February 15, 2011

TO: Fran Kammerer
Office of Environmental Health Hazard Assessment (OEHHA)
California Environmental Protection Agency
1101 I Street, 23rd Floor MS-25B
Sacramento, CA 95812
FKammerer@oehha.ca.gov


These comments are submitted on behalf of the Physicians Committee for Responsible Medicine, with the support of 25,000 physicians and laypersons in California and 125,000 nationwide. PCRM advocates for the development and use of human-relevant, non-animal toxicology testing methods and regulatory policies and procedures that increase their use, while reducing and replacing the use of animal tests. Thank you for this opportunity to comment on the draft Toxics Information Clearinghouse Regulation.

I. General Comments

We recognize that California EPA (CalEPA) cannot, through this regulation, require testing, and that it “imposes no requirements on any person or business since it only identifies hazard traits...that DTSC [Department of Toxic Substances Control] will use in its development of the Toxics Information Clearinghouse [TIC].” However, the DTSC must recognize that the presence of perceived “data gaps” in the TIC, as well as “suggestive evidence” a substance may cause a certain hazard, may both lead to an increase in animal tests.

On the other hand, the explicit use of evidence from ex vivo, in vitro, and non-animal tests is welcome as forward-looking policy. Regulatory agencies from Health Canada to the European Chemicals Agency are investigating ways in which they can use this information to improve substance assessments and we applaud CalEPA for its leadership in this area.

The Initial Statement of Reasons (ISR) clearly articulates the rationale behind the use of “strong” and “suggestive” evidence within the regulation, and these distinctions make sense--consideration of all available evidence ensures a scientifically robust
assessment and can decrease the use of animal tests. This is the regulatory equivalent of “If it walks like a duck, quacks like a duck, and looks like a duck, it must be a duck.” However, the regulation has three deficiencies related to this approach.

First, this regulation designates animal tests as the “strong” evidence gold standard. Iconic examples such as the link between smoking and lung cancer illustrate the danger of this approach—cigarette manufacturers were able to delay regulation of cigarettes for years using “clean” animal tests.\(^1\) Substances that have no effects in a select animal “model” are not necessarily safe, and vice versa. The regulation also consistently places (Q)SAR and in vitro approaches into the “other evidence/suggestive evidence” category (with the exception of genotoxicity), implying that this evidence should be weighed less heavily than evidence from animal tests. In fact, models are more easily accepted for certain hazard endpoints over others, and this should be reflected in the regulation.\(^2\) For example, (Q)SAR models have been used for decades by the US EPA to estimate acute fish toxicity. The results of certain models are considered more acceptable than others. Also within the (Q)SAR field, applicability domain matters. In the past, users of (Q)SAR models did not pay enough attention to the kinds of substances the model was designed to be used for, and so the predictions they obtained were poor. All models—including animal tests—have certain limitations and these limitations should be transparent within the TIC.

Second, while the regulation appropriately recognizes the usefulness of the “toxicity pathway” approach suggested by the National Academy of Sciences\(^3\), it fails to explicitly recognize that some toxicity pathways are known and well-understood, and so can count as “strong” evidence a substances possesses a certain hazard trait without “definitive” in vivo information. Examples include some listed in this very ISR, such as genotoxicity, metabolism to carcinogenic substances (pg. 28), and iodide-uptake inhibition (pg. 37); and others explored in the literature.\(^4,5\)

Third, we find no explicit discussion of how all the available evidence will or should be considered together, whether conflicting or corroborative. The importance of a defined process for considering the weight of evidence (WoE) cannot be overstated, for the benefit of CalEPA, the public, and manufacturers. It is implied that lines of evidence falling in the “strong evidence” categories should be weighed more heavily than evidence falling into the “suggestive evidence” categories. Can multiple lines of “suggestive” evidence combine to form “strong” evidence? Can mechanistic evidence

---


4 Thomas et al 2010.

5 another
alone ever be “strong” evidence? What if “supportive” evidence conflicts with a test in the “strong” evidence category? Should evidence from humans, including from human cell-based assays, be weighed more heavily than evidence from animal tests? At what point does a preponderance of “suggestive” evidence overcome a lack of “strong” evidence?

We do not intend to imply that a WoE should be calculated by counting the number of positive and negative studies, which could lead to an arms-race style testing exercise; wasting time, money, and animal lives. Rather, there is a need to provide more guidance on how one or more lines of evidence should be weighed and interpreted in the context of the hazard trait framework. The “suggestive/strong” framework appropriately allows information from tests or endpoints that might not be “validated” in the traditional sense, but can provide useful information. There is a concern, however, that evidence presented in the TIC that suggests a particular hazard endpoint may lead a producer to conduct an animal test (to “prove” the substance isn’t hazardous) that isn’t warranted because the suggestive evidence is obtained from a test with a high false positive rate (i.e. a positive result is not a true positive).

Tests with high false negative rates are less of a concern here because the TIC simply lists all available information. It doesn’t--and shouldn’t--imply that absence of a definitive test for a particular hazard endpoint is a data gap to be filled. Indeed, some of the hazard traits listed in the regulation (e.g. cardiotoxicity, ototoxicity, hematotoxicity) and some of the lines of evidence for those hazard traits are not included in traditional hazard assessment tests. Expanding the “-icities” commonly tested for in animal studies may not be a goal, but could be an unintended consequence of this regulation.

Related to this point is that the regulation does not address the concept of prioritization of effect. That is, some substances are known to have particular effects on certain organ systems that lead to regulatory limits, like corrosive dermatotoxicity, genotoxicity, or neurotoxicity. One way in which regulatory agencies like the US EPA are planning to use molecular and genomic information from high-throughput \textit{in vitro} assays, at least initially, is to discover clues about a substance’s potential hazards, and to focus more involved \textit{in vivo} testing towards those potential hazards.\textsuperscript{6} For example, if a suite of \textit{in vitro} tests indicates a substances disrupts cellular process related to carcinogenicity, but none related to neurotoxicity, subsequent tests would focus on carcinogenicity and not neurotoxicity.

Central to this concept is recognition that while absence of evidence of a toxic effect does not equal evidence of absence, absence of evidence is not necessarily a data gap to be filled. Checking all of the “boxes” isn’t productive, necessary, or even possible. The backlog of existing chemicals, and the pace of creation of new chemicals and products (including green chemicals) cannot be matched by conducting low-throughput

\textsuperscript{6} \url{http://www.epa.gov/pesticides/science/testing-assessment.html}
animal tests—nor do consumers in California support such an endeavor. The solution is prioritization of substances and effects using quicker tools and approaches, and WoE determinations of existing information. In fact, the accuracy of emerging tools can sometimes be increased by using them in combination. OEHHA clearly understands this reality; however the regulation does not explicitly recognize or discuss these concepts.

These deficiencies are easily rectified by providing interpretation guidelines within the regulation and, perhaps more importantly, within the TIC itself. For all of the reasons discussed above, the interpretation of the results presented (and not presented) in the TIC is important in order to ensure that the TIC is a useful tool for the public, CalEPA, and manufacturers alike. While we recognize that the reliance on assessments from “authoritative organizations” within the regulation and the TIC will provide some interpretation and discussion, this may not be true for all endpoints.

These interpretation guidelines should also discuss the other “side” of risk, i.e. the exposure context. Even though hazard classification schemes like the European CLP do not consider exposure when determining the hazard classification of a substance, this information is still considered and interpreted within the use and potential human or environmental exposure of the substances, and the same must be true here in California.

An endpoint OEHHA should consider introducing into the regulation is a “structural alert,” which refers to particular moieties of a substance’s structure that indicate the potential for a toxic endpoint. While these concepts are relatively new, it reflects a refinement in the state of an evolving science and should be referred to here. The process for updating terms and lines of evidence is not addressed in the ISR or regulation. How often does OEHHA envision updating the regulation to account for advances in testing methods, and, more importantly, the place that those methods and the lines of evidence they produce have in the regulatory context?

II. Specific Comments

In most cases, the use of in vitro or in silico evidence is relegated to “suggestive” or “other” evidence; sometimes this is appropriate. For other hazard endpoints, however, such as with neurotoxicity, inhalation toxicity, certain carcinogenicity mechanisms, or skin sensitization, the only contribution an animal test would provide is metabolism information, which can often be obtained by other in vitro or in silico tests or models.

Initial Statement of Reasons

Notes:

8 Sedykh et al. Environ Health Perspect doi:10.1289/ehp.1002476
Page 15: The use of mechanistic similarity to classify substances is an excellent aspect of this regulation.

Page 28: The statement “Mechanistic evidence alone can provide strong evidence of carcinogenicity,” is part of what we are looking for above. However, this is not evident from the regulation text. Will the ISR be available with the regulation?

Page 37: The paragraph quoting the National Academy of Sciences report about perchlorate and iodide-uptake inhibition is an excellent, clear example of the use of mechanistic data and high-throughput assays; this seems to belong in a more general section and not in the genotoxicity section.

Page 38: While it is true that one mechanism for a substance to cause developmental toxicity or carcinogenicity is gene mutation, not all substances testing positive in the Salmonella assay should be tested in the 2-year cancer bioassay or developmental toxicity test. This example illustrates the importance of discussions of WoE and test attributes.

Page 48: While it is true that “[Q(SAR)] models can be based purely upon statistical correlations”, this is not always the case.

Page 56: There are a host of in vitro tests covering skin sensitization that can be mentioned here and in the regulation,10,11,12,13 including those under the EU Sens-it-iv project,14 which breaks down the steps in the process from exposure to the sensitization reaction, and has a “tool” or test for each step. Taken together, results from these tests could be considered strong evidence of dermal sensitization.

Page 58: The OECD has approved final versions of skin irritation and corrosion test guidelines, which are available here as TGs 430, 431, 435, and 439: http://www.oecd-ilibrary.org/environment/oecd-guidelines-for-the-testing-of-chemicals-section-4-health-effects_20745788.

Page 63: The genotoxicity assays listed here all have varying levels of positive and negative predictivity and this should be recognized to the extent possible.

Page 72: This entire discussion of in vitro tests that provide strong evidence of an immunotoxic effect is excellent.

14 http://www.sens-it-iv.eu/
Page 83: There are several draft and adopted OECD test guidelines that can be referenced here related to eye corrosion and irritation; these can be found here: http://www.oecd-ilibrary.org/environment/oecd-guidelines-for-the-testing-of-chemicals-section-4-health-effects_20745788 and here: http://www.oecd.org/document/55/0,3746,en_2649_34377_2349687_1_1_1_1,00.html.

Page 103: The US EPA has developed a (Q)SAR model that can be used to predict binding of pesticide inerts to the fish estrogen receptor. This can be found, along with all of the other EPA-developed QSARs, in the OECD (Q)SAR Application Toolbox, described here: http://www.oecd.org/document/28/0,3746,en_2649_33713_45310876_1_1_1_1,00.html.

Proposed Regulation

Article 2, section 69402.1

In some regulatory agencies, positive in vivo genotoxicity is enough to classify a substance as carcinogenic. Additionally, there is emerging evidence that for some pharmaceuticals and chemicals, tissue histopathology and blood-based biomarkers collected during short-term studies conclusively predict a carcinogenicity classification.

Article 2, section 69403.2

Within dermatotoxicity, in vitro skin irritation—not just corrosively—is adopted and should be considered “strong evidence” (see OECD TGs above).

Article 2, section 69403.10

In the ISR (pg. 77), in vitro indicators of nephrotoxicity are considered toxicity endpoints, not other relevant data.

Article 2, section 69403.12

In vitro and ex vivo models, such as the Bovine Corneal Opacity and Permeability assay (BCOP) can illustrate many of the adverse changes discussed in this section.

Article 2, section 69403.14

It is unclear why in vitro indicators of reactivity in biological systems would be considered only “other relevant” data; detoxification and metabolism attributes of substances can be discovered and/or confirmed using other lines of evidence.

15 http://www.slidefinder.net/a/approach_risk_assessment_genotoxic_carcinogens/9228021

Article 2, section 69403.15

While *in vitro* evidence is listed under “other relevant data,” in fact many of the endpoints in (b), including airway remodeling, inflammation, fibrosis, respiratory irritation, and tissue damage can be determined using *in vitro and ex vivo* models. These include single cells as well as multi-cell 3D models and complicated air/liquid exposure systems. Some commercial companies producing such models include Epithelix, MatTek, and Biopta; many more models are used in the research setting.

Article 2, section 69404.9

The *in vitro* Fish Embryo Toxicity assay\(^{17}\) has been in use for several years by Germany to test effluent toxicity levels, and should be considered as evidence of wildlife survival impairment; several (Q)SARs, mentioned above, can also be used here for certain substance classes.

Article 5: Exposure Potential hazard traits: We suggest adding biologically-based exposure potential to this list, including the potential for a substance to penetrate the blood-brain barrier, the skin, or the mucosal linings of the gut or airways. These traits are also intrinsic traits that impact whether or to what degree a substance might interact with organisms to create a toxic effect.

Thank you for your attention to these comments. We can be reached using the information below with any questions or concerns.

Kristie Sullivan, MPH
Physicians Committee for Responsible Medicine
Ksullivan@pcrm.org
1-510-923-9446

\(^{17}\) Braunbeck et al. 2005 ALTEX. 22(2):87-102.