tissues. The move toward a more mechanistically-based risk assessment process implies the use of high-throughput and high-content screening (HTS/HCS) assays in mammalian (preferably human) cells, cell cultures and/or tissue surrogates (i.e. *in vitro* tests), while the results from these tests are interpreted by applying a range of computational 'systems toxicology' methods. The latter are used to map the cellular pathways that are perturbed by the toxicant, and to simulate the dynamics of the cellular response. In practice, these methods are intended to determine an acceptable concentration of the toxicant *in vitro*, which then serves as the basis for *in vitro* to *in vivo* extrapolations to establish levels of exposure that will not cause an adverse effect in exposed human beings.

The HTS/HCS programmes that are currently being developed (e.g. as part of the US EPA ToxCast™ programme, http://epa.gov/ncct/ toxcast/) will provide extensive and consistent data for generating new databases (e.g. ToxCastDB, http://actor.epa.gov/actor/faces/ToxMiner/Home.jsp; ACToR, http://actor.epa.gov/actor; ToxRefDB, http://actor.

[a]*Responsibility for the information in this article lies entirely with the authors. The views expressed are those of the authors and do not necessarily reflect the official opinions of their institutions.*

[b]*Current address: European Chemicals Agency, Annakatu 18, PO Box 400, 00121 Helsinki, Finland.*

epa.gov/toxrefdb), and computational tools for analysing and visualising complex datasets (e.g. the US EPA ToxPi tool, http://www.epa.gov/ncct/ToxPi). The availability of new databases should stimulate the development of novel quantitative structure–activity relationship (QSAR) models for predicting the binding of chemicals to specific molecular targets, which, in turn, may trigger toxicity pathways that ultimately lead to adverse outcomes in the organism.

Various terms are being used to capture variants of this general framework, including toxicity pathway (TP), mode-of-action (MoA), adverse outcome pathway (AOP) and source-to-outcome pathway approaches. Although these terms are generally accepted, internationally harmonised definitions of them do not exist (Table 1). The approaches are all based on the assumption that, following the initial interaction of a toxicant with a biological target, a limited set of key events (or intermediate effects) in a biological pathway can be used to describe and predict an adverse outcome. The source-to-outcome pathway approach differs slightly from the others by explicitly including, in addition to information needed for hazard characterisation, the exposure considerations that are critical for risk assessment (6). Irrespective of differences in their scope, these approaches are based on the principle that a detailed and exhaustive understanding of all possible molecular interactions and effects is not necessary for making safety decisions concerning a chemical. In practice, however, this principle is not always easy to apply, since it is often a matter of debate as to which key events are both toxicologically relevant and sufficient in a given decision-making context.

The ongoing discourse about the development and implementation of a new toxicity testing paradigm is closely linked to the requirements and opportunities for the regulatory implementation of alternative assessment methods. To allow for a broad dissemination of information on non-testing (i.e. *in silico*) methods, and to promote their wider use, several freely downloadable or online-accessible tools have been developed by the European Commission (EC) and international organisations. Examples include computational tools made freely available by the EC Joint Research Centre (JRC; 7; http://ihcp.jrc.ec.europa.eu/our_labs/predictive_toxicology/qsar_tools), by CEFIC LRI (http://ambit.sourceforge.net/), as well as tools developed under EC-funded projects, such as CAESAR (http://www.caesar-project.eu/), CADASTER (http://www.cadaster.eu), OSIRIS (http://www.osiris.ufz.de/), and OpenTox (http://www.opentox.org). Some of these tools are being implemented in broader platforms, such as VEGA (http://www.vega-qsar.eu) and Chembench (http://chembench.mml.unc.edu/). A major international project is the development of the Organisation for Economic Co-operation and Development (OECD) QSAR Toolbox (http://www.qsartoolbox.org/) for the mechanistically-based grouping of chemicals into categories intended for use in filling data gaps for chemical hazard assessment. The use of these computational tools, either as stand-alone prediction methods or as 'building blocks' within a testing strategy, requires both practical experience and background knowledge about the applicability domain of each method and its general strengths and limitations.

The various computational tools differ in the extent to which they support their predictions with information on model applicability and prediction reliability. The expertise to generate and use this information for decision-making might not be readily available in many companies. Therefore, the development of appropriate training opportunities could provide valuable guidance for improving such a knowledge base and for strengthening a more consistent and targeted use of alternative methods for

Table 1: Definitions used in mechanistic frameworks for predicting and assessing toxicity

Term	Definition	Ref.
Toxicity pathway (TP)	Perturbation of a normal biochemical pathway, from a molecular initiating event (initial point of chemical–biological interaction), to an effect at the cellular effect.	1
Source-to-outcome pathway	A cascade of measurable events, starting from the release of a chemical into the environment, to an adverse outcome.	13
Mode-of-action (MoA)	Sequence of events, starting with a molecular initiating event, and leading to an adverse effect at the level of whole organism/individual. This term does not (usually) include consideration of exposure or effects at higher levels than the individual.	14
Adverse outcome pathway (AOP)	The linear sequence of events, from the exposure of an individual to a chemical, through to an understanding of the adverse effect at the individual level (for human health effects) or population level (for ecotoxicological effects).	15, 16

the hazard and risk assessment of chemicals. It also implies the need to gain insight into the demand of stakeholders for such training. In a recent study, Gabbert and Benighaus (8) examined the perceptions of stakeholders from industry, regulatory agencies, research institutions and NGOs, on the use and implementation of integrated testing strategies (ITS). They concluded that members from research institutions emphasise additional training opportunities to be important for creating confidence and trust in the use of alternative methods in ITS for chemical hazard assessment. Although this indicates that training can contribute to a wider use of new, animal-free testing methods, certain elements remain unclear: a) what kind of training facilities would be considered relevant; b) what kind of training content stakeholders would prefer; and c) whether or not these preferences are related to particular alternative methods.

In order to address these questions and gain better insight into stakeholder preferences for training aimed at supporting a more consistent use of alternative methods, the JRC conducted in 2011 a survey on the evolving training needs of end-users of alternative methods in industry and scientific and regulatory organisations. This paper presents the outcomes of the survey, and discusses the implications for developing and improving training facilities.

Setup and Evaluation of the Survey on Training Needs

In total, 24 individuals from private enterprises, regulatory agencies, research organisations and NGOs, were contacted by e-mail and asked to respond to the following question:

Bearing in mind the state of the science as well as requirements and opportunities for regulatory implementation, what do you consider to be the main training needs in order to achieve a wider and more consistent use of alternative assessment methods (e.g. in silico*,* in vitro*,* in chemico*) in the context of a new ("21st Century") toxicity assessment paradigm based on Integrated Testing Strategies, including the application of the Mode-of-Action (MoA) and adverse outcome pathway (AOP) concepts?*

The participants were asked to respond by e-mail, and were encouraged to openly describe their (subjective) views about training needs as related to their work experience. Of the 24 persons contacted, 12 responded (a 50% response rate) — five respondents belonged to regulatory agencies, three to the chemical industry, two to national research organisations, one was a member of an NGO, and one worked as a private consultant.

The responses differed considerably with respect to length and detail. In order to allow for a systematic and coherent evaluation of the stakeholder replies, we applied qualitative data analysis, a theory-based social sciences approach that emerged in the 1960s (9, 10) and has become widely used in many different disciplines, such as health research, psychology, sociology, administration, and applied policy research (11, 12). Qualitative data analysis offers a set of principles for structuring and analysing any type of unwieldy and cumbersome textual data (e.g. from interviews, discussions, field notes, etc). Compared to a pure manual evaluation of stakeholder responses, qualitative data analysis facilitates the systematic identification and categorisation of textual content structures, thus reflecting opinions and perceptions. Mapping and comparing such content structures, in turn, allows for a transparent interpretation of textual data with regard to the underlying objective — which was, in our case, the identification of the stakeholders training needs on the use of alternative testing methods.

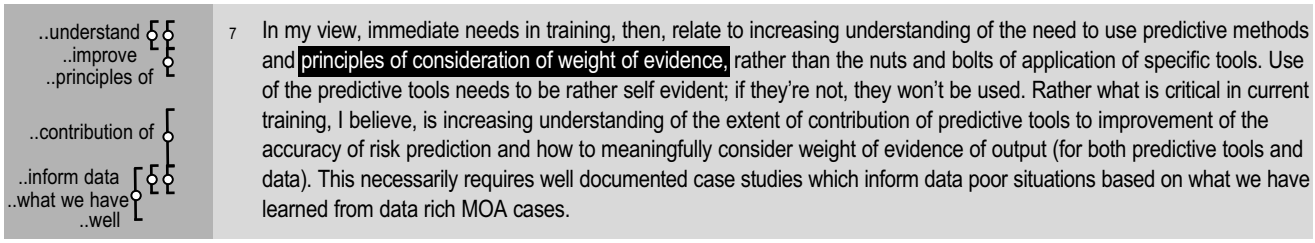
A crucial step for analysing qualitative data is the coding of the textual material. 'Coding' means that the analyst assigns categorical phrases to a text segment, such that its content is adequately described. Generally, a code can be attached to single words, to a part of a sentence, to complete sentences, or even to whole text passages. Coding can be done by either attaching single catchwords or by literally adopting the terms used in the text (8). The aim is to transform and condense textual data (i.e. the stakeholder responses) into analytical data which can then be further analysed. Clearly, coding is an interpretative process. If, for example, the content of a text segment is ambiguous, different codes can be attached to the same text fragment. Likewise, coded text fragments can be allowed to overlap, and equal codes can be assigned to different text fragments with identical contents.

In order to keep the analysis of stakeholder responses as transparent as possible, a two-step coding procedure was adopted. In the first step, all text passages with a content related to our question were selected. By using the qualitative data analysis software MAXQDA (http://www.maxqda.com/), we assigned literal codes to these text segments (an example is provided in Figure 1).

In the second step, the codes were clustered into categories with a similar content. These clusters represent 'core topics', i.e. themes that participants considered relevant when responding to our question. Where necessary, clusters were further divided into sub-categories.

Our analysis of the responses revealed five core topics to which stakeholders paid particular attention. Table 2 shows that most of their statements referred to the desired 'training content'. Besides this topic, stakeholders pointed to the methods that should be addressed in training, and also

Figure 1: The coding of survey responses



A screenshot of results generated with the MAXQDA 10 software (www.maxqda.com).

reflected on the expected training outcomes, the potential target audience and the desired mode of operationalisation.

Given the relatively small number of survey participants, a statistical analysis of code frequencies across topics would be of low informational value. Likewise, we do not claim that our findings provide an exhaustive mapping of stakeholder preferences. Nevertheless, the survey offers interesting insights into the range of themes that the participants considered useful in order to support the application of alternative methods. Therefore, in the following section we have looked into these topics in more detail, from a purely qualitative perspective, i.e. without ranking the topics.

Investigating Stakeholder Preferences for Training on the Use of Alternative Methods

The survey participants pointed to various options for operationalising the training courses. These include not only regular training events, workshops, lectures and laboratory visits, but also IT-supported options, such as webinars and video learning modules. As the target audience, they suggested members from regulatory agencies, as well as from industry, national research organisations and NGOs. Respondents belonging to research organisations and the chemical industry stressed the lack of acceptance of alternative methods by regulatory agencies, and referred to training as a potential means of improving the regulators' general understanding, and acceptance, of alternative methods.

Table 2: Core topics considered relevant by survey participants

Core topic	Code frequency
Operationalisation	12
Target audience	14
Expected outcomes	30
Methods addressed	52
Training content	87

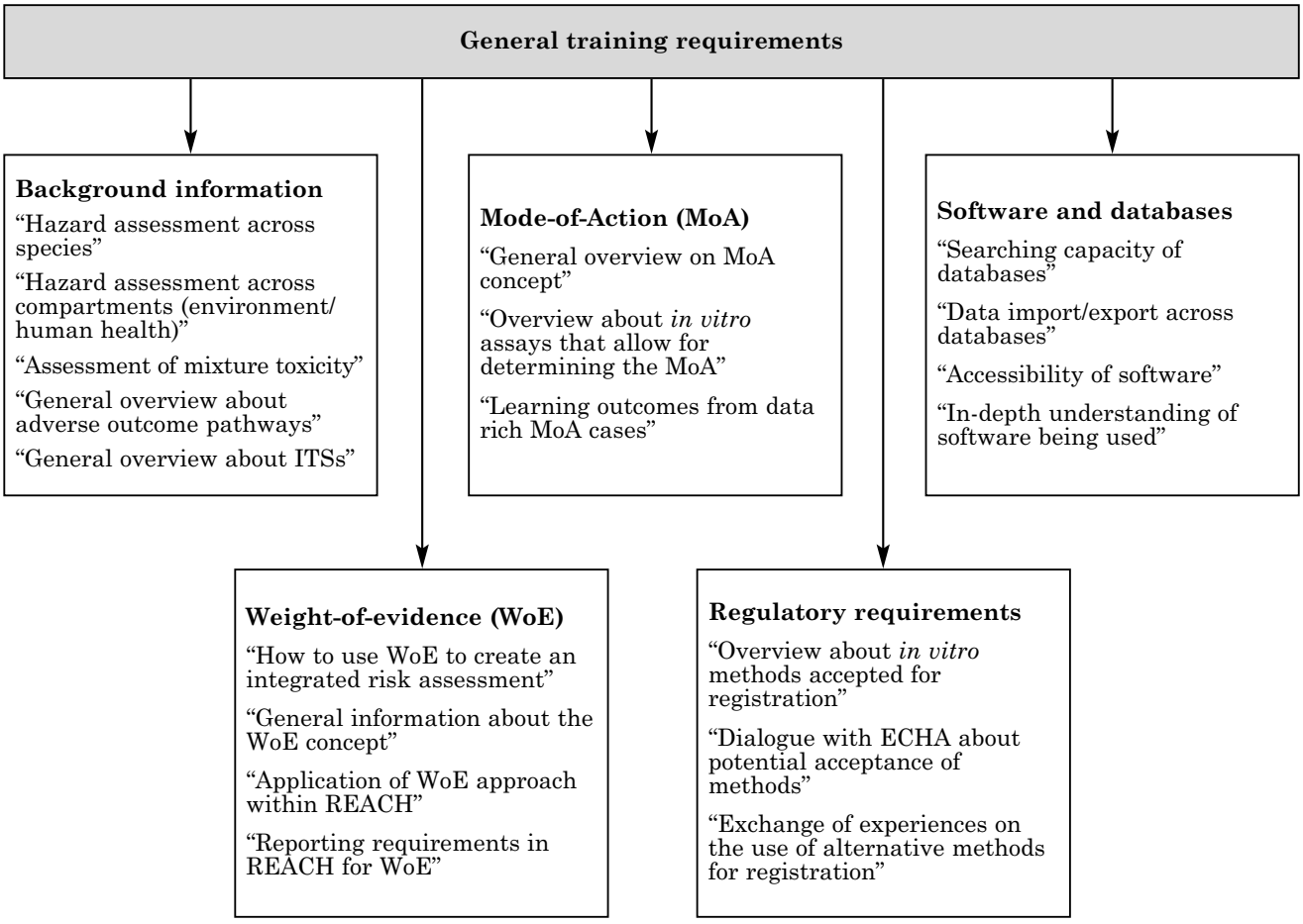
The numbers represent the frequencies of text codes belonging to a particular theme.

Consequently, stakeholders regarded progress toward the regulatory acceptance of alternative testing methods for the *reduction*, *refinement* or *replacement* of animal testing to be the key expected outcomes of increased training opportunities.

We noted that the stakeholders addressed a wide variety of alternative testing methods. These include high-throughput technologies, epidemiological assessments, *in vitro* methods, and non-testing methods (chemical grouping and read-across approaches, methods for deriving threshold-of-toxicological concern values, 'omics' technologies, and *in chemico* and *in silico* methods). Within the group of *in silico* methods, they mentioned specific QSAR tools as examples (Oncologic, ECOSAR, TEST, LAZAR, ToxPredict, CAESAR, TOPKAT, and DEREK). This illustrates that the stakeholders are well-informed about the diversity of existing alternative methods. However, responses referring to the preferred content of training indicated that the applicability of these methods and their overall usefulness for hazard and risk assessment is much less clear among the stakeholders.

The replies about training content could be clustered into two main categories. The first category comprises general training requirements for improving a person's knowledge base (Figure 2). Topics in this category provide background information that is relevant for the hazard assessment of chemicals in a regulatory context, but not relevant *per se* for applying alternative methods. Besides these general training needs, stakeholders pointed to several topics that are considered to facilitate the practical use of alternative methods (Figure 3). Figures 2 and 3 show the topics and sub-topics resulting from our qualitative data analysis in bold, whereas the coded text fragments are shown in quotation marks.

Figure 2: General training requirements



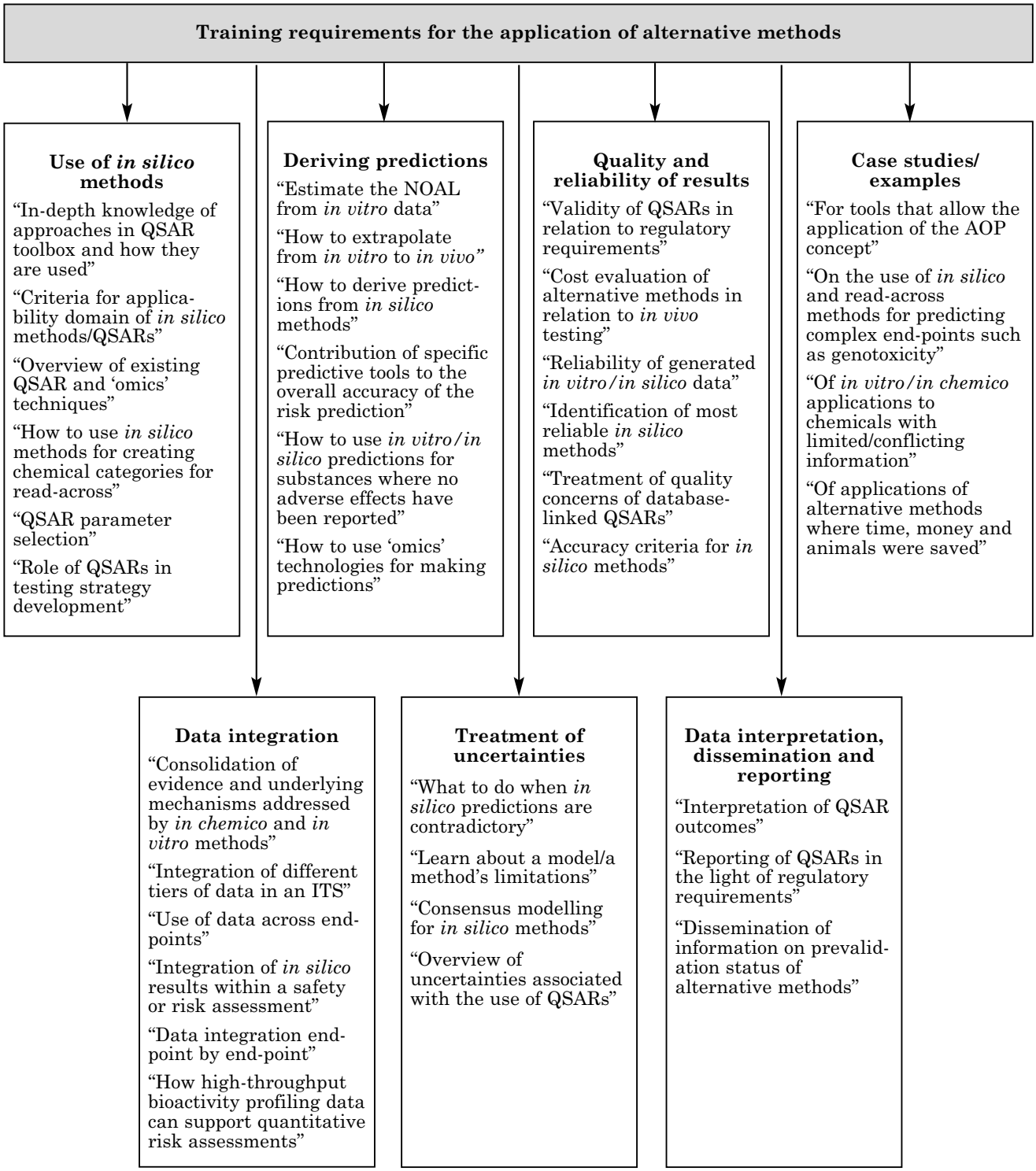
The topics and sub-topics resulting from the qualitative data analysis are shown in bold, whereas the coded text fragments are shown in quotes. The topics in this category provide background information that is relevant for the hazard assessment of chemicals in a regulatory context, but not relevant per se *for applying alternative methods.*

The general training requirements highlighted by stakeholders reflect their interest in gaining a more sophisticated insight into the key conceptual, operational and regulatory issues underlying chemical hazard assessment. Though general information for each of these topics is available and accessible, for example through guidance documents from regulatory bodies or the scientific literature, this information may still be too fragmented and dispersed. In addition, training opportunities offer an important possibility for stakeholder interaction, i.e. for communicating and exchanging experiences.

The analysis of training requirements associated with the application of alternative methods illustrates that stakeholders gave primary focus to methodological problems, such as, for example, the selection of relevant model parameters, the treatment of uncertainties, the integration of data, the evaluation of the quality and the reliability of method outcomes, and its interpretation for the hazard and/or risk assessment of a chemical. Furthermore, the stakeholders strongly emphasised the need for practical case studies — examples and exercises where the use and the limitations of a method can be demonstrated.

Besides identifying the topics considered important by the survey participants, we analysed whether the stakeholders prioritised particular types of alternative methods when reflecting on training needs. This was done by examining overlapping coded text segments categorised under "methods addressed" and "training requirements for applying alternative methods" (Table 2). We observed that stakeholder preferences for training are clearly associated with *in silico* methods and, within this group, with QSARs. More specifically, *in silico* methods were explicitly pointed out when

Figure 3: Training requirements for the application of alternative methods



The topics and sub-topics resulting from the qualitative data analysis are shown in bold, whereas the coded text fragments are shown in quotes.

stakeholders addressed the "use of *in silico* methods", "data integration", "treatment of uncertainties", and "quality and reliability of results". Apparently, the stakeholders considered the need for training with respect to these topics to be most relevant. This implies that the potential contribution of training courses to supporting the use of alternative methods would be greatest, if these practical issues were well-elaborated. *In vitro* methods, to the contrary, were predominantly linked to training options addressing "data integration" (e.g. the consolidation of evidence gained from *in vitro* and *in silico* and empirical data into a hazard/risk assessment) and "deriving predictions" (e.g. how to predict rodent carcinogens from an *in vitro* battery, or how to extrapolate from *in vitro* data to *in vivo* effects).

Discussion and Conclusions

This paper presents the outcome of a stakeholder survey performed by the JRC in order to learn about the training requirements that can support a more consistent use of alternative methods. Stakeholders from regulatory agencies, the chemical industry, research organisations, NGOs and consultancies, were asked to explain their training preferences, while reflecting on personal experiences within their work environment. The stakeholders pointed to a broad range of training topics, with a clear focus on the practical use and application of alternative methods. Though stakeholders addressed a range of different alternative methods, we observed a preference for training on the use of *in silico* approaches and *in vitro* methods. Interestingly, several of the stakeholders explicitly underlined the need for transparent proof-of-concept and hands-on exercises related to the practical use of *in silico* methods, rather than providing general background information on these methods. Training topics that were frequently specified included various aspects of *in silico* investigations, ranging from data integration, deriving predictions, the assessment and treatment of uncertainties, the interpretation and reporting of outcomes, and the development of case studies for practical applications of the tools.

Although our findings do not claim to be representative of all stakeholders and thus can provide only a preliminary snapshot of the training requirements, we conclude that there is a need for increased attention to be paid to the development of well-targeted training activities and media that support the use and implementation of alternative methods. Further research could, for example, increase the number of stakeholders involved, to enable the analysis of the similarities and differences of training requirements across stakeholder groups. Furthermore, the issue of how training needs can be combined and accommodated in targeted training solutions, so that they optimally respond to user needs, requires further attention.

In the development of future training resources, it will be important to draw on the experience of previously-held courses, such as those organised within the framework of completed EC-funded projects such as ANTARES (http://www.antares-life.eu/index.php), CADASTER (http://www.cadaster.eu), OSIRIS (http://www.osiris.ufz.de/), and OpenTox (http://www.opentox.org), the training courses on computational systems biology currently offered by the US-based Hamner Institutes for Health Sciences (http://www.thehamner.org), as well as ongoing projects such as those organised by the EU SEURAT-1 cluster (http://www.seurat-1.eu/). One lesson that has been gained by the authors in organising and providing training courses to support the use of alternative methods — which is confirmed by the outcomes of our survey — is that there is a need to focus not only on the technical use of alternative methods, but also on how to interpret the data provided by such methods. The latter is often more challenging, since the development of the tools has outpaced our understanding of how to use them for different regulatory purposes and in different contexts. In view of this, more documentation should be offered along with the software tools, to help with the interpretation of the results obtained from them, and thus provide users with convenient and relevant guidance to justify the *in silico* approach used and the use of the predictions obtained.

Besides the need for training, stakeholders pointed to other issues that, in their view, still hamper a coherent implementation of alternative methods at the international level. Several of the participants in the survey emphasised the disparity between research activities and regulatory applications of alternative methods, which forms a significant hurdle to the broader use of the methods. Likewise, the lack of, or uncertain, regulatory acceptance of alternative approaches by the regulatory authorities (especially outside a single jurisdiction such as the EU) makes it more attractive for industry to follow traditional testing strategies (because they are less risky).

Finally, a challenge lies in reconciling and possibly combining the use of freely available tools (which are becoming increasingly available) with commercial tools (for which there is a longer history of use in many organisations). There is, sometimes, a perception that the latter are more reliable, more user-friendly or somehow more mature than the former, although one could also argue that it is simply a matter of time before the former are seen to 'catch up'. The authors think that this challenge could be partially addressed by encouraging the providers of proprietary software tools to publicly release some of their models, such as restricted versions of their

products, and/or parts of their structure–activity datasets. In addition, in the interests of transparency, the providers could submit QSAR Model Reporting Formats (QMRFs) for their models to the JRC QSAR Model Database, a web-based inventory of harmonised documentation for QSAR models (http://qsardb.jrc. ec.europa.eu/qmrf).

To conclude, the wider and more consistent implementation of alternative methods represents a multifaceted challenge, but it depends partially on the need to narrow the (arguably increasing) gap between the rapid development of new alternative methods and the experience and confidence of stakeholders in applying these methods. Thus, the further uptake of alternative methods as parts of testing strategies that reduce or eliminate the need for animal testing, depends not only on the careful analysis of the requirements of different user groups, but also on the exploration of how to provide optimal and tailor-made training solutions. Specific educational programmes on systems toxicology will have to be developed at universities and research organisations, in order to effectively prepare and support the next generation of researchers and risk assessors.

Acknowledgement

This work was partly funded by the European Union FP6 integrated project OSIRIS (http://www.osiris-reach.eu/).

Received 24.05.12; received in final form 25.11.12; accepted for publication 26.11.12.

References

1. NRC (2007). *Toxicity Testing in the 21st Century: A Vision and a Strategy*, 196pp. Washinton, DC, USA: National Academies Press.
2. Collins, F.S., Gray, G.M. & Bucher, J.R. (2008). Toxicology. Transforming environmental health protection. *Science, New York* **319**, 906–907.
3. Andersen, M.E., Al-Zoughool, M., Croteau, M., Westphal, M. & Krewski, D. (2010). The future of toxicity testing. *Journal of Toxicology & Environmental Health, Part B* **13**, 163–196.
4. Hartung, T. & McBride, M. (2011). Food for thought... On mapping the human toxome. *ALTEX* **28**, 83–93.
5. Boekelheide, K. & Andersen, M.E. (2010). A mechanistic redefinition of adverse effects — a key step in the toxicity testing paradigm shift. *ALTEX* **27**, 243–252.
6. Georgopoulos, P.G. (2008). A multiscale approach for assessing the interactions of environmental and biological systems in a holistic health risk assessment framework. *Water, Air & Soil Pollution Focus* **8**, 3–21.
7. Pavan, M. & Worth, A.P. (2008). Publicly-accessible QSAR software tools developed by the Joint Research Centre. *SAR & QSAR in Environmental Research* **19**, 785–799.
8. Gabbert, S. & Benighaus, C. (2012). *Quo vadis* integrated testing strategies? Experiences and observations from the work floor. *Journal of Risk Research* **15**, 583–599.
9. Glaser, B. & Strauss, A. (1967). *The Discovery of Grounded Theory: Strategies for Qualitative Research*, 284pp. Piscataway, NJ, USA: Aldine Transaction.
10. Charmaz, K. (2006). *Constructing Grounded Theory: A Practical Guide Through Qualitative Analysis (Introducing Qualitative Methods Series)*, 224pp. Thousand Oaks, California, USA: Sage Publications Ltd.
11. Hallberg, L.R-M. (2006). The "core category" of grounded theory: Making constant comparisons. *International Journal of Qualitative Studies on Health & Well-being* **1**, 141–148.
12. Ritchie, J. & Spencer, L. (1994). Qualitative data analysis for applied policy research. In *Analysing Qualitative Data* (ed. A. Bryman & R.G. Burgess), pp. 173–194. London, UK: Routledge.
13. Krewski, D., Westphal, M., Al-Zoughool, M., Croteau, M.C. & Andersen, M.E. (2011). New directions in toxicity testing. *Annual Review of Public Health* **32**, 161–178.
14. US EPA Risk Assessment Forum (2005). *Guidelines for Carcinogen Risk Assessment*, 151pp. Washington, DC, USA: US Environmental Protection Agency. Available at: http://www.epa.gov/raf/publications/pdfs/CANCER_GUIDELINES_FINAL_3-25-05.PDF (Accessed 26.11.12).
15. Ankley, G.T., Bennett, R.S., Erickson, R.J., Hoff, D.J., Hornung, M.W., Johnson, R.D., Mount, D.R., Nichols, J.W., Russom, C.L., Schmieder, P.K., Serrrano, J.A., Tietge, J.E. & Villeneuve, D.L. (2010). Adverse outcome pathways: A conceptual framework to support ecotoxicology research and risk assessment. *Environmental Toxicology & Chemistry* **29**, 730–741.
16. Schultz, T.W. (2010). Adverse outcome pathways: A way of linking chemical structure to *in vivo* toxicological hazards. In *In Silico Toxicology: Principles and Applications* (ed. M.T.D. Cronin & J.C. Madden), pp. 346–371. Cambridge, UK: The Royal Society of Chemistry.