PEST COMMITTEE MEETING OF 7 JUNE 2018

EU AUTHORISATION PROCEDURE FOR PESTICIDES - EFSA OPINION ON DRAFT ASSESSMENT REPORTS AND ECHA CLASSIFICATION OF ACTIVE SUBSTANCES

PREPARATORY QUESTIONS

In the context of the PEST Committee meeting of 7 June, a public hearing with four invited parties on "*EU authorisation procedure for pesticides - EFSA opinion on draft assessment reports and ECHA classification of active substances*" will take place.

To prepare for this public hearing, political groups have submitted the following questions by Members of the PEST committee. We would kindly ask the invited parties to provide written answers to these questions, which address many of the topics at stake, by 5 June 2018, 12.00 h to the PEST Committee Secretariat: pest-secretariat@europarl.europa.eu.

For Mr. Christopher Portier:

46. In an open letter to the European Commission, signed by you and a number of other scientists, you have criticized the use by EFSA of industry studies by saying that "compliance with guidelines and Good Laboratory Practice does not guarantee validity and relevance of the study design, statistical rigor and attention to sources of bias." (https://www.efsa.europa.eu/sites/default/files/Prof_Portier_letter.pdf)

    Could you please elaborate on this and the potential problems with the GLP guidelines and their implementation?

I seldom have concerns about the quality of the regulatory studies submitted by industry for pesticide registration. The quoted phrase comes from our 2015 paper in the Journal of Epidemiology and Community Health and is in the context of our disagreement with EFSA that "data from studies published in the peer reviewed literature should automatically receive less weight than guideline studies". Regulatory studies are generally limited to specific dosing regimens, specific strains and specific species and seeing no positive responses in these studies does not guarantee safety. Non-regulatory studies are more flexible, using different strains, different species and different dosing regimens that may be more indicative of the likely response to be seen in humans. Giving less weight to a peer-reviewed study that saw different findings from a regulatory study simply because it is not a GLP study is not good scientific practice nor good public health policy. Finally, let's be certain we all understand why GLP is used in these studies; GLP provides a framework for being certain that the laboratories that did these studies (generally under contract to a corporation that wishes to profit from the product being tested) have actually done what is required. Not having GLP does not mean a study has flaws; it means some aspects of the conduct of the study may be difficult to evaluate.

What is the nature of the research institutions (laboratories) that carry out the industry studies?

In general, these are highly qualified laboratories that have highly qualified scientific staff. They have extensive experience in conducting the studies. For example, some of the contract laboratories used by industrial groups for safety evaluations are also used by the US National Toxicology Program for safety evaluations. Thus, the data produced by these laboratories is generally of high quality. However, having high quality data is not enough; one most analyse and interpret these data appropriately and generally that is done by the company contracting for the study or an independent contractor. This is where bias can enter into an evaluation.

Who certifies them and how is this financed?

This varies. As a general rule, most laboratories doing chronic animal carcinogenicity studies are AAALAC certified (AALAC is the Association for Assessment of Laboratory Animal Care). In addition, the organizations contracting the work routinely send out teams of scientists to check on aspects of how the work is conducted in these laboratories to check that protocols are being followed, etc. In general, the sponsoring authority pays for the review of the laboratory by outside reviewers. For AALAC certification, the laboratory generally pays for the evaluation.

What are the controls of the compliance with the GLP standards?

Again, this varies. I am not certain how this is done in the EU, but in the US there are records requirements that allow auditors to assess compliance with GLP. In addition, as mentioned above, most groups sponsoring GLP-based laboratory studies require inspections by outside scientists.

Do you need more studies to assess a substance?

This always depends on what you already know. For cancer, you can certainly usually find something new to study on any given agent at any time. The mass of literature on ionizing radiation and dioxins can attest to that reality. However, for determining carcinogenicity, you can certainly manage to make a reasonable decision from the data available at the time of registration. As new data appears, it can certainly alter the original decision and all available data should be considered.

47. Can you please explain the difference between IARC and JMPR?

It is, of course, better for you to address these questions to IARC and JMPR. I have served on several IARC Working Groups and have never served on a JMPR review. Thus, my comparison of the two groups is in one case based upon experience and in the other based upon simply reading documents. I have served on evaluations for the Joint Committee on Food Additives (JCFA) which, to my understanding, has similar processes to JMPR. So, from a very high level, here are my impressions of the major differences between the two:

IARC is focused on only cancer and all exposures that could lead to cancers. JMPR is focused on all endpoints but for only pesticide residues. Members of IARC Working Groups (WG) are generally well-known experts in their fields with extensive publication records and generally come from academia or government research laboratories. JMPR reviewers are more likely to be from national regulatory agencies. An IARC WG is interested in answering the question of whether a substance could cause cancer to humans; this includes in occupational settings as well as general public exposures. JMPR reviewers are tasked with

setting a guidance level for where they believe a pesticide residue in food poses minimal to no risk to humans. IARC WGs generally contain more epidemiologists that JMPR reviews (sometimes JMPR reviews do not include epidemiologists) and IARC WGs generally address fewer agents in a single meeting than do JMPR reviews.

48. Please explain the IARC policy on conflicts of interest.

That is a topic for IARC to explain.

49. Please explain how IARC deals with experts of its committees being members or employees of other organisations.

All IARC WG members are employees of other organizations. IARC requires participants to fill in a Declaration of Interests form (http://monographs.iarc.fr/ENG/Meetings/vol112-doi.pdf) then evaluates that form to determine if the participant has a serious potential conflict of interest. Depending on the severity of the potential conflict and the expertise being provided by the expert, IARC may exclude them from the WG or allow them to attend the WG meeting as a non-voting expert (Special Expert). In some cases, where IARC has learned that an expert has a potential conflict after forming a WG, they have removed that expert from the WG but retained them as a Special Expert.

50. If IARC had assessed the same studies as EFSA/ECHA, would it have come to the same conclusions?

I assume this question pertains to glyphosate. In my opinion, if IARC were to convene a WG now to review all of the available information, they would again determine glyphosate is Class 2A, a probable human carcinogen. This would include the analyses of data in the Appendices to the Greim et al. (2015) paper provided there is sufficient time to do these analyses (even though certain aspects of these studies are still publicly unavailable) and provided the WG determines there is sufficient detail for an assessment. I note that evaluations made by IARC WGs are the opinions of the WG and that IARC is there to support the systematic review of the scientific evidence and the consensus evaluations of the WG.

51. Can IARC please comment on the Reuters reporting about important differences in the draft IARC report and the final IARC report, especially with regard to the deletion of exonerating content?

I do not speak for IARC hence I cannot answer this question on their behalf. However, I have extensive experience in reviewing scientific literature for the purpose of determining safety; I ran two US federal agencies whose missions included evaluation of scientific evidence, I have served in EPA, WHO, IARC and other panels doing such evaluations and I have managed the evaluation of the health and safety on electric and magnetic fields from powerline frequencies on behalf of the US Government. In every case, evaluation language and conclusions evolve over the course of the evaluation. Someone has to draft the initial discussion of the science and the evaluations. Others then come in, comment on the text and conclusions and the document evolves. The fact that early drafts of a safety evaluation do not match the final evaluation should be expected. In addition, because many of these early drafts are pre-decisional, most agencies consider them confidential and generally do not share them with the public, even when requested.

52. One of the key elements of the PEST committee mandate is to investigate conflicts of interest within the EU agencies and the national authorities. Can Prof. Portier provide an

explanation of how conflicts of interest are handled within the WHO IARC assessment process?

I cannot explain how conflicts of interest are handled within IARC – that is again a question best asked of IARC. However, let's be certain we have a good definition of "conflict of interest". The US National Academies of Sciences (http://www.nationalacademies.org/coi/bi-coi_form-1.pdf) define a "conflict of interest" as "any financial or other interest which conflicts with the service of the individual because it (1) could significantly impair the individual's objectivity or (2) could create an unfair competitive advantage for any person or organization". They go on to specify that "The term "conflict of interest" means something more than individual bias. There must be an interest, ordinarily financial, that could be directly affected by the work of the committee." And, "Conflict of interest requirements are objective and prophylactic. They are not an assessment of one's actual behaviour or character, one's ability to act objectively despite the conflicting interest, or one's relative insensitivity to particular dollar amounts of specific assets because of one's personal wealth. Conflict of interest requirements are objective standards designed to eliminate certain specific, potentially compromising situations from arising, and thereby to protect the individual, the other members of the committee, the institution, and the public interest. The individual, the committee, and the institution should not be placed in a situation where others could reasonably question, and perhaps discount or dismiss, the work of the committee simply because of the existence of conflicting interests."

In particular, what are its disclosure requirements for the employment of WHO IARC experts within other organisations (public and private)?

See answer to question 49.

53. Can Prof. Portier explain the drafting process in the WHO IARC and, more specifically, the differences between the draft report and the final report evaluating the carcinogenicity of glyphosate as reported by Reuters (October 19, 2017)?

Literature for an IARC evaluation is collected by IARC and the WG members and divided into 4 separate areas (exposure, epidemiology, cancer in experimental animals and mechanisms). Members of the WG with expertise in these areas are tasked with drafting a review of specific literature in their area of expertise. These initial reviews are reviewed by at least one other member of the WG, changes are made and the final draft versions are brought together as the initial draft of the WG report. During the actual WG meeting, the scientists initially work in one of the four areas of science described above to discuss the quality and meaning of individual studies and then to make preliminary decisions regarding the strength of the evidence in each area using guidance from the IARC Preamble. The whole WG meets occasionally during the 8-day meeting to discuss progress and discuss recommendations coming out of the individual groups. The individual groups bring forward suggested evaluations that are subject to change before the final evaluation. The final evaluation is the consensus of the entire WG. Since many experts have scientific credentials in multiple areas, it is not unusual for draft evaluations for a specific subgroup to be changed when the entire WG is discussing the overall findings. For more details, I suggest you ask IARC.

54. Could Prof. Portier explain how many hours the WHO IARC worked on its evaluation of the carcinogenicity of glyphosate, and how many staff were involved in the assessment?

No, I cannot. The number of hours spent by each WG member prior to the WG meeting cannot be known. They need to read the literature, read the draft documents and come prepared to discuss the science. During a WG meeting, there are usually eight full days available for discussions and evaluations. The full list of participants in an IARC Monograph Meeting, including IARC staff, are available in the Monograph.

55. The Commission considers the WHO IARC evaluations to play an important screening assessment of the carcinogenic potential of agents, but that this should not be compared with the more comprehensive hazard assessment conducted by national competent authorities and EFSA, which are designed to support the EU's regulatory process for pesticides. In response to this, can you explain why you adjudge the WHO IARC classification to be by far the more credible assessment of glyphosate when compared with the renewal assessment report by the German Federal Institute for Risk Assessment (BfR)?

My concerns and those of my colleagues are explained in the multiple letters sent to EU officials and publication regarding the scientific limitations of the evaluation conducted by the BfR. Rather than repeating each argument and the literature associated with that argument, I would refer you to our previous letters and publications. However, I will summarize.

The evaluation of the epidemiological data is inconsistent and subjective. The authors of the document give greater weight to a single cohort study with inappropriate justification. This over emphasis on a single study is not consistent with acceptable scientific practice. In addition, the authors place far too much emphasis on individual studies being positive or negative ignoring the general trends in the data and the supportive role of different epidemiology studies showing consistent findings. They afford very little weight to the finding in the meta-analysis of a significant positive associative. Finally, EFSA refers to the findings as "very limited", a category not listed in the CLP Guidance (FYI, the EFSA review also ignored the defined categories from the CLP Guidance).

The evaluation of the animal carcinogenicity data is weak, relying on evaluations provided by the applicants without any additional analyses to verify the findings. They almost entirely rely upon findings using pair-wise comparisons when it is well known that trend tests provide a greater ability to identify a positive finding (when one really exists) without an increase in the false-positive rate. Only after the IARC review occurred did BfR reanalyze a subset of the data using a trend test and then generally discarded these positive findings since there were no pair-wise positive finding, or there were no non-decreasing dose-response, or the data were in the range of the historical control data, or the data on tumor precursors (usually undefined) were not significant. In essence, they created an almost impossible hurdle for a study to be positive and on a study-by-study basis excluded findings. In comparing across studies, they failed to account for significant study differences comparing Wistar rat findings with Sprague-Dawley rat findings, comparing CD-1 mouse findings with Swiss mouse findings, comparing 18-month studies with 24-month studies, accepting the presence of a virus in one study for which there was no evidence then changing the reason to reject the study when challenged regarding the virus, etc. They also disregarded positive findings in males if there were no equivalent findings in females. They consistently overrode a positive finding in an evaluation against the concurrent control using an inappropriate procedure using the historical controls. They failed to appropriately use historical controls to evaluate rare tumours.

In essence, in order to be considered a positive finding in the BfR analysis, a tumor must: 1) have both a positive trend test and a positive pairwise comparison at a dose below 1000 mg/kg/day (even though OECD guidelines they cite state otherwise and none of the studies show any concerns about overt toxicity at their highest doses); ***AND*** *2) all studies in the same species must show the exact same findings (even though we know study findings can differ for a variety of reasons such as study duration, strain, genetic drift, etc.);* ***AND*** *3) females must have the same results as males (even though this is not a routine observation since cancer rates and sensitivities differ across sexes; would we reject a positive epidemiology study in males if we had a negative study in females?);* ***AND*** *4) there must be clear, significant dose-response of pre-malignant lesions (even though most guidelines refer to a positive findings in pre-malignant lesions as strengthening the overall finding and do not refer to a lack of pre-malignant lesions as reducing the importance of a positive finding);* ***AND*** *5) the response must be outside of the historical control range (the IARC Preamble, developed by a team of prominent independent scientists, rejects this approach and OECD references papers that reject this approach in their recommendations for the use of historical controls);* ***AND*** *6) the responses must increase or stay level as the dose increases (non-monotonic dose-response; random drops below response at a lower dose are not acceptable). This creates an almost impossible hurdle for a study to get through for a finding to be positive. Thus, they are able to dismiss all positive findings in each individual study.*

Finally, their subjective evaluation across studies is inappropriate and a more objective evaluation should have been used such as pooled analyses or meta-analyses.

Many of the short-comings in the analysis of the animal carcinogenicity data are present in the analysis of the mechanistic data (historical controls used inappropriately, no trend analyses, etc.). In addition, there is a selective bias to exclude positive findings from peer-reviewed studies and accept negative findings from industry studies regardless of the fact that studies are measuring different, non-complimentary outcomes using different species and/or cell lines, different dosing protocols, differing exposure periods, etc.

The IARC review had none of these short-comings.

56. Could Prof. Portier please explain the difference between the WHO IARC assessment and the Joint FAO/WHO Meeting on Pesticide Residues (JMPR) assessment?

See my answer to question 47.

Both are specialised agencies within the WHO but they have come to different conclusions regarding the carcinogenicity of glyphosate.

They have not reached differing conclusions. IARC has concluded that glyphosate is "probably carcinogenic to humans, Class 2A" whereas the JMPR has concluded that "glyphosate is unlikely to pose a carcinogenic risk to humans from exposure through the diet". These conclusions are not in opposition; a chemical can be a probable human carcinogen without posing a risk to humans at levels currently seen in food. That said, the JMPR review suffers from all of the same scientific shortcomings as the BfR review and, as such, should not be relied upon.

57. If the WHO IARC had been given an opportunity to scrutinise the industry generated data compliant with OECD test guidelines and GLP when preparing its evaluation on glyphosate, would that have changed its final assessment?

See my answer to question 50. I will also point out that the IARC Preamble precludes the use of scientific data that is not publicly available. Thus, even if the IARC WG had seen all of the studies provided by industry, they would not have used these studies without the ability to make them public. This one rule guarantees a greater degree of transparency in the IARC review as compared to others.

58. In your analysis of the raw data you identified some new incidents of tumors in animal experiments that had not been reported at all in the summaries of the studies that were provided by the applicant. Would that change the outcome of the analysis?

Yes. There were positive tumor findings in multiple studies (e.g. malignant lymphomas, hemangiosarcomas, kidney tumors, skin keratoaconthomas) that strengthen a finding of sufficient evidence in animals. In addition, the large number of positive tumor findings from my analyses result in a miniscule probability that all of the observed significant increases in trends are the result of random false positive errors.

59. Do you think that classification of an active substance can and should take into account the data on the effects of the formulation?

Yes. For two reasons: the epidemiology and human accidental exposures will only be to the formulation (these are relevant data in the target population) and because a comparison of similar studies can provide insight into glyphosate toxicity. For example, given two studies with the exact same protocols where one uses only pure glyphosate and the other uses a glyphosate formulation but study subjects get the same exposure to glyphosate, there are five possible outcomes: 1) if both studies show no increased hazard as a function of exposure, glyphosate has no effect; 2) if the glyphosate-alone study is positive and the study using the formulation is negative the interpretation is difficult with regard to glyphosate toxicity; 3) if both are positive at approximately the same hazard level glyphosate is the cause; 4) if both studies are positive with the formulation showing greater toxicity, glyphosate is positive and the non-glyphosate materials in the formulation enhance that effect; 5) if both are positive but the glyphosate effect is much larger, glyphosate is positive and the non-glyphosate materials in the formulation reduce the effect. Thus, in almost every case, a clear interpretation relative to pure glyphosate is enhanced by the evaluation including the formulated product. The evaluation becomes much more difficult when you have hundreds of studies and multiple formulations, but can still be done.

60. What would you propose to regulators to do in order to revise the assessment of glyphosate and other active substances in Europe?

Follow their guidelines as written, use sound scientific and statistical procedures, have outside experts formally evaluate each aspect of the assessment (epidemiology, animal experiments, mechanistic experiments, cancer, reproductive effects, etc.) and make certain there are no conflicts of interest as defined above. I realize a formal external review for each compound evaluation is to some degree impractical, but for a ubiquitous exposure like glyphosate, this should be mandatory.

61. You have been attacked for having conflicts of interest. What is your response?

At the time of my participation in the IARC review, it is my opinion that I had no conflict of interest (see the definition above). My only position at the time of that review was a 20% appointment with the US Environmental Defense Fund to advise them on issues relating to air pollution, hydraulic fracturing, climate change and testing methods for evaluating

chemical toxicity. EDF had (at that time) and still has (at this time) no opinion on the carcinogenicity of glyphosate. However, EDF is an environmental non-governmental organization and, as such, IARC was concerned enough to have me serve as an Expert Scientist where I was not allowed to join in any discussions on the evaluations nor could I write any text for the monograph; my role was to simply advice the WG on scientific issues relating to the studies they were reviewing. Currently, I am serving as an expert witness in several law suits in the US and I would currently have a conflict of interest for any independent, committee-based evaluations relative to the carcinogenicity of glyphosate.

For all experts:

62. EFSA and ECHA have shown great efforts to address possible conflicts of interests. However, ever stricter rules will likely result in difficulties to recruit qualified staff for the agencies.

I find this notion offensive both personally and on behalf of all my colleagues working in government positions worldwide tasked with the evaluation of the toxicity of chemicals and other agents and I disagree with the assumption that only scientists with conflicts of interest are qualified. Having spent 34 years in government service and run two large government public health agencies, I can say, without doubt, that the scientists I worked with or who worked for me were qualified and without conflict of interest. My colleagues and I were barred from having any outside employment without a careful review for conflict of interest. My colleagues and I were not allowed to accept any funding from any industry that in any way related to the toxicity of chemicals including travel funds. My colleagues and I could not even accept a dinner invitation if the total value of our portion of the cost of the dinner exceeded $25 unless we paid for it ourselves. My colleagues and I were required to submit annual reports on all stocks and related holdings that were evaluated for conflict of interest; if a conflict were found, we were required to sell the stock. In addition, as Director, my annual financial report was summarized and made public. Yet I, like virtually all of my colleagues, continued to work hard on the issues we were faced with and, in my opinion, did a good job of getting to the right answers. Simply put, it is not possible to have it both ways. The agencies we entrust to protect the public and workers from exposures in their environment cannot have any real or perceived conflict of interest if the agency and their evaluations are to be trusted. I never had a problem recruiting scientists who would follow the rules imposed upon those trusted with the public's health and any scientist who felt those rules were too strict did not belong working for such an agency.

How can a fair balance be achieved between ensuring sufficient independence of the agencies and ensuring that the staff is sufficiently familiar with and experienced in their fields of expertise?

These two concepts are unrelated and not well-described. What do agencies need to be independent of? They need to be independent of inappropriate political and business influence. This is achieved by ensuring the laws governing the roles of the various players involved in the evaluations of environmental exposures are clear and precise. This involves not only having strong policies regarding conflicts of interests by employees, but rules relating to transparency in the evaluations and full disclosure of all communications regarding a specific evaluation. Expertise in a field is obtained through hard work and education. This is achieved and maintained by providing sufficient resources for continuing education and opportunities to interact with peers at open scientific meetings.

What measures have proven to be best practices?

They are outlined above: draw very clear lines that are easily understood and easily managed. Because we are focused on pesticides, employees should be banned from attending any meeting funded substantially by the pesticide industry unless the funds to attend that meeting are coming from agency budgets and there is a clear advantage to public health for the employee attending the meeting. Employees should not hold outside jobs related to their areas of expertise with the exception of un-related activities like teaching or research positions at universities (and then, only for very limited compensation). Stocks, bonds and related assets should be reviewed annually for conflicts of interest with greater scrutiny for senior managers. Finally, there should be restrictions for a fixed period of time for employees who either leave industry to work for an agency or leave an agency to work for industry (revolving doors).

Do you consider further improvements necessary and if so, how could the systems in place be further strengthened without compromising the agencies' ability to work?

I do not have sufficient understanding of the systems in place to comment on this issue.

63. In recent political and public debates, the presence of "data gaps" identified by EFSA have caused discussions about the validity of Commission approvals. However, data gaps do not necessarily mean that authorisation procedures cannot be positively concluded. Do you have recommendations on how to deal with data gaps in political and public communication so that they do not undermine the trust in the approval system?

Decisions have to be made and there will always be data gaps. In determining whether to approve a pesticide for use, the agency must evaluate the importance of the data gaps and act accordingly. In my opinion, an agency needs clear guidelines categorizing the importance of a data gap and the actions to be taken once a data gap is classified. I can envision a category in which a data gap is so significant that a decision cannot be made without having that data (call this Category 1). In this case, the guidelines would state that the evaluation process should be put on hold and the data must be obtained if the process is to be completed. Guidance for materials already in commerce would also need to be provided for when the product is allowed to remain in commerce (say Category 1b) or when it is removed (Category 1a). This must be clearly articulated to the public and removal must be one of the options. Other categories would involve less critical data gaps that must be filled in a fixed period of time (say Category 2), but would come with a temporary approval for use (Category 2b) or exclusion from use (Category 2a). Finally, there would be data gaps that would be mentioned as a guide for scientists looking for interesting projects that might eventually have regulatory importance (Category 3). In all cases, it would be important to communicate the impact the requested science would have on the outcome of the review. If that is unknown, then why is this a data gap for a regulatory decision? Of course, greater thought needs to be put into how to define these different categories, but the general idea is that categories help to provide a communications tool that spells out the fact that some gaps are very important and others less so.

64. Part of the SAM's mandate for the review of the PPP authorisation system is to assess options for arbitration in case of diverging assessments by different competent authorities. In this regard, the biocidal products regulation has been mentioned as a possible positive example. Could you explain how arbitration is handled under the

biocidal products regulation and how this is currently done for plant protection products? Are there other best practices that would be applicable to the PPP authorisation? In your opinion, would it make sense to extend arbitration also to scientific bodies other than competent authorities, particularly with a view to increase public trust in the soundness of the authorisation system?

I have no comment. However, referring to the previous question, "plant protection product" or "PPP" is not a good term for communication to the public since they more readily understand the term pesticide or herbicide. Would we refer to vaccines as "human immune system stimulators"? For a public health agency to refer to a pesticide as a PPP suggests to the public that they are trying to hide the true nature of a potentially harmful substance.

65. Could you elaborate the consequences for the independence of the risk assessment if EFSA would be exclusively responsible for both commissioning required studies for the authorisation procedure and for carrying out the risk assessment? Could that lead to a higher direct exposure of EFSA to the applicant as well as to a less stringent peer-review of the assessment of the application?

This is a hard question to answer in a simple manner. Since these agencies are tasked with protecting the health of the public, I would first ask myself what the public expects. That answer is simpler. They expect an assessment to be objective, devoid of any undue influence and protective of public health.

The problem is not so much with who produces the data but is more closely aligned with who owns, analyses and interprets the data. In most cases, given careful rules for GLP and guidelines on things like animal pathology, cell counting, etc., regulatory studies can be conducted by industry that are well done and have minimal bias. However, the analyses and interpretation of these studies should be done by the regulatory scientists to avoid biases entering into the evaluation of these data; it is easier for a corporation to bias an analysis and the interpretation of the data than it is to bias the actual raw study results. If all the agency does is to read a report produced by industry without doing their own analyses, any bias toward null findings remains.

A better solution than the agencies contracting to have the studies done is for the agencies to receive these studies electronically from the contract laboratory in a format that allows the agency to analyse and evaluate the data. For transparency purposes, all of these data should also be openly available in the same electronic format for others to evaluate. Even if the agency has to contract out the statistical analysis of the data, strong conflict of interest rules could ensure that the contracting organization was independent of industry influence.

As best I can tell, the systematic identification of studies to be included in an evaluation is done by industry and not by the regulatory authority. This can also lead to bias in the evaluation and the potential for carelessness if agency scientists do not go back and read the original manuscripts. There are numerous contract firms which will scan the literature and perform a systematic review and provide the agency with the results; this is how the National Toxicology Program starts reviews for the Report on Carcinogens and it is how the Agency for Toxic Substances and Disease Registry starts reviews for their ToxProfiles. Assuming the agency is careful in their choice of a contract firm that has no ties to the regulated industry, this is a better solution to obtaining the literature as it excludes the regulated industry from the overall search. The initial draft of the risk

assessment should also be developed by the agency scientists independent of any summaries provided by industry.

From a practical perspective, having run the US National Toxicology Program, I know how many resources it takes to adequately address the complexities of managing an experimental study at a contract laboratory and unless the EU is prepared to spend hundreds of millions of euros to achieve the goal of all studies funded and managed through their agencies, the result would end up being lower quality studies. Even if the industry was forced to pay for the studies, there would be considerable resources needed to manage the overall process. Another practical objection to an approach working only through the regulatory authority is the setting of priorities; who's study gets done first. Businesses set their own priorities based upon expected gain whereas agencies would have to develop their own priorities that could slow down the development and approval of truly beneficial products.

There is one case where it is impractical to expect industry-sponsored studies and that is the case of multiple exposures. Pesticides are never used in isolation and there may be cases where a regulator is concerned about people being exposed to two compounds simultaneously. If these products are owned by different companies, these companies have no reason to cooperate on doing a joint study. In this case, having resources within an agency to contract out such a study would be an advantage.

Finally, it might be useful to consider the development of a European National Toxicology Program, tasked with conducting toxicological evaluations of important substances (possibly through contract laboratories as is the case for most US NTP studies). The placement of such a program within the European science community would be critical to the success of the program. In the US, when we were developing the US NTP, we decided to place it within the National Institutes of Health rather than any regulatory agency. This insured the science would be independent of regulatory pressures and the results would be equally shared across all regulatory agencies. The results are clear since the US NTP studies and their evaluations are considered the "gold standard" in toxicology. I would be happy to expand more on the develop, management and quality assurance at the US NTP.

66. Could you explain how you apply the weight of evidence approach in the context of your assessments? Do you see ways to improve the communication to strengthen trust in the reliability of your assessment in this regard?

Systematic review is an approach that is gaining regulatory acceptance that has some advantages over the current process. In systematic review, studies are evaluated in a specific manner that is clearly expressed. Weights are assigned to each study characteristic and the final weight can be used to determine if a study is included or excluded from a final evaluation. The expansion of the use of meta-analyses and pooled analyses beyond epidemiology into toxicology will provide a more objective evaluation of the available data with the weights appropriately described in the meta- or pooled-analysis.